22nd International Conference on Formal Power Series and Algebraic Combinatorics

22e Conférence Internationale sur les Séries Formelles et la Combinatoire Algébrique

San Francisco State University

August 2–6, 2010

2–6 août 2010

1365–8050 © 2010
DMTCS, Nancy, France

Discrete Mathematics and Theoretical Computer Science
http://www.dmtcs.org/dmtcs-ojs/index.php/proceedings/issue/archive

Program Committee – Comité de programme

Sara Billey (University of Washington, USA, co-chair)
Vic Reiner (University of Minnesota, USA, co-chair)
Michael Albert (University of Otago, New Zealand)
Francesco Brenti (Universita' di Roma "Tor Vergata", Italy)
Sylvie Corteel (Universit Paris XI, France)
Sergi Elizalde (Dartmouth College, USA)
Dmitry Feichtner-Kozlov (Universitt Bremen, Germany)
Ira Gessel (Brandeis University, USA)
Curtis Greene (Haverford College, USA)
Jim Haglund (University of Pennsilvania, USA)
Sam Hsiao (Bard College, USA)
Jakob Jonsson (KTH, Sweden)
Gil Kalai (Hebrew University, Israel)
Carly Klivans (University of Chicago, USA)
Fu Liu (University of California, Davis, USA)
Soichi Okada (Nagoya University, Japan)
Julian Pfeifle (Universitat Politcnica de Catalunya, Spain)
Alexander Postnikov (Massachusetts Institute of Technology, USA)
David Speyer (Massachusetts Institute of Technology, USA)
Volkmar Welker (Philipps-Universitt Marburg, Germany)
Alex Yong (University of Illinois at Urbana-Champaign, USA)
Mike Zabrocki (York University, Canada)

Organization Committee – Comité d'organisation

Federico Ardila (San Francisco State University, USA, co-chair)
Matthias Beck (San Francisco State University, USA, co-chair)
Nantel Bergeron (York University, Canada)
Thomas Bliem (San Francisco State University, USA and Universitt zu Kln, Germany)
Tristram Bogart (San Francisco State University, USA and Queen's University, Canada)
Serkan Hosten (San Francisco State University, USA)
Brant Jones (University of California, Davis, USA)
Fu Liu (University of California, Davis, USA)
Peter McNamara (Bucknell University, USA)
Ellen Veomett (California State University, East Bay, USA)
Mike Zabrocki (York University, Canada)

Proceedings prepared by Peter McNamara with help from translators Anouk Bergeron-Brlek, Emmanuel Briand, Emily Dryden, Alain Goupil, Matthieu Josuat-Vergès, Pierre-Loïc Méliot and Rosa Sena-Dias

We were very pleased to receive a record number of submissions this year covering the spectrum of combinatorics and its applications. The presentations will cover many exciting new results and build connections between research areas. We hope you will be able to attend all the lectures and the two poster sessions. We encourage you to ask questions, discuss the material during breaks and participate fully in the FPSAC experience.

Our thanks goes out to everyone attending FPSAC 2010, and especially the members of the Program and Organizing Committees. We also thank the following organizations for their financial support: the National Science Foundation, the National Security Agency, the San Francisco State University College of Science and Engineering and Department of Mathematics, the Fields Institute, Elsevier, and Lindo Systems. Finally, we hope everyone will join us in thanking Federico Ardila and Matthias Beck for taking on the huge job of co-chairing the Organizing Committee.

Sara Billey & Vic Reiner
Program Committee co-chairs

Department of Mathematics
College of Science & Engineering

Table of Contents – Table des matières

Invited Speakers – Conférenciers invités

Talks – Exposés

Posters – Affiches

Part I
Invited Speakers – Conférenciers invités

Recent developments on log-concavity and q-log-concavity of combinatorial polynomials

Bill Chen

Nankai University, China

Abstract. In this talk, I wish to report some recent work with my students and colleagues at Nankai University on log-concavity and q-log-concavity of combinatorial polynomials. While this is a classical subject of algebraic combinatorics, interesting problems and techniques continue to emerge.
(1) We proved the unimodality conjecture on balanced colorings of the n-cube proposed by Palmer, Read and Robinson, and obtained a log-concavity theorem for sufficiently large n.
(2) We proved the ratio monotone property of the Boros-Moll polynomials which is stronger than the log-concavity. We further proved the 2-log-concavity which was considered as a difficult problem. The 2-log-convexity of the Apery numbers has been established. We obtained the reverse ultra log-concavity of the Boros-Moll polynomials, and confirmed Moll's minimum conjecture. A combinatorial approach has been found to justify the log-concavity and other properties of the Boros-Moll polynomials.
(3) By using the Littlewood-Richardson rule, we obtained certain Schur positivity results that lead to a proof of the q-log-convexity conjecture for the Narayana polynomials.
(4) By establishing the strong q-log-concavity of q-Narayana numbers $N_q(n,k)$ for fixed k, we confirmed the 2-fold case of the q-log-concavity conjecture for the Gaussian coefficients proposed by McNamara and Sagan.
(5) We found a unified approach to the q-log-convexity of the Bell polynomials, the Bessel polynomials, the Ramanujan polynomials and the Dowling polynomials.
(6) We shall also mention some open problems.

A probabalistic interpretation of the Macdonald polynomials

Persi Diaconis

Stanford University, USA

Abstract. The two-parameter Macdonald polynomials are a central object of study in algebraic combinatorics. Arun Ram and I have found a simple Markov chain on partitions (with stationary distributon the Macdonald weight) whose eigenfunctions are the co-efficients of the Macdonald polynomials when expanded in the power sums. In turn, this Markov chain is a special case of classical algorithms in statistical mechanics (Swedsen-Wang and auxiliary variable algorithms). The wealth of knowledge about Macdonald polynomials allows a sharp analysis of the rate of convergence of the Markov chain to stationarity.

Hypergeometric series with algebro-geometric dressing

Alicia Dickenstein

Universidad de Buenos Aires, Argentina

Abstract. Important classical functions as well as generating functions associated to combinatorial objects are given by series whose coefficients satisfy hypergeometric recurrences. Following Gelfand, Kapranov and Zelevinsky, who defined A-hypergeometric systems satisfied by suitable homogeneous versions of classical hypergeometric functions, we will present joint work with Eduardo Cattani, Laura Matusevich, Fernando Rodrguez Villegas, Timur Sadykov and Bernd Sturmfels, to describe the structure of rational hypergeometric series in two variables and the ocurrence of finite solutions to hypergeometric recurrences.

Normal polytopes

Joseph Gubeladze †

Department of Mathematics, San Francisco State University, San Francisco, CA 94132, USA

Abstract. In Section 1 we overview combinatorial results on normal polytopes, old and new. These polytopes represent central objects of study in the contemporary discrete convex geometry, on the crossroads of combinatorics, commutative algebra, and algebraic geometry. In Sections 2 and 3 we describe two very different possible ways of advancing the theory of normal polytopes to next essential level, involving arithmetic and topological aspects.

Keywords: lattice polytopes, normal polytopes, tight polytopes, unimodular cover, integral Carathéodory property

1 Normal polytopes: old and new

All our polytopes are assumed to be convex.

Let $P \subset \mathbb{R}^d$ be a lattice polytope and denote by L the affine lattice in $\mathbb{Z}^d$, generated by the lattice points in P; i. e., $L = v + \sum_{x,y \in P \cap \mathbb{Z}^d} \mathbb{Z}(x-y) \subset \mathbb{Z}^d$, where v is some (equivalently, any) lattice point in P. Observe, $P \cap L = P \cap \mathbb{Z}^d$. Here is the central definition:

(a) P is *integrally closed* if the following condition is satisfied:

$$c \in \mathbb{N},\ z \in cP \cap \mathbb{Z}^d \quad \Longrightarrow \quad \exists x_1, \dots, x_c \in P \cap \mathbb{Z}^d \quad x_1 + \cdots + x_c = z.$$

(b) P is *normal* if the following condition is satisfied:

$$c \in \mathbb{N},\ z \in cP \cap L \quad \Longrightarrow \quad \exists x_1, \dots, x_c \in P \cap L \quad x_1 + \cdots + x_c = z.$$

The normality property is invariant under affine-lattice isomorphisms of lattice polytopes, and the property of being integrally closed is invariant under an affine change of coordinates, leaving the lattice structure $\mathbb{Z}^d \subset \mathbb{R}^d$ invariant.

A lattice polytope $P \subset \mathbb{R}^d$ is integrally closed if and only if it is normal and L is a direct summand of $\mathbb{Z}^d$. Obvious examples of normal but not integrally closed polytopes are the s. c. empty lattice simplices of large volume. No classification of such simplices is known in dimensions ≥ 5, the main difficulty being the lack of satisfactory characterization of their lattice widths; see [13, 20].

A normal polytope $P \subset \mathbb{R}^d$ can be made into a full-dimensional integrally closed polytope by changing the lattice of reference $\mathbb{Z}^d$ to L and the ambient Euclidean space $\mathbb{R}^d$ to the subspace $\mathbb{R}L$. In particular,

†Supported by NSF grant DMS - 1000641

normal and integrally closed polytopes refer to same isomorphism classes of lattice polytopes. In the combinatorial literature the difference between 'normal' and 'integrally closed' is sometimes blurred.

Normal/integrally closed polytopes enjoy popularity in algebraic combinatorics and they have been showcased on recent workshops ([1, 2]). These polytopes represent the homogeneous case of the Hilbert bases of finite positive rational cones and the connection to algebraic geometry is that they define projectively normal embeddings of toric varieties. There are many challenges of number theoretic, ring theoretic, homological, and K-theoretic nature, concerning the associated objects: Ehrhart series', rational cones, toric rings, and toric varieties; see [7].

If a lattice polytope is covered by (in particular, subdivided into) integrally closed polytopes then it is integrally closed as well. The simplest integrally closed polytopes one can think of are *unimodular simplices*, i. e., the lattice simplices $\Delta = conv(x_1, \dots, x_k) \subset \mathbb{R}^d$, $\dim \Delta = k-1$, with $x_1 - x_j, \dots, x_{j-1} - x_j, x_{j+1} - x_j, \dots, x_k - x_j$ a part of a basis of $\mathbb{Z}^d$ for some (equivalently, every) j.

Unimodular simplices are the smallest 'atoms' in the world of normal polytopes. But the latter is not built out exclusively of these atoms: not all 4-dimensional integrally closed polytopes are triangulated into unimodular simplices [9, Prop. 1.2.4], and not all 5-dimensional integrally closed polytopes are covered by unimodular simplices [5] – contrary to what had been conjectured before [19]. Further 'negative' results, such as [4] and [8] (the latter disproving a conjecture from [10]), contributed to the current thinking in the area that there is no succinct geometric characterization of the normality property. One could even conjecture that in higher dimensions the situation gets as bad as it can; see Section 2 for details.

'Positive' results in the field mostly concern special classes of lattice polytopes that are normal, or have unimodular triangulations or unimodular covers. Knudsen-Mumford's classical theorem ([7, Sect. 3B], [14, Chap. III],) says that every lattice polytope P has a multiple cP for some $c \in \mathbb{N}$ that is triangulated into unimodular simplices. The existence of a dimensionally uniform lower bound for such c seems to be a very hard problem. More recently, it was shown in [6] that there exists a dimensionally uniform exponential lower bound for unimodularly covered multiple polytopes. By improving one crucial step in [6], von Thaden was able to cut down the bound to a degree 6 polynomial function in the dimension [7, Sect. 3C], [22].

The results above on multiple polytopes yield no new examples of normal polytopes, though. In fact, an easy argument ensures that for any lattice d-polytope P the multiples cP, $c \geq d-1$, are integrally closed [9, Prop. 1.3.3], [11]. One should remark that there is no algebraic obstruction to the existence of (even quadratic regular) unimodular triangulations for all multiples cP, $c \geq d-1$: the nice homological properties that the corresponding toric rings would have (according to Sturmfels' theory [21]) should such triangulations existed, are all present [9].

Lattice polytopes with long edges of independent length are considered in [12], where it is shown that if the edges of a lattice d-polytope P have lattice lengths $\geq 2d^2(d+1)$ then P is integrally closed. In the special case when P is a simplex one can do better: P is covered by lattice parallelepipeds, provided the edges of P have lattice lengths $\geq d(d+1)$. Lattice parallelepipeds are the simplest integrally closed polytopes after unimodular simplices.

Currently, one problem attracts much attention in the field. Namely, Oda's question asks whether all smooth polytopes are integrally closed. A lattice polytope $P \subset \mathbb{R}^d$ is called smooth if the primitive edge vectors at every vertex of P define a part of a basis of $\mathbb{Z}^d$. Smooth polytopes correspond to projective embeddings of smooth projective toric varieties. Oda's question remains (as of writing this) wide open – so far no smooth polytope just without a unimodular triangulation has been found. The research was triggered by a faulty attempt in the mid 1990s to answer the question in the positive. The Frobenius splitting techniques that was used then was recently revived in Payne's work, leading to new classes of normal (and even Koszul) polytopes, associated to root systems of various types [17].

For further classes of polytopes, associated to root systems and admitting unimodulaar triangulations as certificate of normality, see [3, 15, 16, 18].

In the next two sections we describe two very different possible ways of advancing the theory of normal polytopes to next essential level.

2 Carathéodory rank

What follows next is the homogeneous special case of a more general story that concerns rational positive cones and their Hilbert bases.

An arithmetic version of a unimodular cover is the *integral Carathéodory property (ICP):* a lattice d-polytope P has (ICP) if for every natural number c and every lattice point $z \in cP$ there exist lattice points $x_1, \dots, x_{d+1} \in P$ and integers $a_1, \dots, a_{d+1} \geq 0$ such that $z = a_1x_1 + \cdots + a_{d+1}x_{d+1}$ and $a_1 + \cdots + a_{d+1} = c$.

For a lattice polytope P its *Carathéodory rank,* denoted by $CR(P)$, is the smallest natural number k such that for every natural number c and every lattice point $z \in cP$ there exist lattice points $x_1, \dots, x_k \in P$ and integers $a_1, \dots, a_k \geq 0$ such that $z = a_1x_1 + \cdots + a_kx_k$ and $a_1 + \cdots + a_k = c$.

Sebö has shown [19] that $CR(P) \leq 2 \dim P$ for arbitrary integrally closed polytope P. If P has a unimodular cover then it has (ICP) too, i. e., $CR(P) = \dim P + 1$. It is known that (ICP) implies 'integrally closed' [5]. The converse is not true: there are integrally closed 5-polytopes without (ICP) [8]. That (ICP) and the existence of a unimodular cover are different conditions was discovered only recently: there are 5-polytopes with (ICP) but without unimodular cover [4].

We conjecture that Sëbo's estimate for Carathéodory rank is asymptotically sharp:

$$\lim_{d\to\infty} d^{-1} \max CR(P) = 2,$$

where, for each fixed d, P runs over the integrally closed d-polytopes. This conjecture, in particular, says that there are essentially new types of counterexamples to (ICP) in higher dimensions, not obtained by trivial extensions of counterexamples in lower dimensions.

3 Do normal polytopes model quantum states?

The method, by which counterexamples to the unimodular cover property and (ICP) were found, was to check s. c. *tight polytopes* for these properties. An integrally closed polytope $P \subset \mathbb{R}^d$ is called tight if, whatever lattice point $x \in P \cap \mathbb{Z}^d$ we choose, the convex hull of $\left(P \cap \mathbb{Z}^d\right) \setminus \{x\}$ in $\mathbb{R}^d$ is not an integrally closed polytope. This moves center stage the descending sequences of lattice integrally closed polytopes in $\mathbb{R}^d$ of type

$$P_1 \supset P_2 \supset \cdots \supset P_k, \qquad \# \left(P_i \cap \mathbb{Z}^d\right) = \# \left(P_{i+1} \cap \mathbb{Z}^d\right) + 1, i = 1, \dots, k-1.$$

That such a sequence may halt at all at a positive dimensional polytope, or equivalently, that there exist nontrivial tight polytopes is already something not quite obvious. This phenomenon shows up in dimensions ≥ 4. There are no tight polygons, and the existence of tight 3-polytopes is not known.

Consider the poset $\mathsf{Pol}(d)$ of lattice integrally closed polytopes in $\mathbb{R}^d$, where the order relation is generated by the elementary relations of type $P < Q$, $\#\left(Q \cap \mathbb{Z}^d\right) = \#\left(P \cap \mathbb{Z}^d\right) + 1$. Minimal elements of $\mathsf{Pol}(d)$ are exactly the tight polytopes in $\mathbb{R}^d$. Informally, the poset $\mathsf{Pol}(d)$ offers a global picture of the interaction of polytopal shapes in $\mathbb{R}^d$ with the integer lattice $\mathbb{Z}^d$.

Here is a list of a several interesting questions one can ask about $\mathsf{Pol}(d)$: Do there exist maximal elements in $\mathsf{Pol}(d)$? What is the homotopy type of $\mathsf{Pol}(d)$? Is it contractible? Does it have isolated points?

We suggest the following game: think of the chains in $\mathsf{Pol}(d)$ as quantum processes, individual polytopes as quantum states, and elementary relations $P < Q$ as quantum jumps. The vocabulary can be extended to accommodate such terminology as energy and time (both discrete), observables, entanglement, tunneling, uncertainty principle, fluctuations. Real fun starts when one thinks of the potentially nontrivial homotopy groups of $\mathsf{Pol}(d)$ as a force that permeates all of the geometric realization space of $\mathsf{Pol}(d)$ and keeps the world of integrally closed polytopes from collapsing into a point.

References

[1] Workshop: Combinatorial Challenges in Toric Varieties. The workshop held April 27 to May 1, 2009, Organized by Joseph Gubeladze, Christian Haase, and Diane Maclagan, American Institute of Mathematics, Palo Alto. http://www.aimath.org/pastworkshops/toricvarieties.html.

[2] Mini-Workshop: Projective Normality of Smooth Toric Varieties. *Oberwolfach Rep.*, 4(3):2283–2319, 2007. Abstracts from the mini-workshop held August 12–18, 2007, Organized by Christian Haase, Takayuki Hibi and Diane Maclagan, Oberwolfach Reports. Vol. 4, no. 3.

[3] Matthias Beck and Serkan Hoşten. Cyclotomic polytopes and growth series of cyclotomic lattices. *Math. Res. Lett.*, 13:607–622, 2006.

[4] Winfried Bruns. On the integral Carathéodory property. *Experiment. Math.*, 16:359–365, 2007.

[5] Winfried Bruns and Joseph Gubeladze. Normality and covering properties of affine semigroups. *J. Reine Angew. Math.*, 510:161–178, 1999.

[6] Winfried Bruns and Joseph Gubeladze. Unimodular covers of multiples of polytopes. *Doc. Math.*, 7:463–480, 2002.

[7] Winfried Bruns and Joseph Gubeladze. *Polytopes, rings, and K-theory*. Springer Monographs in Mathematics. Springer, Dordrecht, 2009.

[8] Winfried Bruns, Joseph Gubeladze, Martin Henk, Alexander Martin, and Robert Weismantel. A counterexample to an integer analogue of Carathéodory's theorem. *J. Reine Angew. Math.*, 510:179–185, 1999.

[9] Winfried Bruns, Joseph Gubeladze, and Ngô Viêt Trung. Normal polytopes, triangulations, and Koszul algebras. *J. Reine Angew. Math.*, 485:123–160, 1997.

[10] W. Cook, J. Fonlupt, and A. Schrijver. An integer analogue of Carathéodory's theorem. *J. Combin. Theory Ser. B*, 40:63–70, 1986.

[11] Günter Ewald and Uwe Wessels. On the ampleness of invertible sheaves in complete projective toric varieties. *Results Math.*, 19:275–278, 1991.

[12] Joseph Gubeladze. Convex normality of rational polytopes with long edges. preprint.

[13] Christian Haase and Günter M. Ziegler. On the maximal width of empty lattice simplices. *European J. Combin.*, 21:111–119, 2000. Combinatorics of polytopes.

[14] G. Kempf, Finn Faye Knudsen, D. Mumford, and B. Saint-Donat. *Toroidal embeddings. I.* Lecture Notes in Mathematics, Vol. 339. Springer-Verlag, Berlin, 1973.

[15] Tomonori Kitamura. Gröbner bases associated with positive roots and Catalan numbers. *Osaka J. Math.*, 42(2):421–433, 2005.

[16] Hidefumi Ohsugi and Takayuki Hibi. Unimodular triangulations and coverings of configurations arising from root systems. *J. Algebraic Combin.*, 14(3):199–219, 2001.

[17] Sam Payne. Lattice polytopes cut out by root systems and the Koszul property. *Adv. Math.*, 220:926–935, 2009.

[18] Victor Reiner and Volkmar Welker. On the Charney-Davis and Neggers-Stanley conjectures. *J. Combin. Theory Ser. A*, 109(2):247–280, 2005.

[19] András Sebö. Hilbert bases, Carathéodory's theorem and combinatorial optimization. *Proc. of the IPCO conference (Waterloo, Canada)*, pages 431–455, 1990.

[20] András Sebő. An introduction to empty lattice simplices. In *Integer programming and combinatorial optimization (Graz, 1999)*, volume 1610 of *Lecture Notes in Comput. Sci.*, pages 400–414. Springer, Berlin, 1999.

[21] Bernd Sturmfels. Gröbner bases of toric varieties. *Tohoku Math. J. (2)*, 43(2):249–261, 1991.

[22] Michael von Thaden. Improved bounds for covers and triangulations of lattice polytopes. Preprint.

Combinatorial representation theory of algebras: the example of j-trivial monoids

Florent Hivert

Université de Rouen, France

Abstract. The representation theory of algebras is a very important source of interesting combinatorics. The symmetric groups leading to the combinatorics of tableaux is certainly the most striking example, but it is far from being unique: other examples include the various Hecke algebras, descent algebras... Another interesting feature of the representation theory of the finite dimensional algebras is that it is mostly effective. As a consequence, with the appropriate tools one can very easily use computers for exploration.

The goal of the talk is to discuss these features together with the simple remark that several recently studied algebras are in fact monoid algebras: examples are 0-Hecke algebras, degenerated Ariki-Koiki algebras, Solomon-Tits algebras. Apparently, the fact that they are indeed monoid algebras wasn't used in those studies. However, from recent results in semigroups theory it seem that a lot of representation theory of a semigroup algebra is of combinatorial nature (provided the representation theory of groups is known). One can expect, for example, that for aperiodic semigroup (semigroup which doesn't contains non trivial groups), most of the combinatorial information (dimensions of the simple/projective indecomposable, induction/restriction constants/Cartan's invariants) can be computed without using any linear algebra.

In this talk, we will focus on the so-called J-trivial monoids, which are the monoids M such that the product has the following triangular properties: there exists a partial ordering $\leq$ on M such that for a $x, y \in M$, one had $xy \leq x$ and $xy \leq y$. A typical example is the 0-Hecke monoid of a Coxeter group. We will show that for such a monoid, most of the combinatorial data of the representation theory including the Cartan's invariant matrix and the quiver can be expressed by counting particular kinds of elements in the monoid itself.

This is a joint work with Tom Denton, Anne Schilling and Nicolas Thiéry.

The discrete geometry of moment polytopes

Tara Holm

Cornell University, USA

Abstract. Moment polytopes are convex polytopes associated to Hamiltonian systems in symplectic geometry. The discrete geometry of the moment polytope can tell us a good deal about the global geometry and topology of the corresponding system. I will give a brief introduction to symplectic geometry to show how moment polytopes arise, followed by a survey demonstrating how combinatorial techniques can be brought to bear on topological questions.

Posets and curvature

Jon McCammond

University of California, Santa Barbara, USA

Abstract. It is extremely common to convert a partially ordered set into a simplicial complex and to correlate the topology of the resulting space with the structure of the original poset. In this talk, I discuss how the adding a metric perspective and introducing notions such as curvature from geometric group theory help to enhance this connection. (Joint work with Tom Brady)

Rigidity, sparsity and pebble games

Ileana Streinu

Smith College, USA

Abstract. A famous open problem, going back to the work of James Clerk Maxwell in 1864, is to give a combinatorial characterization for generically rigid frameworks made from bars connected by rotatable joints. The same question can be asked for a long list of other geometrically constrained systems, but only very few answers are known. They include bar-and-joint frameworks in dimensions one and two, body-and-bar structures in arbitrary dimensions and a few other isolated instances such as skeleta of triangulated 3D polyhedra. The underlying combinatorial structure, in all these cases, is a graph satisfying some linear sparsity conditions. The pebble games are simple construction rules characterizing exactly those classes of sparse (hyper)graphs which are matroids. On the other hand, for Maxwell's problem the necessary (but not sufficient) sparsity condition falls "just below" the matroidal range. This talk will present these varied facets of combinatorial rigidity (accompanied by a variety of visual and physical props), and will conclude with some recent results.

The Worm order and its applications

Peter Winkler

Dartmouth College, USA

Abstract. Let x and y be two words in a linearly-ordered alphabet (such as the real numbers). We say that x is below y in the worm order if they can be "scheduled" in such a way that x is always less than or equal to y. It turns out that in any submodular system there is a maximal chain that is minimal in the worm order, among all paths from 0 to 1. One consequence is a set of general conditions under which parallel scheduling can be done without backward steps. Among the applications are a fast algorithm for scheduling multiple processes without overusing a resource; a theorem about searching for a lost child in a forest; and a closed-form expression for the probability of escaping from the origin in a form of coordinate percolation. Joint work in part with Graham Brightwell (LSE) and in part with Lizz Moseman (USMA).

Part II
Talks – Exposés

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 14–25

A Pieri rule for skew shapes

Sami H. Assaf[1 †] and Peter R. W. McNamara[2]

[1] *Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*
[2] *Department of Mathematics, Bucknell University, Lewisburg, PA 17837, USA*
sassaf@math.mit.edu, peter.mcnamara@bucknell.edu

Abstract. The Pieri rule expresses the product of a Schur function and a single row Schur function in terms of Schur functions. We extend the classical Pieri rule by expressing the product of a skew Schur function and a single row Schur function in terms of skew Schur functions. Like the classical rule, our rule involves simple additions of boxes to the original skew shape. Our proof is purely combinatorial and extends the combinatorial proof of the classical case.

Résumé. La règle de Pieri exprime le produit d'une fonction de Schur et de la fonction de Schur d'une seule ligne en termes de fonctions de Schur. Nous étendons la règle classique de Pieri en exprimant le produit d'un fonction gauche de Schur et de la fonction de Schur d'une ligne en termes de fonctions gauches de Schur. Comme la règle classique, notre règle implique l'ajout de cases à la forme gauche initiale. Notre preuve est purement combinatoire et étend celle du cas classique.

Keywords: Pieri rule, skew Schur functions, Robinson-Schensted

1 Introduction

The basis of Schur functions is arguably the most interesting and important basis for the ring of symmetric functions. This is due not just to their elegant combinatorial definition, but more broadly to their connections to other areas of mathematics. For example, they are intimately tied to the cohomology ring of the Grassmannian, and they appear in the representation theory of the symmetric group and of the general and special linear groups.

It is therefore natural to consider the expansion of the product $s_\lambda s_\mu$ of two Schur functions in the basis of Schur functions. The Littlewood-Richardson rule [LR34, Sch77, Tho74, Tho78], which now comes in many different forms ([Sta99] is one starting point), allows us to determine this expansion. However, more basic than the Littlewood-Richardson rule is the Pieri rule, which gives a simple, beautiful and more intuitive answer for the special case when $\mu = (n)$, a partition of length 1. Though we will postpone the preliminary definitions to Section 2 and the statement of the Pieri rule to Section 3, stating the rule in a rough form will give its flavor. For a partition λ and a positive integer n, the Pieri rule states that

† Supported by NSF Postdoctoral Fellowship DMS-0703567

$s_\lambda s_n$ is a sum of Schur functions s_{λ^+}, where λ^+ is obtainable by adding cells to the diagram of λ according to a certain simple rule. The Pieri rule's prevalence is highlighted by its adaptions to many other settings, including Schubert polynomials [LS82, LS07, Man98, Sot96, Win98], Hall-Littlewood polynomials [Mor64], Jack polynomials [Las89, Sta89], LLT polynomials [Lam05], and Macdonald polynomials [Koo88, Mac87, Mac95].

It is therefore surprising that there does not appear to be a known adaption of the Pieri rule to the most well-known generalization of Schur functions, namely skew Schur functions. We fill this gap in the literature with a natural extension of the Pieri rule to the skew setting. Reflecting the simplicity of the classical Pieri rule, the skew Pieri rule states that for a skew shape λ/μ and a positive integer n, $s_{\lambda/\mu}s_n$ is a signed sum of skew Schur functions s_{λ^+/μ^-}, where λ^+/μ^- is obtainable by adding cells to the diagram of λ/μ according to a certain simple rule. Our proof is purely combinatorial, using a sign-reversing involution that reflects the combinatorial proof of the classical Pieri rule.

After reading an early version of the full version [AM09] of this manuscript, which included an algebraic proof of the case $n = 1$ due to Richard Stanley, Thomas Lam provided a complete algebraic proof of our skew Pieri rule, which is included as an appendix to [AM09]. It is natural to ask if our skew Pieri rule can be extended to give a "skew" version of the Littlewood-Richardson rule, and we included such a rule as a conjecture in [AM09]. This conjecture has been proved by Lam, Aaron Lauve and Frank Sottile in [LLS09] using Hopf algebras. It remains an open problem to find a combinatorial proof of the skew Littlewood-Richardson rule.

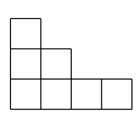
The remainder of this paper is organized as follows. In Section 2, we give the necessary symmetric function background. In Section 3, we state the classical Pieri rule and introduce our skew Pieri rule. In Section 4, we give a variation from [SS90] of the Robinson-Schensted-Knuth algorithm, along with relevant properties. This algorithm is then used in Section 5 to define the sign-reversing involution that proves the skew Pieri rule.

1.1 Acknowledgments

We are grateful to a number of experts for informing us that they too were surprised by the existence of the skew Pieri rule, and particularly to Richard Stanley for providing an algebraic proof of the $n = 1$ case that preceded our combinatorial proof. This research was performed while the second author was visiting MIT, and he thanks the mathematics department for their hospitality. Computations were performed using [Buc99, Ste].

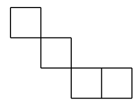
2 Preliminaries

We follow the terminology and notation of [Mac95, Sta99] for partitions and tableaux, except where specified. Letting $\mathbb{N}$ denote the nonnegative integers, a *partition* λ of $n \in \mathbb{N}$ is a weakly decreasing sequence $(\lambda_1, \lambda_2, \ldots \lambda_l)$ of positive integers whose sum is n. It will be convenient to set $\lambda_k = 0$ for $k > l$. We also let $\emptyset$ denote the unique partition with $l = 0$. We will identify λ with its *Young diagram* in "French notation": represent the partition λ by the unit square cells with top-right corners $(i, j) \in \mathbb{N} \times \mathbb{N}$ such that $1 \leq i \leq \lambda_j$. For example, the partition $(4, 2, 1)$, which we abbreviate as 421, has Young diagram



.

Define the *conjugate* or *transpose* λ^t of λ to be the partition with λ_i cells in column i. For example, $421^t = 3211$. For another partition μ, we write $\mu \subseteq \lambda$ whenever μ is contained within λ (as Young diagrams); equivalently $\mu_i \leq \lambda_i$ for all i. In this case, we define the *skew shape* λ/μ to be the set theoretic difference $\lambda - \mu$. In particular, the partition λ is the skew shape $\lambda/\emptyset$. We call the number of cells of λ/μ its *size*, denoted $|\lambda/\mu|$. We say that a skew shape forms a *horizontal strip* (respectively *vertical strip*) if it contains no two cells in the same column (resp. row). A k-*horizontal strip* is a horizontal strip of size k, and similarly for vertical strips. For example, the skew shape 421/21 is a 4-horizontal strip:

 .

With another skew shape σ/τ, we let $(\lambda/\mu) * (\sigma/\tau)$ denote the skew shape obtained by positioning λ/μ so that its bottom right cell is immediately above and left of the top left cell of σ/τ. For example, the horizontal strip 421/21 above could alternatively be written as $(21/1) * (2)$ or as $(1) * (31/1)$.

A *Young tableau* of shape λ/μ is a map from the cells of λ/μ to the positive integers. A *semistandard Young tableau* (SSYT) is such a filling which is weakly increasing from left-to-right along each row and strictly increasing up each column, such as

$$\begin{array}{|c|c|c|c|}\cline{1-1} 5 & \multicolumn{3}{c}{} \\ \cline{1-3} 3 & 3 & 5 & \multicolumn{1}{c}{} \\ \hline \multicolumn{1}{c|}{} & 1 & 2 & 7 \\ \cline{2-4}\end{array} \quad .$$

The *content* of an SSYT T is the sequence π such that T has π_i cells with entry i; in this case $\pi = (1, 1, 2, 0, 2, 0, 1)$.

We let Λ denote the ring of symmetric functions in the variables $x = (x_1, x_2, \ldots)$ over $\mathbb{Q}$, say. We will use three familiar bases from [Mac95, Sta99] for Λ: the elementary symmetric functions e_λ, the complete homogeneous symmetric functions h_λ and, most importantly, the *Schur functions* s_λ. The Schur functions form an orthonormal basis for Λ with respect to the Hall inner product and may be defined in terms of SSYTs by

$$s_\lambda = \sum_{T \in \mathrm{SSYT}(\lambda)} x^T, \tag{2.1}$$

where the sum is over all SSYTs of shape λ and where x^T denotes the monomial $x_1^{\pi_1} x_2^{\pi_2} \cdots$ when T has content π. Replacing λ by λ/μ in (2.1) gives the definition of the *skew Schur function* $s_{\lambda/\mu}$, where the sum is now over all SSYTs of shape λ/μ. For example, the SSYT shown above contributes the monomial $x_1 x_2 x_3^2 x_5^2 x_7$ to $s_{431/1}$.

3 The skew Pieri rule

The celebrated Pieri rule gives an elegant method for expanding the product $s_\lambda s_n$ in the Schur basis. This rule was originally stated in [Pie93] in the setting of Schubert Calculus. Recall that the single row Schur function s_n equals the complete homogeneous symmetric function h_n. Recall also the involution ω on Λ, which may be defined by sending the Schur function s_λ to s_{λ^t} or equivalently by sending h_k to e_k. Thus the Schur function s_{1^n} equals the elementary symmetric function e_n, where 1^n denotes a single column of size n.

Theorem 3.1 ([Pie93]) *For any partition λ and positive integer n, we have*

$$s_\lambda s_n = s_\lambda h_n = \sum_{\lambda^+/\lambda \text{ } n\text{-hor. strip}} s_{\lambda^+}, \tag{3.1}$$

where the sum is over all partitions λ^+ such that λ^+/λ is a horizontal strip of size n.

Applying the involution ω to (3.1), we get the dual version of the Pieri rule:

$$s_\lambda s_{1^n} = s_\lambda e_n = \sum_{\lambda^+/\lambda \text{ } n\text{-vert. strip}} s_{\lambda^+} \ , \tag{3.2}$$

where the sum is now over all partitions λ^+ such that λ^+/λ is a vertical strip of size n.

A simple application of Theorem 3.1 gives

$$s_{322}s_2 = s_{3222} + s_{3321} + s_{4221} + s_{432} + s_{522},$$

as represented diagrammatically in Figure 1.

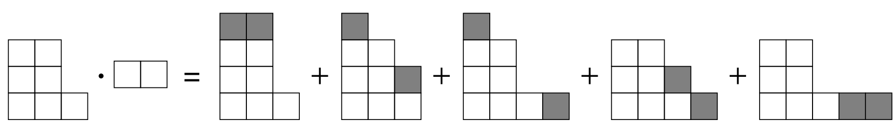


Fig. 1: The expansion of $s_{322}s_2$ by the Pieri rule.

Given the simplicity of (3.1), it is natural to hope for a simple expression for $s_{\lambda/\mu}s_n$ in terms of skew Schur functions. This brings us to our main result.

Theorem 3.2 *For any skew shape λ/μ and positive integer n, we have*

$$s_{\lambda/\mu}s_n = s_{\lambda/\mu}h_n = \sum_{k=0}^{n}(-1)^k \sum_{\substack{\lambda^+/\lambda \text{ } (n-k)\text{-hor. strip} \\ \mu/\mu^- \text{ } k\text{-vert. strip}}} s_{\lambda^+/\mu^-} \ , \tag{3.3}$$

where the second sum is over all partitions λ^+ and μ^- such that λ^+/λ is a horizontal strip of size $n-k$ and μ/μ^- is a vertical strip of size k.

Observe that when $\mu = \emptyset$, Theorem 3.2 specializes to Theorem 3.1. Again, we can apply the ω transformation to obtain the dual version of the skew Pieri rule.

Corollary 3.3 *For any skew shape λ/μ and any positive integer k, we have*

$$s_{\lambda/\mu}s_{1^n} = s_{\lambda/\mu}e_n = \sum_{k=0}^{n}(-1)^k \sum_{\substack{\lambda^+/\lambda \text{ } (n-k)\text{-vert. strip} \\ \mu/\mu^- \text{ } k\text{-hor. strip}}} s_{\lambda^+/\mu^-},$$

where the sum is over all partitions λ^+ and μ^- such that λ^+/λ is a vertical strip of size $n-k$ and μ/μ^- is a horizontal strip of size k.

Example 3.4 *A direct application of Theorem 3.2 gives*

$$s_{322/11}s_2 = s_{3222/11} + s_{3321/11} + s_{4221/11} + s_{432/11} + s_{522/11} \\ - s_{3221/1} - s_{332/1} - s_{422/1} + s_{322},$$

as represented diagrammatically in Figure 2.

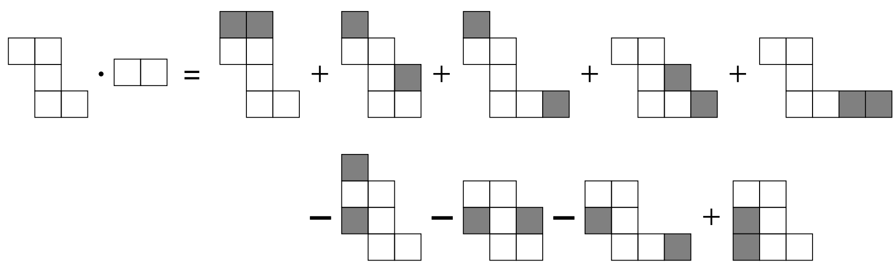


Fig. 2: The expansion of $s_{322/11}s_2$ by the skew Pieri rule.

As (3.3) contains negative signs, our approach to proving Theorem 3.2 will be to construct a sign-reversing involution on SSYTs of shapes of the form λ^+/μ^-. We will then provide a bijection between SSYTs of shape λ^+/μ^- that are fixed under the involution and SSYTs of shape $(\lambda/\mu) * (n)$. The result then follows from the fact that $s_{(\lambda/\mu)*(n)} = s_{\lambda/\mu}s_n$.

4 Row insertion

In order to describe our sign-reversing involution, we will need the Robinson-Schensted-Knuth (RSK) row insertion algorithm on SSYTs [Rob38, Sch61, Knu70]. For a thorough treatment of this algorithm along with many applications, we recommend [Sta99]. In fact, we will use an analogue of the algorithm for SSYTs of *skew* shape from [SS90]. There, row insertion comes in two forms, external and internal row insertion. External row insertion, which we now define, is just like the classical RSK insertion.

Definition 4.1 *Let T be an SSYT of arbitrary skew shape and choose a positive integer k. Define the* external row insertion *of k into T, denoted $T \leftarrow_0 k$, as follows: if k is weakly larger than all entries in row 1 of T, then add k to the right end of the row and terminate the process. Otherwise, find the leftmost cell in row 1 of T whose entry, say k', is greater than k. Replace this entry by k and then row insert k' into T at row 2 using the procedure just described. Repeat the process until some entry comes to rest at the right end of a row.*

Example 4.2 *Let $\lambda/\mu = 7541/32$ and $\lambda^+/\mu^- = 7542/31$ so that the outlined entries below are those in λ/μ. The result of externally row inserting a 2 is shown below, with changed cells circled.*

$$\begin{array}{cccccc} 4 & 5 & \bigcirc & & & \\ 2 & 2 & 7 & 7 & & \\ & 1 & 2 & 3 & 4 & \\ & & 2 & 2 & 3 & 6 \end{array} \quad \leftarrow_0 2 \quad = \quad \begin{array}{cccccc} 4 & 5 & 7 & & & \\ 2 & 2 & 4 & 7 & & \\ & 1 & 2 & 3 & 3 & \\ & & 2 & 2 & 2 & 6 \end{array} \tag{4.1}$$

An *inside corner* (resp. *outside corner*) of an SSYT T is a cell that has no cell of T immediately below or to its left (resp. above or to its right). Therefore, inside and outside corners are those individual cells whose removal from T still yields an SSYT of skew shape.

Definition 4.3 *Let T be an SSYT of arbitrary skew shape and let T have an inside corner in row r with entry k. Define the* internal row insertion of k from row r into T, *denoted $T \leftarrow_r k$, as the removal of k from row r and its insertion, using the rules for external row insertion, into row $r+1$. The insertion proceeds until some entry comes to rest at the right end of a row.*

We could regard external insertions as internal insertions from row 0, explaining our notation. We will simply write $T \leftarrow k$ when specifying the type or row of the insertion is unnecessarily cumbersome.

Example 4.4 *Taking T as the SSYT on the right in* (4.1)*, the result of internally row inserting the 1 from row 2 is shown below.*

$$\begin{array}{ccccccc} \bigcirc & & & & & & \\ \textcircled{4} & 5 & 7 & & & & \\ \textcircled{2} & 2 & 4 & 7 & & & \\ & \textcircled{1} & 2 & 3 & 3 & & \\ & & & 2 & 2 & 2 & 6 \end{array} \quad \leftarrow_2 1 \quad = \quad \begin{array}{ccccccc} \textcircled{4} & & & & & & \\ \textcircled{2} & 5 & 7 & & & & \\ \textcircled{1} & 2 & 4 & 7 & & & \\ & \bigcirc & 2 & 3 & 3 & & \\ & & & 2 & 2 & 2 & 6 \end{array} \tag{4.2}$$

For both types of insertion, we must be a little careful when inserting an entry into an empty row, say row i: in this case $\lambda_i = \mu_i$ and the entry must be placed in column $\lambda_i + 1$.

Note that an internal insertion results in the same multiset of entries while an external insertion adds an entry. It is not difficult to check that either operation results in an SSYT.

We will also need to invert row insertions, again for skew shapes and following [SS90].

Definition 4.5 *Let T be an SSYT of arbitrary skew shape and choose an outside corner c of T, say with entry k. Define the* reverse row insertion of c from T, *denoted $T \to c$, by deleting c from T and reverse inserting k into the row below, say row r, as follows: if $r = 0$, then the procedure terminates. Otherwise, if k is weakly smaller than all entries in row r, then place k at the left end of row r and terminate the procedure. Otherwise, find the rightmost cell in row r whose entry, say k', is less than k. Replace this entry by k and then reverse row insert k' into row $r-1$ using the procedure just described.*

Example 4.6 *In* (4.1)*, reverse row insertion of the cell containing the circled 7 from the SSYT on the right results in the SSYT on the left, and similarly in* (4.2) *for the circled 4.*

As with row insertion, it follows from the definition that the resulting array will again be an SSYT. Observe that the first type of termination mentioned in Definition 4.5 corresponds to reverse external row insertion, and we then say that k *lands in row* 0. The second type of termination corresponds to reverse internal row insertion, and we then say that k *lands in row* r. In both cases, we will call the entry k left at the end of the procedure the *final entry of* $T \to c$. The following lemma, which follows immediately from Definitions 4.1, 4.3 and 4.5, formalizes the bijectivity of row and reverse row insertion.

Lemma 4.7 *Let T be an SSYT of skew shape.*

a. If S is the result of $T \leftarrow k$ for some positive integer k, then $S \to c$ results in T, where c is the unique non-empty cell of S that is empty in T.

b. *If S is the result of $T \to c$ for some removable cell c of T and the final entry k of $T \to c$ lands in row $r \geq 0$, then $S \leftarrow_r k$ results in T.*

For both row insertion and reverse row insertion, we will often want to track the cells affected by the procedure. Therefore define the *bumping path of the row insertion* $T \leftarrow k$ (resp. the *reverse bumping path of a reverse row insertion* $T \to c$) to be the set of cells in T, as well as those empty cells, where the entries differ from the corresponding entries in $T \leftarrow k$ (resp. $T \to c$). The cells of the bumping paths for row insertion and reverse row insertion are circled in (4.1) and (4.2). Note that the fact that the bumping path and reverse bumping path are equal in each of these examples is a consequence of Lemma 4.7.

It is easy to see that the bumping paths always move weakly right from top to bottom in the case of column-strict tableaux. The following bumping lemma will play a crucial role in defining our sign-reversing involution and in proving its relevant properties.

Lemma 4.8 *Let T be an SSYT of skew shape and let k, k' be positive integers. Let B be the bumping path of $T \leftarrow k$ and let B' be the bumping path of $(T \leftarrow k) \leftarrow k'$.*

a. *If B is strictly left of B' in any row r, then B is strictly left of B' in every row they both occupy. Moreover, the top cells of B and B' form a horizontal strip.*

b. *If both row insertions are external, then B is strictly left of B' in every row they both occupy if and only if $k \leq k'$.*

c. *Suppose C' is the reverse bumping path of $T \to c'$ with final entry k' and C is the reverse bumping path of $(T \to c') \to c$ with final entry k. If c is strictly left of c', then C is strictly left of C' in every row they both occupy. If, in addition, both reverse row insertions land in row 0, then $k \leq k'$.*

To foreshadow the role of Lemma 4.8 in the following section, we give a proof of the classical Pieri rule using this result.

Proof of Theorem 3.1: The formula is proved if we can give a bijection between SSYTs of shape $\lambda * (n)$ and SSYTs of shape λ^+ such that λ^+/λ is a horizontal strip of size n. Let $k_1 \leq \cdots \leq k_n$ be entries of (n) from left to right. Repeated applications of (a) and (b) of Lemma 4.8 ensure that row inserting these entries into an SSYT of shape λ will add a horizontal strip of size n to λ. By Lemma 4.7, this establishes a bijection where the inverse map is given by reverse row inserting the cells of λ^+/λ from right to left. □

5 A sign-reversing involution

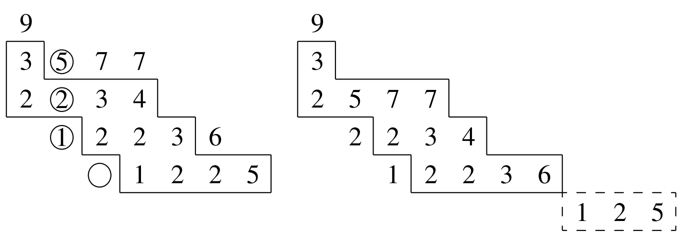
Throughout this section, fix a skew shape λ/μ. We will be interested in SSYTs of shape λ^+/μ^-, where we always assume that λ^+/λ is a horizontal strip, μ/μ^- is a vertical strip, and $|\lambda^+/\lambda| + |\mu/\mu^-| = n$. Our goal is to construct a sign-reversing involution on SSYTs whose shapes take the form λ^+/μ^-, such that the fixed points are in bijection with SSYTs of shape $(\lambda/\mu) * (n)$.

Our involution is reminiscent of the proof of the classical Pieri rule given in Section 4. By Lemma 4.7, reverse row insertion gives a bijective correspondence provided we record the final entry and its landing row. Our strategy, then, is to reverse row insert the cells of λ^+/λ from right to left, recording the entries as we go. If at some stage we land in row $r \geq 1$, we will then re-insert all the previous final entries. More formally, we have the following definition of a *downward slide of T*.

Definition 5.1 *Let T be an SSYT of shape λ^+/μ^-. Define the* downward slide *of T, denoted $\mathrm{D}(T)$, as follows: construct $T \to c_1$ where c_1 is the rightmost cell of λ^+/λ, and let k_1 denote the final entry. If k_1 lands in row 0, then continue with c_2 the second rightmost cell of λ^+/λ and k_2 the final entry of $(T \to c_1) \to c_2$. Continue until the first time k_m lands in row $r \geq 1$ and set $m' = m - 1$, or set $m = m' = |\lambda^+/\lambda|$ if no such k_m exists. Then $\mathrm{D}(T)$ is given by*

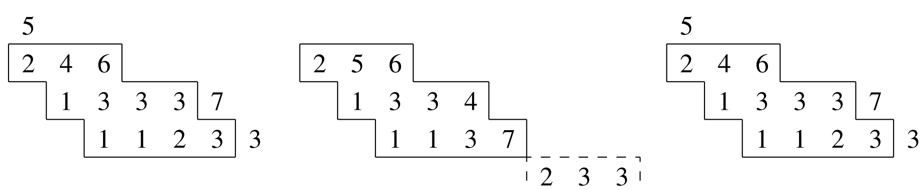
$$(\cdots(((\cdots(T \to c_1) \to c_2 \cdots) \to c_m) \leftarrow k_{m'}) \cdots) \leftarrow k_1.$$

Example 5.2 *With T shown on the left below, we exhibit the construction of $\mathrm{D}(T)$ in two steps. We find that $m = 4$ and the middle SSYT shows the result of $(((T \to c_1) \to c_2) \to c_3) \to c_4$. The entries that land in row 0 are recorded in the dashed box. Then the SSYT on the right is $\mathrm{D}(T)$. The significance of the circles will be explained later.*



(5.1)

Alternatively, if T is the SSYT shown on the left below, we find that $m = 3$ and that all three final entries land in row 0. Then $m' = |\lambda^+/\lambda|$ and Lemma 4.7 ensures that $D(T) = T$. Below in the middle, we have shown $((T \to c_1) \to c_2) \to c_3$. The position of the dashed box is intended to be suggestive: together with the entries in the outlined shape, we see that we have an SSYT of shape $(\lambda/\mu) * (n) = (653/21) * (3)$.



The final reverse bumping path in a downward slide will play an important role in the sign-reversing involution. Therefore, with notation as in Definition 5.1, if $m < |\lambda^+/\lambda|$, then we refer to the reverse bumping path of $T \to c_m$ as the *downward path of* T. The cells of the downward path of T are circled above. Say that the downward path of T *exits right* if its bottom cell (which may be empty) is strictly below the bottom cell μ/μ^-. Our terminology is justified since one can show that the exits right condition is equivalent to the bottom cell of the downward path being weakly right of the bottom cell of μ/μ^-. The importance of the exits right condition is revealed by the following result.

Proposition 5.3 *Suppose T is an SSYT of shape λ^+/μ^- such that the downward path of T, if it exists, exits right. Then $\mathrm{D}(T)$ is an SSYT of shape λ'/μ', where λ'/λ (resp. μ/μ') is a horizontal (resp. vertical) strip.*

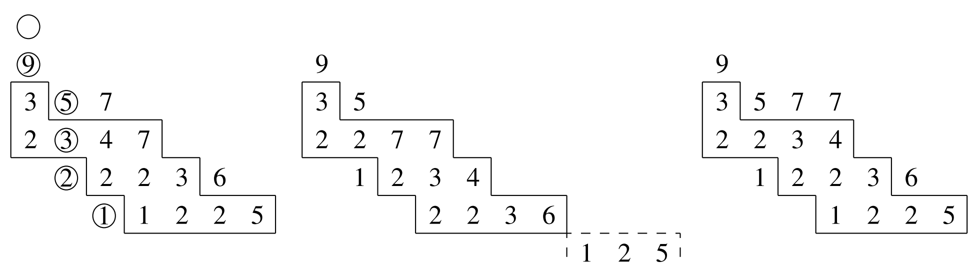
Supposing $\mathrm{D}(S) = T$ with $T \neq S$, the next step is to invert the downward slides for such T. Since any such T necessarily has $\mu^- \neq \mu$, the idea is to internally row insert the bottom cell of μ/μ^-. However, before doing this we must reverse row insert certain cells of λ^+/λ, as in a downward slide. To describe which cells to reverse insert, we define the *upward path of* T to be the bumping path that would result from internal row insertion of the entry in the bottom cell of μ/μ^-. Roughly, we will reverse row insert anything that is weakly right of this upward path.

Definition 5.4 *Let T be an SSYT of shape λ^+/μ^- such that $\mu^- \neq \mu$. Define the* upward slide of T, *denoted $\mathrm{U}(T)$, as follows: construct $T \to c_1$ where c_1 is the rightmost cell of λ^+/λ, and let k_1 denote its final entry and B_1 its bumping path. If B_1 fails to stay weakly right of the upward path of T, then set $m = m' = 0$. Otherwise, consider c_2, the second rightmost cell of λ^+/λ, and k_2, the final entry of $(T \to c_1) \to c_2$, and B_2, the corresponding bumping path. Continue until the last time B_m stays weakly right of the upward path of T or until no cell of λ^+/λ remains. Suppose that after the reverse row insertions, the bottom cell of μ/μ^- is in row r and has entry k. Then $\mathrm{U}(T)$ is given by*

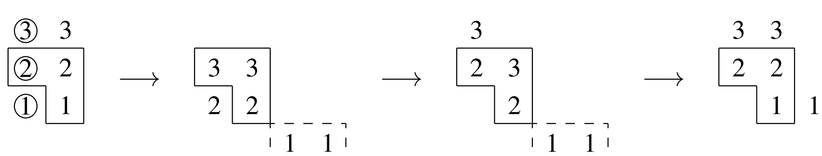
$$(\cdots((((\cdots(T \to c_1) \to c_2 \cdots) \to c_m) \leftarrow_r k) \leftarrow k_{m'}) \cdots) \leftarrow k_1 \tag{5.2}$$

where we set $m' = m$ if k_m lands in row 0, and $m' = m - 1$ otherwise.

Example 5.5 *Letting T be the rightmost SSYT of* (5.1)*, the cells of the upward path of T are circled below. We determine $\mathrm{U}(T)$ in three steps. We find that $m = 3$ and the middle SSYT of* (5.1) *shows $((T \to c_1) \to c_2) \to c_3$. Then $(((T \to c_1) \to c_2) \to c_3) \leftarrow k$ is shown in the middle below, while $\mathrm{U}(T)$ is shown on the right. Comparing with Example 5.2, we observe that the upward slide in this case does indeed invert the downward slide.*



There are also instances where the entry k of Definition 5.4 is different from the entry originally at the bottom of the upward path. For example, the same three-step process for constructing $\mathrm{U}(T)$ is shown below for an example with $m = 2$. There, $k = 2$, even though the original upward path of T had 1 as its bottom entry.



As with downward slides and before presenting our involution, we must ensure that the result of an upward slide is always a tableau of the appropriate skew shape.

Proposition 5.6 *Suppose T is an SSYT of shape λ^+/μ^- such that in the upward slide of T, all the final entries of the reverse row insertions land in row 0. Then $\mathrm{U}(T)$ is an SSYT of shape λ'/μ', where λ'/λ (resp. μ/μ') is a horizontal (resp. vertical) strip.*

Our involution will consist of either applying a downward slide or an upward slide. The decision for which slide to apply is roughly based on which of the downward path of T or the upward path of T lies further to the right.

Definition 5.7 *Consider the set of SSYTs T of shape λ^+/μ^- such that that λ^+/λ is a horizontal strip and μ/μ^- is a vertical strip. Define a map ϕ on such T by*

$$\phi(T) \;=\; \begin{cases} \mathrm{D}(T) & \text{if } T \text{ has no upward path or} \\ & \text{the downward path of } T \text{ exists and exits right,} \\ \mathrm{U}(T) & \text{otherwise.} \end{cases}$$

Theorem 5.8 *The map ϕ defines an involution on the set of SSYTs with shapes of the form λ^+/μ^- where λ^+/λ is a horizontal strip and μ/μ^- is a vertical strip.*

We now have all the ingredients needed to prove the skew Pieri rule.

Proof of Theorem 3.2: Using the expansion of s_{λ^+/μ^-} in terms of SSYTs as in (2.1), observe that if $\phi(T) \neq T$, then the T and $\phi(T)$ occur with different signs in the right-hand side of (3.3). Since ϕ clearly preserves the monomial associated to an SSYT, the monomials corresponding to T and $\phi(T)$ in the right-hand side of (3.3) will cancel out. Because $s_{\lambda/\mu}s_n = s_{(\lambda/\mu)*(n)}$, it remains to show that there is a monomial-preserving bijection from fixed points of ϕ to SSYTs of shape $(\lambda/\mu) * (n)$.

Note that T is a fixed point of ϕ only if T has neither an upward path nor a downward path. This happens if and only if $\mu^- = \mu$ and when reverse row inserting the cells of λ^+/λ from right to left, every final entry lands in row 0. In particular, the entries of T remaining after reverse row inserting the cells of λ^+/λ form an SSYT of shape λ/μ. Say the final entries of the reverse row insertions are $k_n, \ldots, k_1$ in the order removed. By Lemma 4.8(c), since λ^+/λ is a horizontal strip, we have $k_1 \leq \cdots \leq k_n$ and so these entries form an SSYT of shape (n). By Lemma 4.7, this process is invertible and therefore establishes the desired bijection. □

Remark 5.9 *We proved Theorem 3.2 by working with SSYTs. In particular, we showed that the two sides of* (3.3) *were equal when expanded in terms of monomials. Alternatively, we could consider the expansions of both sides of* (3.3) *in terms of Schur functions. The Littlewood-Richardson rule states that the coefficient of s_ν in the expansion of any skew Schur function $s_{\lambda/\mu}$ is the number of* Littlewood-Richardson fillings *(*LR-fillings*) of shape λ/μ and content ν. (The interested reader unfamiliar with LR-fillings can find the definition in [Sta99].) It is not hard to check that our maps* D *and* U *send LR-fillings to LR-fillings, and bumping within LR-fillings has some nice properties. For example, the entries along a (reverse) bumping path are always $1, 2, \ldots, r$ from bottom to top for some r.*

However, we chose to give our proof in terms of SSYTs because one does not need to invoke the power of the Littlewood-Richardson rule to prove the classical Pieri rule, and we wanted the same to apply to the skew Pieri rule.

References

[AM09] Sami H. Assaf and Peter R. W. McNamara. A Pieri rule for skew shapes. Preprint. With an appendix by Thomas Lam. `arXiv:0908.0345`, 2009.

[Buc99] Anders S. Buch. Littlewood-Richardson calculator, 1999. Available from `http://www.math.rutgers.edu/~asbuch/lrcalc`.

[Knu70] Donald E. Knuth. Permutations, matrices, and generalized Young tableaux. *Pacific J. Math.*, 34:709–727, 1970.

[Koo88] Tom H. Koornwinder. Self-duality for q-ultraspherical polynomials associated with root system a_n. Unpublished manuscript. `http://staff.science.uva.nl/~thk/art/informal/dualmacdonald.pdf`, 1988.

[Lam05] Thomas Lam. Ribbon tableaux and the Heisenberg algebra. *Math. Z.*, 250(3):685–710, 2005.

[Las89] Michel Lassalle. Une formule de Pieri pour les polynômes de Jack. *C. R. Acad. Sci. Paris Sér. I Math.*, 309(18):941–944, 1989.

[LLS09] Thomas Lam, Aaron Lauve, and Frank Sottile. Skew Littlewood-Richardson rules from Hopf algebras. Preprint. `arXiv:0908.3714`, 2009.

[LR34] D.E. Littlewood and A.R. Richardson. Group characters and algebra. *Philos. Trans. Roy. Soc. London, Ser. A*, 233:99–141, 1934.

[LS82] Alain Lascoux and Marcel-Paul Schützenberger. Polynômes de Schubert. *C. R. Acad. Sci. Paris Sér. I Math.*, 294(13):447–450, 1982.

[LS07] Cristian Lenart and Frank Sottile. A Pieri-type formula for the K-theory of a flag manifold. *Trans. Amer. Math. Soc.*, 359(5):2317–2342 (electronic), 2007.

[Mac87] I. G. Macdonald. The symmetric functions $P(x; q, t)$: facts and conjectures. Unpublished manuscript, 1987.

[Mac95] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[Man98] Laurent Manivel. *Fonctions symétriques, polynômes de Schubert et lieux de dégénérescence*, volume 3 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 1998.

[Mor64] A. O. Morris. A note on the multiplication of Hall functions. *J. London Math. Soc.*, 39:481–488, 1964.

[Pie93] Mario Pieri. Sul problema degli spazi secanti. *Rend. Ist. Lombardo (2)*, 26:534–546, 1893.

[Rob38] G. de B. Robinson. On the Representations of the Symmetric Group. *Amer. J. Math.*, 60(3):745–760, 1938.

[Sch61] C. Schensted. Longest increasing and decreasing subsequences. *Canad. J. Math.*, 13:179–191, 1961.

[Sch77] M.-P. Schützenberger. La correspondance de Robinson. In *Combinatoire et représentation du groupe symétrique (Actes Table Ronde CNRS, Univ. Louis-Pasteur Strasbourg, Strasbourg, 1976)*, pages 59–113. Lecture Notes in Math., Vol. 579. Springer, Berlin, 1977.

[Sot96] Frank Sottile. Pieri's formula for flag manifolds and Schubert polynomials. *Ann. Inst. Fourier (Grenoble)*, 46(1):89–110, 1996.

[SS90] Bruce E. Sagan and Richard P. Stanley. Robinson-Schensted algorithms for skew tableaux. *J. Combin. Theory Ser. A*, 55(2):161–193, 1990.

[Sta89] Richard P. Stanley. Some combinatorial properties of Jack symmetric functions. *Adv. Math.*, 77(1):76–115, 1989.

[Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[Ste] John R. Stembridge. SF, posets and coxeter/weyl. Available from `http://www.math.lsa.umich.edu/~jrs/maple.html`.

[Tho74] Glânffrwd P. Thomas. *Baxter algebras and Schur functions*. PhD thesis, University College of Swansea, 1974.

[Tho78] Glânffrwd P. Thomas. On Schensted's construction and the multiplication of Schur functions. *Adv. in Math.*, 30(1):8–32, 1978.

[Win98] Rudolf Winkel. On the multiplication of Schubert polynomials. *Adv. in Appl. Math.*, 20(1):73–97, 1998.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 26–34

The expansion of Hall-Littlewood functions in the dual Grothendieck polynomial basis

Jason Bandlow[1] and Jennifer Morse[2]

[1] *University of Pennsylvania, Philadelphia, PA, USA*
[2] *Drexel University, Philadelphia, PA, USA*

Abstract. A combinatorial expansion of the Hall-Littlewood functions into the Schur basis of symmetric functions was first given by Lascoux and Schützenberger, with their discovery of the charge statistic. A combinatorial expansion of stable Grassmannian Grothendieck polynomials into monomials was first given by Buch, using set-valued tableaux. The dual basis of the stable Grothendieck polynomials was given a combinatorial expansion into monomials by Lam and Pylyavskyy using reverse plane partitions. We generalize charge to set-valued tableaux and use all of these combinatorial ideas to give a nice expansion of Hall-Littlewood polynomials into the dual Grothendieck basis.

Résumé. En associant une charge à un tableau, une formule combinatoire donnant le développement des polynômes de Hall-Littlewood en termes des fonctions de Schur a été obtenue par Lascoux et Schützenberger. Une formule combinatoire donnant le développement des polynômes de Grothendieck Grassmanniens stables en termes des fonctions monomiales a quant à elle été obtenue par Buch à l'aide de tableaux à valeurs sur des ensembles. Finalement, une formule faisant intervenir des partitions planaires inverses a été obtenue par Lam et Pylyavskyy pour donner le développement de la base duale aux polynômes de Grothendieck stables en termes de monômes. Nous généralisons le concept de charge aux tableaux à valeurs sur des ensembles et, en nous servant de toutes ces notions combinatoires, nous obtenons une formule élégante donnant le développement des polynômes de Hall-Littlewood en termes de la base de Grothendieck duale.

Keywords: symmetric functions, Hall-Littlewood polynomials, Grothendieck polynomials, charge statistic

1 Introduction

The *Hall-Littlewood functions* are symmetric functions with a wealth of applications. In various forms, they interpolate between the complete homogeneous and Schur basis of symmetric functions, provide a polynomial realization of the Hall algebra, and have several interpretations as characters of representations. Lascoux and Schützenberger gave an expansion of the Hall-Littlewood functions into Schur functions, in terms of a statistic on tableaux called *charge* [LS78]. We write this expansion as

$$H_\mu[X;t] = \sum_\lambda \sum_{T \in SST(\lambda,\mu)} t^{ch(T)} s_\lambda \tag{1}$$

where $SST(\lambda, \mu)$ is the set of all semi-standard tableaux of shape λ and evaluation μ. From this expansion, the following properties of the Hall-Littlewood function are more or less immediate:

- the coefficient of s_λ is a polynomial in t with non-negative integer coefficients,
- specializing $t = 1$ gives $H_\mu[X; 1] = h_\mu$, the complete homogeneous symmetric function, and
- specializing $t = 0$ gives $H_\mu[X; 0] = s_\mu$.

See [Mac95] for more details about Hall-Littlewood functions.

The *Grothendieck* polynomials were introduced by Lascoux and Schützenberger [LS83] as power series representatives for the K-theory classes of the structure sheaves of Schubert varieties. The stable Grothendieck polynomials introduced by Fomin and Kirillov [FK94] are the stable limit of these as the number of variables approaches infinity. These functions, written G_λ, are non-homogeneous symmetric functions, which cannot be written as a finite sum of Schur functions. The function G_λ is equal to s_λ in its lowest degree homogeneous component, and (in terms of Schur functions or monomials) is sign-alternating by degree in the higher homogeneous components. Buch gave an expansion of the stable Grothendieck polynomials into monomial symmetric functions by introducing set-valued tableaux [Buc02]. We write this expansion as

$$G_\lambda = \sum_\mu (-1)^{|\mu|-|\lambda|} k_{\lambda,\mu} m_\mu \tag{2}$$

where $k_{\lambda,\mu}$ denotes the number of set-valued tableaux of shape λ and evaluation μ.

Lam and Pylyavskyy studied the dual basis to the stable Grothendieck polynomials under the Hall inner product [LP07]. They expanded these into monomials using a special evaluation of reverse plane partitions. We denote the dual basis to $\{G_\lambda\}$ by $\{g_\lambda\}$. These are Schur positive, non-homogeneous symmetric functions, with g_λ equal to s_λ in the top degree. We note that equation 2 and a simple fact about dual bases immediately implies

$$h_\mu = \sum_\lambda (-1)^{|\mu|-|\lambda|} k_{\lambda,\mu} g_\lambda \ . \tag{3}$$

In this work, we give a generalization of the charge statistic to set-valued tableaux. In particular, we define the reading word of a set-valued tableau, and then define the charge to be the charge of the reading word. We then prove the common generalization of equations 1 and 3:

$$H_\mu[X; t] = \sum_\lambda (-1)^{|\mu|-|\lambda|} \sum_{T \in SVT(\lambda,\mu)} t^{ch(T)} g_\lambda \tag{4}$$

where $SVT(\lambda, \mu)$ denotes the set of set-valued tableaux of shape λ and evaluation μ. We find it remarkable that such a nice formula exists, as we are unaware of any direct connection between the Hall-Littlewood functions and K-theory.

2 Definitions and notation

2.1 Symmetric function basics

We begin by setting our notation with some standard definitions. An introduction to symmetric functions can be found in [Mac95] or [Sta99].

Definition 1 *The* Young diagram *of a partition* $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_k)$ *is a left- and bottom-justified array of* 1×1 *square cells in the first quadrant of the coordinate plane, with* λ_i *cells in the* i^{th} *row from the bottom.*

Example 1 *The Young diagram of the partition* $(3, 2)$ *is* $\begin{array}{|c|c|c|}\cline{1-2} & & \multicolumn{1}{c}{} \\ \hline & & \\ \hline\end{array}$.

Definition 2 *A* semi-standard tableau *of shape* λ *is a filling of the cells in the Young diagram of* λ *with positive integers, such that the entries*

- *are weakly increasing while moving rightward across any row, and*
- *are strictly increasing while moving up any column.*

Example 2 *A semi-standard tableau of shape* $(3, 2)$ *is* $\begin{array}{|c|c|c|}\cline{1-2} 2 & 3 & \multicolumn{1}{c}{} \\ \hline 1 & 1 & 2 \\ \hline\end{array}$.

The *evaluation* of a semi-standard tableau is the sequence $(\alpha_i)_{i\in\mathbb{N}}$ where α_i is the number of cells containing i. The evaluation of the tableau in example 2 is $(2, 2, 1)$ (trailing 0's have been omitted, as is customary). We use the notation $SST(\lambda)$ to mean the set of all semi-standard tableaux of shape λ, and $SST(\lambda, \mu)$ to mean the set of all semi-standard tableaux of shape λ and evaluation μ.

The *Schur functions* have many definitions, one of which is in terms of semi-standard tableaux.

Definition 3 *The* Schur function s_λ *is defined by*

$$s_\lambda = \sum_{T\in SST(\lambda)} \mathbf{x}^{ev(T)} .$$

The notation $\mathbf{x}^{ev(T)}$ means $x_1^{\alpha_1} x_2^{\alpha_2} \cdots$, where $(\alpha_1, \alpha_2, \cdots)$ is the evaluation of T. The Schur functions are elements of $\mathbb{C}[[x_1, x_2, \cdots]]$, the power series ring in infinitely many variables, and are well known to be a basis for the *symmetric functions* (*i.e.,* those elements of $\mathbb{C}[[x_1, x_2, \cdots]]$ which are invariant under any permutation of their indices).

Example 3 *The Schur function* $s_{(2,1)}$ *can be written as*

$$s_{(2,1)} = x_1^2 x_2 + x_1 x_2^2 + 2x_1x_2x_3 + \cdots$$

corresponding to the tableaux

$$\begin{array}{|c|c|}\cline{1-1} 2 & \multicolumn{1}{c}{} \\ \hline 1 & 1 \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 2 & \multicolumn{1}{c}{} \\ \hline 1 & 2 \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 3 & \multicolumn{1}{c}{} \\ \hline 1 & 2 \\ \hline\end{array} \quad \begin{array}{|c|c|}\cline{1-1} 2 & \multicolumn{1}{c}{} \\ \hline 1 & 3 \\ \hline\end{array} \cdots .$$

Another basis for the symmetric functions is the *monomial symmetric functions*.

Definition 4 *The* monomial symmetric function m_λ *is defined by*

$$m_\lambda = \sum_\alpha \mathbf{x}^\alpha ,$$

where the sum is over all sequences α *which are a rearrangement of the parts of* λ*. (Here* λ *is thought of as having finitely many non-zero parts, followed by infinitely many* 0 *parts.)*

Example 4 *The monomial symmetric function $m_{(2,1)}$ can be written as*

$$m_{(2,1)} = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2 + \dots .$$

The *Kostka numbers* give the change of basis matrix between the Schur and monomial symmetric functions. For two partitions λ, μ, we define the number $K_{\lambda,\mu}$ to be number of semi-standard tableaux of shape λ and evaluation μ. From the previous definitions, one can see that a consequence of the symmetry of the Schur functions is that

$$s_\lambda = \sum_\mu K_{\lambda,\mu} m_\mu \ . \tag{5}$$

There is a standard inner product on the vector space of symmetric functions (known as the *Hall inner product*), defined by setting

$$\langle s_\lambda, s_\mu \rangle = \begin{cases} 1 & \text{if } \lambda = \mu \\ 0 & \text{otherwise.} \end{cases}$$

The following proposition is a basic, but very useful, fact of linear algebra.

Proposition 1 *If $(\{f_\lambda\}, \{f_\lambda^*\})$ and $(\{g_\lambda\}, \{g_\lambda^*\})$ are two pairs of dual bases for an inner-product space, and*

$$f_\lambda = \sum_\mu M_{\lambda,\mu} g_\mu \ ,$$

then

$$g_\mu^* = \sum_\lambda M_{\lambda,\mu} f_\lambda^* \ .$$

We define the set of *homogeneous symmetric functions*, $\{h_\lambda\}$, to be the dual-basis to the monomial symmetric functions. An immediate consequence of proposition 1 is that

$$h_\mu = \sum_\lambda K_{\lambda,\mu} s_\lambda \ . \tag{6}$$

2.2 Hall-Littlewood symmetric functions

The Hall-Littlewood functions belong to the space $\mathbb{C}[t][[x_1, x_2, \dots]]$ of formal power series in infinitely many variables with coefficients in the polynomial ring $\mathbb{C}[t]$. There are multiple (unequal, but related) definitions of these functions in the literature. The version we concern ourselves with here are most commonly written as H_λ or Q'_λ. In [LS78], Lascoux and Schützenberger found a purely combinatorial presentation of these functions. The key notion is a statistic on semi-standard tableaux known as *charge*. Before defining charge, we need the notion of the *reading word* of a tableau.

Definition 5 *The* reading word *of a tableau T, which we denote by $w(T)$, is the sequence $(w_1, w_2, \dots, w_n)$ obtained by listing the elements of T starting from the top-left corner, and reading across each row and then continuing down the rows.*

Example 5 *We have* $w\left(\begin{array}{|c|c|c} \hline 2 & 3 & \\ \hline 1 & 1 & \multicolumn{1}{c|}{2} \\ \hline \end{array}\right) = (2, 3, 1, 1, 2).$

We will first define the charge of a word, and then define the charge of a tableau to be the charge of its reading word. For our purposes, it will be sufficient to define charge only on words whose evaluation is a partition. While this can be extended to all semi-standard tableau, this requires a substantially more complicated definition. We begin by defining the charge of a *permutation*; this is a word with evaluation $(1, 1, \dots, 1)$. If w is a permutation of length n, then the charge of w is given by $\sum_{i=1}^{n} c_i(w)$ where $c_1(w) = 0$ and $c_i(w)$ is defined recursively as

$$c_i(w) = c_{i-1}(w) + \chi\,(i \text{ appears to the right of } i-1 \text{ in } w)\,. \tag{7}$$

Here we use the notation that when P is a proposition, $\chi(P)$ is equal to 1 if P is true and 0 if P is false.

Example 6 *A straightforward computation shows that*

$$ch(3, 5, 1, 4, 2) = 0 + 1 + 1 + 2 + 2 = 6.$$

We will now describe the decomposition of a word with partition content into *charge subwords*, each of which are permutations. The charge of a word will then be defined as the sum of the charge of its charge subwords. To find the first charge subword $w^{(1)}$ of a word w, we begin at the *right* of w (*i.e.,* at the last element of w) and move leftward through the word, marking the first 1 that we see. After marking a 1, we continue to travel to the left, now marking the first 2 that we see. If we reach the beginning of the word, we loop back to the end. We continue in this manner, marking successively larger elements, until we have marked the largest letter in w, at which point we stop. The subword of w consisting of the marked elements (with relative order preserved) is the first charge subword. We then remove the marked elements from w to obtain a word w'. The process then continues iteratively, with the second charge subword being the first charge subword of w', and so on.

Example 7 *We illustrate the first charge subword of* $w = (5, 2, 3, 4, 4, 1, 1, 1, 2, 2, 3)$ *by labeling the relevant elements in bold:* $(\mathbf{5}, \mathbf{2}, 3, 4, \mathbf{4}, 1, 1, \mathbf{1}, 2, 2, \mathbf{3})$. *If we remove the bold letters, and bold the second charge subword, we obtain* $(\mathbf{3}, \mathbf{4}, 1, \mathbf{1}, 2, \mathbf{2})$. *It is now easy to see that the third and final charge subword is* $(\mathbf{1}, \mathbf{2})$. *Thus we have the following computation of the charge of* w*:*

$$\begin{aligned} ch(w) &= ch(5, 2, 4, 1, 3) + ch(3, 4, 1, 2) + ch(1, 2) \\ &= (0 + 0 + 1 + 1 + 1) + (0 + 1 + 1 + 2) + (0 + 1) \\ &= 8 \end{aligned}$$

We can now define the Hall-Littlewood polynomials.

Definition 6 *The Hall-Littlewood polynomial* $H_\lambda[X; t]$ *is defined by*

$$H_\mu[X; t] = \sum_\lambda \sum_{T \in SST(\lambda, \mu)} t^{ch(T)} s_\lambda$$

Note the similarity of this definition and equation (6). In particular, if we set $t = 1$ in definition 6, we obtain equation (6) exactly.

2.3 Grothendieck polynomials

To define the Grothendieck polynomials, we need first need a definition of set-valued tableaux, due to Buch [Buc02].

Definition 7 *A* set-valued tableau *of shape λ is a filling of the cells in the Young diagram of λ with sets of positive integers, such that*

- *the maximum element in any cell is weakly smaller than the minimum element of the cell to its right, and*
- *the maximum element in any cell is strictly smaller than the minimum entry of the cell above it.*

Another way to think about this definition is that if we select a single element from each cell (in any possible way) we will always end up with a semi-standard tableau.

Example 8

A set-valued tableau of shape $(3,2)$ *is*
$$\begin{array}{|c|c|c} \cline{1-2} 3 & 4,5,6 & \\ \hline 1,2 & 2,3 & \multicolumn{1}{c|}{3} \\ \hline \end{array}\ .$$

We have omitted the set braces, '$\{$' and '$\}$', here and throughout for clarity.

The *evaluation* of a set-valued tableaux S is the composition $\alpha = (\alpha_i)_{i\geq 1}$ where α_i is the total number of times i appears in S. For example, the evaluation of the tableau in example 8 is $(1,2,3,1,1,1)$. The collection of all set-valued tableaux of shape λ will be denoted $SVT(\lambda)$ and the collection of all set-valued tableaux of shape λ and evaluation α will be denoted $SVT(\lambda,\alpha)$. We write $k_{\lambda,\mu}$ for the number of set-valued tableaux of shape λ and evaluation μ.

We will use set-valued tableaux to define the Grothendieck polynomials.

Definition 8 *We define the polynomials $G_\lambda(X)$ by*

$$G_\lambda = \sum_\mu (-1)^{|\mu|-|\lambda|} k_{\lambda,\mu} m_\mu$$

We note that when $|\mu| = |\lambda|$, we must have one element in every cell; hence G_λ is equal to s_λ plus higher degree terms. Since the G_λ are known to be symmetric functions, they must therefore form a basis.

Applying proposition 1 to this definition gives

$$h_\mu = \sum_\mu (-1)^{|\mu|-|\lambda|} k_{\lambda,\mu} g_\lambda \tag{8}$$

where the $\{g_\lambda\}$ are the dual basis to the $\{G_\lambda\}$.

The dual Grothendieck polynomials g_λ were studied by Lam and Pylyavskyy [LP07]. They gave an expansion of the g_λ into monomials via *reverse plane partitions*.

Definition 9 *A* reverse plane partition *of shape λ is a filling of the cells in the Young diagram of λ with positive integers, such that the entries are weakly increasing in rows and columns.*

Example 9 *A reverse plane partition of shape* $(3,2)$ *is* $\begin{array}{|c|c|c} \hline 1 & 2 & \multicolumn{1}{|c}{} \\ \hline 1 & 1 & \multicolumn{1}{|c|}{2} \\ \hline \end{array}$.

Following Lam and Pylyavskyy (and differing from some other conventions) we define the *evaluation* of a reverse plane partition P to be the composition $\alpha = (\alpha_i)_{i\geq 1}$ where α_i is the total number of columns in which i appears. For example, the evaluation of the reverse plane partition in example 9 is $(2,2)$. The collection of all reverse plane partitions of shape λ will be denoted $RPP(\lambda)$ and the collection of all reverse plane partitions of shape λ and evaluation α will be denoted $RPP(\lambda, \alpha)$.

Theorem 1 (Lam-Pylyavskyy) *The polynomials* g_λ *have the expansion*

$$g_\lambda = \sum_{T \in RPP(\lambda)} \mathbf{x}^{ev(P)} .$$

We note that when $|\mu| = |\lambda|$, the entries must be strictly increasing up columns; hence g_λ is equal to s_λ plus lower degree terms.

3 Main result

Before we can state our result, we must define the charge of a set-valued tableau. This is accomplished by defining the reading word of a set-valued tableau.

Definition 10 *The* reading word *of a set-valued tableau* T*, which we denote by* $w(T)$*, is the sequence* $(w_1, w_2, \ldots, w_n)$ *obtained by listing the elements of* T *starting from the top-left corner, and reading each row according to the following procedure and then continuing down the rows. In each row, we first ignore the smallest element of each cell, and read the remaining elements from right to left, and from largest to smallest within each cell. Then we read the smallest element of each cell from left to right, and proceed to the next row.*

Example 10 *The reading word of*

$$\begin{array}{|c|c|c} \hline 3 & 4,5,6 & \multicolumn{1}{|c}{} \\ \hline 1,2 & 2,3 & \multicolumn{1}{|c|}{3} \\ \hline \end{array}$$

is $(6,5,3,4,3,2,1,2,3)$.

We now define the charge of set-valued tableau to be the charge of its reading word. We may now state our main theorem.

Theorem 2 *We have the following expansion of Hall-Littlewood functions into dual Grothendieck functions:*

$$H_\mu[X;t] = \sum_\lambda (-1)^{|\mu|-|\lambda|} \sum_{S \in SVT(\lambda,\mu)} t^{ch(S)} g_\lambda . \tag{9}$$

As this is an extended abstract, we provide only a sketch of the proof. Expanding the s_λ on the right hand side of equation 1 gives

$$H_\mu[X;t] = \sum_\lambda \sum_{T \in SSYT(\lambda,\mu)} t^{ch(T)} s_\lambda \tag{10}$$

$$= \sum_\lambda \sum_{T \in SSYT(\lambda,\mu)} t^{ch(T)} \sum_{Q \in SSYT(\lambda)} \mathbf{x}^{ev(Q)}. \tag{11}$$

If we expand the g_λ in (9) according to theorem 1, we obtain

$$H_\mu[X;t] = \sum_\lambda (-1)^{|\mu|-|\lambda|} \sum_{S \in SVT(\lambda,\mu)} t^{ch(S)} \sum_{R \in RPP(\lambda)} \mathbf{x}^{ev(R)} \;. \tag{12}$$

Now we define $\mathfrak{S}_\mu$ to be the set of pairs (S,R) where S is a set-valued tableau of evaluation μ and R is a reverse plane-partition of the same shape as S. We define the sign of such a pair to be $(-1)^{|\mu|-|\lambda|}$ (where λ is the mutual shape of S and R) and the weight of such a pair to be $t^{ch(S)}\mathbf{x}^{ev(R)}$. Comparing (10) and (12), we see that we can complete the proof by finding a sign-reversing, weight-preserving involution on $\mathfrak{S}_\mu$ whose fixed points are pairs (S,R) where both S and R are semi-standard in the usual sense. We describe this involution below.

3.1 Definition of the involution

Given a pair (S,R) we wish to construct a pair $\iota(S,R) = (S',R')$ of opposite sign and equal weight. We start from the top of both tableaux and work our way down until we find the first row where at least one of the following conditions hold:

1. A cell in S contains more than one element.
2. A cell in R contains the same element as the cell immediately above it.

If no such row exists, the pair is a fixed point. Otherwise, we define $row(S,R)$ to be this row. If only condition (1) holds in $row(S,R)$, we will perform an operation we call *expansion* to define (S',R'). Alternatively, if only condition (2) holds in $row(S,R)$, we perform an operation called *contraction* to define (S',R'). If both conditions hold, we will either expand or contract; the method for determining which will be described following the description of the operations.

We first describe expansion, beginning with the construction of S'. Let $i = row(S,R)$, and define $x(S,R)$ to be the largest element in row i of S which is contained in a multi-element cell (henceforth, *multicell*). Let $\widehat{S}$ be the semi-standard tableau consisting of the rows of S which are strictly above row i. We form S' from S by removing $x(S,R)$ from the multi-cell in row i which contains it, and replacing $\widehat{S}$ with the Schensted insertion $x(S,R) \to \widehat{S}$. Let c be the cell $S' \setminus S$. We form R' from R by placing an empty marker in the cell c and sliding this marker to the south-west using jeu-de-taquin. When the marker reaches row i of R, we replace it with the entry in the cell directly above it. This is R'.

We now describe contraction, beginning with the construction of R'. Again, we let $i = row(S,R)$. We begin by replacing with an empty marker the rightmost cell in row $i+1$ of R which has the same value as the cell immediately below it. Using reverse jeu-de-taquin, slide this marker up and to the right until it exits the diagram. This is R'. Let c be the cell $R \setminus R'$. To construct S', we again let $\widehat{S}$ be the semi-standard tableau consisting of the rows of S which are strictly above row i. Then we perform reverse Schensted insertion on the element in cell c to get a semi-standard tableaux $\widehat{S}'$ and an element $y(S,R)$. Finally, we place $y(S,R)$ inside an existing cell in row i; there will be a unique cell in this row into which we can place $y(S,R)$ so that the result is a valid set-valued row. S' is then defined by placing $\widehat{S}'$ on top of this modified row i, on top of the remaining lower rows of S.

If both conditions 1 and 2 hold in row $row(S,R)$, we must decide whether to perform expansion or contraction. We expand if $x(S,R) \geq y(S,R)$ and contract otherwise. This justifies the claim that, in the contraction case, there is a unique cell in row i into which we can place the element $y(S,R)$; there will never be an element $\geq y(S,R)$ in a multi-cell in row i. Thus $y(S,R)$ can placed (and must be placed) in the rightmost cell such that all of its elements are $\leq y(S,R)$. Such a cell must exist since in S, $y(S,R)$ was an element of row $i+1$.

As this is an extended abstract, we omit the proof that this is a weight-preserving sign-reversing involution. However, we give a simple example below.

Example 11 *The involution ι exchanges the two pairs below:*

$$\left(\begin{array}{|c|c|c|} \hline 3,4 & 5 \\ \cline{1-2} 1 & 1,2 & 2 \\ \hline \end{array} , \begin{array}{|c|c|c|} \hline 1 & 2 \\ \cline{1-2} 1 & 1 & 2 \\ \hline \end{array} \right) \leftrightarrow \left(\begin{array}{|c|c|c|} \cline{1-1} 4 \\ \cline{1-2} 3 & 5 \\ \hline 1 & 1,2 & 2 \\ \hline \end{array} , \begin{array}{|c|c|c|} \cline{1-1} 1 \\ \cline{1-2} 1 & 2 \\ \hline 1 & 1 & 2 \\ \hline \end{array} \right)$$

These pairs have opposite sign, and common weight $t^2 x_1^2 x_2^2$.

References

[Buc02] Anders Skovsted Buch. A Littlewood-Richardson rule for the K-theory of Grassmannians. *Acta Math.*, 189(1):37–78, 2002.

[FK94] Sergey Fomin and Anatol N. Kirillov. Grothendieck polynomials and the Yang-Baxter equation. In *Formal power series and algebraic combinatorics/Séries formelles et combinatoire algébrique*, pages 183–189. DIMACS, Piscataway, NJ, 1994.

[LP07] Thomas Lam and Pavlo Pylyavskyy. Combinatorial Hopf algebras and K-homology of Grassmannians. *Int. Math. Res. Not. IMRN*, (24):Art. ID rnm125, 48, 2007.

[LS78] Alain Lascoux and Marcel-Paul Schützenberger. Sur une conjecture de H. O. Foulkes. *C. R. Acad. Sci. Paris Sér. A-B*, 286(7):A323–A324, 1978.

[LS83] Alain Lascoux and Marcel-Paul Schützenberger. Symmetry and flag manifolds. In *Invariant theory (Montecatini, 1982)*, volume 996 of *Lecture Notes in Math.*, pages 118–144. Springer, Berlin, 1983.

[Mac95] Ian G. Macdonald. *Symmetric Functions and Hall Polynomials*. The Clarendon Press, Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky.

[Sta99] Richard P. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, Cambridge, United Kingdom, 1999.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 35–46

Counting unicellular maps on non-orientable surfaces

Olivier Bernardi[1†] and Guillaume Chapuy[2‡]

[1]*Department of Mathematics, MIT, 77 Massachusetts Avenue, Cambridge MA 02139, USA*
[2]*Department of Mathematics, Simon Fraser University, Burnaby, B.C. V5A 1S6, Canada*

Abstract. A unicellular map is the embedding of a connected graph in a surface in such a way that the complement of the graph is a topological disk. In this paper we give a bijective operation that relates unicellular maps on a non-orientable surface to unicellular maps of a lower topological type, with distinguished vertices. From that we obtain a recurrence equation that leads to (new) explicit counting formulas for non-orientable precubic (all vertices of degree 1 or 3) unicellular maps of fixed topology. We also determine asymptotic formulas for the number of all unicellular maps of fixed topology, when the number of edges goes to infinity. Our strategy is inspired by recent results obtained for the orientable case [Chapuy, PTRF 2010], but significant novelties are introduced: in particular we construct an involution which, in some sense, "averages" the effects of non-orientability.

Résumé. Une carte unicellulaire est le plongement d'un graphe connexe dans une surface, tel que le complémentaire du graphe est un disque topologique. On décrit une opération bijective qui relie les cartes unicellulaires sur une surface non-orientable aux cartes unicellulaires de type topologique inférieur, avec des sommets marqués. On en déduit une récurrence qui conduit à de (nouvelles) formules closes d'énumération pour les cartes unicellulaires précubiques (sommets de degré 1 ou 3) de topologie fixée. On obtient aussi des formules asymptotiques pour le nombre total de cartes unicellulaires de topologie fixée, quand le nombre d'arêtes tend vers l'infini. Notre stratégie est motivée par de récents résultats dans le cas orientable [Chapuy, PTRF, 2010], mais d'importantes nouveautés sont introduites: en particulier, on construit une involution qui, en un certain sens, "moyenne" les effets de la non-orientabilité.

Keywords: One-face map, ribbon graph, non-orientable surface, bijection, involution

1 Introduction

A *map* is an embedding of a connected graph in a (2-dimensional, compact, connected) surface considered up to homeomorphism. By *embedding*, we mean that the graph is drawn on the surface in such a way that the edges do not intersect and the *faces* (connected components of the complementary of the graph) are simply connected. Maps are sometimes referred to as *ribbon graphs, fat-graphs*, and can be defined combinatorially rather than topologically as is recalled in Section 2. A map is *unicellular* if is has a single face. For instance, the unicellular maps on the sphere are the plane trees.

†Supported by ANR project A3.
‡Supported by a CNRS/PIMS postdoctoral fellowship.

In this paper we consider the problem of counting unicellular maps by the number of edges, when the topology of the surface is fixed. In the orientable case, this question has a respectable history. The first formula for the number $\epsilon_g(n)$ of orientable unicellular maps with n edges and $n+1-2g$ vertices (hence genus g) was given by Lehman and Walsh in [WL72], as a sum over the integer partitions of size g. Independently, Harer and Zagier found a simple recurrence formula for the numbers $\epsilon_g(n)$ [HZ86]. Part of their proof relied on expressing the generating function of unicellular maps as a matrix integral. Other proofs of Harer-Zagier's formula were given in [Las01, GN05]. Recently, Chapuy [Cha09], extending previous results for cubic maps [Cha10], gave a bijective construction that relates unicellular maps of a given genus to unicellular maps of a smaller genus, hence leading to a new recurrence equation for the numbers $\epsilon_g(n)$. In particular, the construction in[Cha09] gives a combinatorial interpretation of the fact that for each g the number $\epsilon_g(n)$ is the product of a polynomial in n times the n-th Catalan number.

For non-orientable surfaces, results are more recent. The interpretation of matrix integrals over the Gaussian Orthogonal Ensemble (space of real symmetric matrices) in terms of maps was made explicit in [GJ97]. Ledoux [Led09], by means of matrix integrals and orthogonal polynomials, obtained for unicellular maps on general surfaces a recurrence relation which is similar to the Harer-Zagier one. As far as we know, no direct combinatorial nor bijective technique have successfully been used for the enumeration of a family of non-orientable maps until now.

A unicellular map is *precubic* if it has only vertices of degree 1 and 3: precubic unicellular maps are a natural generalization of binary trees to general surfaces. In this paper, we show that for all $h \in \frac{1}{2}\mathbb{N}$, the number of precubic unicellular maps of size m on the non-orientable surface of Euler Characteristic $2-2h$ is given by an explicit formula, which has the form of a polynomial in m times the mth Catalan number for $h \in \mathbb{N}$, and of a polynomial times 4^m if $h \notin \mathbb{N}$. These formulas, and our main results, are presented in Section 3. Our approach, which is completely combinatorial, is based on two ingredients. The first one, inspired from the orientable case [Cha10, Cha09], is to consider some special vertices called *intertwined nodes*, whose deletion reduces the topological type h of a map. The second ingredient is of a different nature: we show that, among non-orientable maps of a given topology and size, the *average number* of intertwined nodes per map can be determined explicitly. This is done thanks to an *averaging involution*, which is described in Section 4. This enables us to find a simple recurrence equation for the number of precubic unicellular maps by the number of edges and the topological type. As in the orientable case, an important feature of our recurrence is that it is recursive *only* on the topological type, contrarily to equations of the Harer-Zagier type [HZ86, Led09], where also the number of edges vary. It is then easy to iterate the recurrence, to obtain the promised counting formulas for precubic maps.

In the case of *general* (not necessarily precubic) unicellular maps, our approach does not work *exactly*, but it does work, in some sense, *asymptotically*. We obtain, with the same technique, the asymptotic number of non-orientable unicellular maps of fixed topology, when the number of edges tends to infinity. As far as we know, these formulas, and the ones for precubic maps, never appeared before in the literature.

2 Topological considerations

2.1 Classical definitions of surfaces and maps

Surfaces. Our *surfaces* are compact, connected, 2-dimensional manifolds. We consider surfaces up to homeomorphism. For any non-negative integer g, we denote by $\mathbb{S}_g$ the g-torus, that is, the orientable surface obtained by adding g *handles* to the sphere. For any positive *half-integer* h, we denote by $\mathbb{N}_h$

the non-orientable surface obtained by adding $2h$ *cross-caps* to the sphere. Hence, $\mathbb{S}_0$ is the sphere, $\mathbb{S}_1$ is the torus, $\mathbb{N}_{1/2}$ is the projective plane and $\mathbb{N}_1$ is the Klein bottle. The *type* of the surface $\mathbb{S}_h$ or $\mathbb{N}_h$ is the number $h \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\}$ By the theorem of classification, each orientable surface is homeomorphic to one of the $\mathbb{S}_g$ and each non-orientable surface is homeomorphic to one of the $\mathbb{N}_h$ (see e.g. [MT01]).

Maps as graphs embedding. Our *graphs* are finite and undirected; loops and multiple edges are allowed. A *map* is an embedding (without edge-crossings) of a connected graph into a surface, in such a way that the *faces* (connected components of the complement of the graph) are simply connected. Maps are always considered up to homeomorphism. A map is *unicellular* if it has a single face.

Each edge in a map is made of two *half-edges*, obtained by removing its middle-point. The *degree* of a vertex is the number of incident half-edges. A *leaf* is a vertex of degree 1. A *corner* in a map is an angular sector determined by a vertex, and two half-edges which are consecutive around it. The total number of corners in a map equals the number of half-edges which is twice the number of edges. A map is *rooted* if it carries a distinguished half-edge called the *root*, together with a distinguished side of this half-edge. The vertex incident to the root is the *root vertex*. The unique corner incident to the root half-edge and its distinguished side is the *root corner*. *From now on, all maps are rooted.*

The *type* of a map is the type of the underlying surface. If $\mathfrak{m}$ is a map, we let $v(\mathfrak{m})$, $e(\mathfrak{m})$, $f(\mathfrak{m})$ and $h(\mathfrak{m})$ be its numbers of vertices, edges, faces, and its type. These quantities satisfy the *Euler formula*:

$$e(\mathfrak{m}) = v(\mathfrak{m}) + f(\mathfrak{m}) + 2 - 2h(\mathfrak{m}). \tag{1}$$

Maps as graphs with rotation systems and twists. Let G be a graph. To each edge e of G correspond two *half-edges*, each of them incident to an endpoint of e (they are both incident to the same vertex if e is a loop). A *rotation system* for G is the choice, for each vertex v of G, of a cyclic ordering of the half-edges incident to v. We now explain the relation between maps and rotation systems. Our surfaces are locally orientable and an *orientation convention* for a map $\mathfrak{m}$ is the choice of an orientation, called *counterclockwise orientation*, in the vicinity of each vertex. Any orientation convention for the map $\mathfrak{m}$ induces a rotation system on the underlying graph, by taking the counterclockwise ordering of appearance of the half-edges around each vertex. Given an orientation convention, an edge $e = (v_1, v_2)$ of $\mathfrak{m}$ is a *twist* if the orientation conventions in the vicinity of the endpoints v_1 and v_2 are not simultaneously extendable to an orientation of a vicinity of the edge e; this happens exactly when the two sides of e appear in the same order when crossed clockwise around v_1 and clockwise around v_2. Therefore a map together with an orientation convention defines both a rotation system and a subset of edges (the twists). The *flip* of a vertex v consists in inverting the orientation convention at that vertex. This changes the rotation system at v by inverting the cyclic order on the half-edges incident to v, and changes the set of twists by the fact that non-loop edges incident to e become twist if and only if they were not twist (while the status of the other edges remain unchanged). The next lemma is a classical topological result (see e.g. [MT01]).

Lemma 1 *A map (and the underlying surface) is entirely determined by the triple consisting of its (connected) graph, its rotation system, and the subset of its edges which are twists. Conversely, two triples define the same map if and only if one can be obtained from the other by flipping some vertices.*

By the lemma above, we can represent maps of positive types on a sheet of paper as follows: we draw the graph (with possible edge crossings) in such a way that the rotation system at each vertex is given by the counterclockwise order of the half-edges, and we indicate the twists by marking them by a cross (see e.g. Figure 1). The faces of the map are in bijection with the *borders* of that drawing, which are obtained by walking along the edge-sides of the graph, and using the crosses in the middle of twisted edges as ”crosswalks” that change the side of the edge along which one is walking (Figure 1). Observe that the number of faces of the map gives the type of the underlying surface using Euler formula.

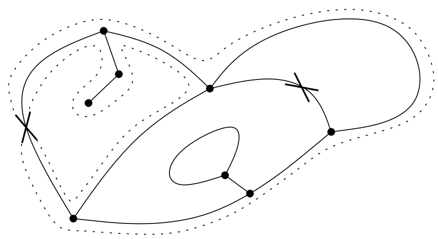


Fig. 1: A representation of a map on the Klein bottle with three faces. The border of one of them is distinguished in dotted lines.

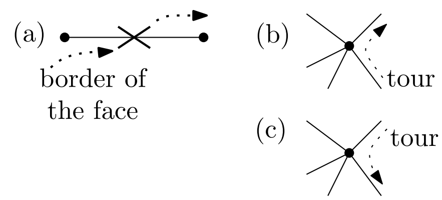


Fig. 2: (a) a twist; (b) a left corner; (c) a right corner.

2.2 Unicellular maps, tours, and canonical rotation system

Tours of unicellular maps. Let $\mathfrak{m}$ be a unicellular map. By definition, $\mathfrak{m}$ has a unique face. The *tour* of the map $\mathfrak{m}$ is done by following the edges of $\mathfrak{m}$ starting from the root corner along the distinguished side of the root half-edge, until returning to the root-corner. Since $\mathfrak{m}$ is unicellular, every corner is visited once during the tour. An edge is said *two-ways* if it is followed in two different directions during the tour of the map (this is always the case on orientable surfaces), and is said *one-way* otherwise. The tour induces an *order of appearance* on the set of corners, for which the root corner is the least element. We denote by $c < d$ if the corner c appears before the corner d along the tour. Lastly, given an orientation convention, a corner is said *left* if it lies on the left of the walker during the tour of map, and *right* otherwise (Figure 2).

Canonical rotation-system. As explained above, the rotation system associated to a map is defined up to the choice of an orientation convention. We now explain how to choose a particular convention which will be well-suited for our purposes. A map is said *precubic* if all its vertices have degree 1 or 3, and its root-vertex has degree 1. Let $\mathfrak{m}$ be a precubic unicellular map. Since the vertices of $\mathfrak{m}$ all have an odd degree, there exists a unique orientation convention at each vertex such that the number of left corners is more than the number of right corners (indeed, by flipping a vertex we exchange its left and right corners). We call *canonical* this orientation convention. From now on, *we will always use the canonical orientation convention*. This defines canonically a rotation system, a set of *twists*, and a set of *left/right corners*. Observe that the root corner is a left corner (as is any corner incident to a leaf) and that vertices of degree 3 are incident to either 2 or 3 left corners. We have the following additional property.

Lemma 2 *In a (canonically oriented) precubic unicellular map, two-ways edges are incident to left corners only and are not twists.*

Proof: Let e be a two-ways edge, and let c_1, c_2 be two corners incident to the same vertex and separated by e (c_1 and c_2 coincide if v has degree 1). Since e is two-ways, the corners c_1, c_2 are either simultaneously left or simultaneously right. By definition of the canonical orientation, they have to be simultaneously left. Thus two-way edges are only incident to left corners. Therefore two-ways edges are not twists since following a twisted edge always leads from a left corner to a right corner or the converse. □

2.3 Intertwined nodes.

We now define a notion of *intertwined node* which generalizes the definition given in [Cha10] for precubic maps on orientable surfaces.

Definition 1 *Let $\mathfrak{m}$ be a (canonically oriented) precubic unicellular map, let v be a vertex of degree 3, and let c_1, c_2, c_3 be the incident corners in counterclockwise order around v, with the convention that c_1 is the first of these corners to appear during the tour of $\mathfrak{m}$.*

- *The vertex v is an* intertwined node *if c_3 appears before c_2 during the tour of $\mathfrak{m}$.*
- *Moreover, we say that v has flavor* **A** *if it is incident to three left corners. Otherwise, v is incident to exactly one right corner, and we say that v is of flavor* **B**, **C**, *or* **D** *respectively, according to whether the right corner is c_1, c_2 or c_3.*

Observe that the definition of the canonical orientation was a prerequisite to define intertwined nodes. We will now show that intertwined nodes are exactly the ones whose deletion decreases the type of the map without disconnecting it.

The *opening* of an intertwined node of a map $\mathfrak{m}$ is the operation consisting in splitting this vertex into three (marked) vertices of degree 1, as in Figure 3. That is, we define a rotation system and set of twists of the embedded graph $\mathfrak{n}$ obtained in this way (we refrain from calling it a map yet, since it is unclear that it is connected) as the rotation system and set of twists inherited from the original map $\mathfrak{m}$.

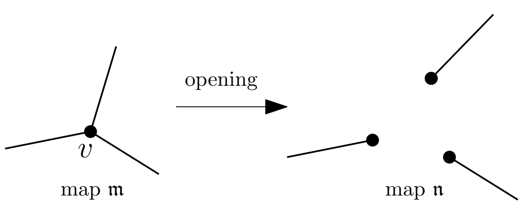


Fig. 3: Opening an intertwined node.

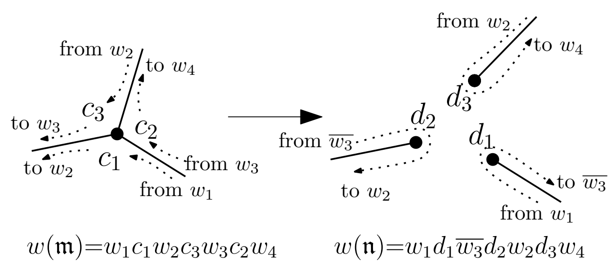


Fig. 4: The tours of $\mathfrak{m}$ and $\mathfrak{n}$, in the case of flavor **B**.

Proposition 1 *Let n be a positive integer and let g be in $\{1, 3/2, 2, 5/2, \ldots\}$. For each flavor F in $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$, the opening operation gives a bijection between the set of precubic unicellular maps with n edges, type h, and a distinguished intertwined node of flavor F, and the set of precubic unicellular maps with n edges, type $h-1$ and three distinguished vertices of degree 1. The converse bijection is called the* gluing of flavor *F.*

Moreover, if a precubic unicellular map $\mathfrak{m}$ is obtained from a precubic unicellular map $\mathfrak{n}$ of lower type by a gluing of flavor F, then $\mathfrak{m}$ is orientable if and only if $\mathfrak{n}$ is orientable and $F = \mathbf{A}$.

We omit the proof of the Proposition. However, let us give a "picture" of what happens, in the case of flavor **B**. If $\mathfrak{m}$ is a unicellular map, and v is an intertwined node of $\mathfrak{m}$, then the sequence of corners appearing during the tour of $\mathfrak{m}$ has the form $w(\mathfrak{m}) = w_1c_1w_2c_3w_3c_2w_4$, where c_1, c_2, c_3 are as in Definition 1, and w_1, w_2, w_3, w_4 are sequences of corners. Now, let $\mathfrak{n}$ be the map obtained by opening $\mathfrak{m}$ at v. If v has flavor **B**, then by following the edges of the map $\mathfrak{n}$, starting from the root, one gets the sequence of corners $w(\mathfrak{n}) = w_1d_1\overline{w}_3d_2w_2d_3w_4$, where $\overline{w_3}$ is the *mirror word* of w_3, as can be seen from Figure 4 (we used three new letters d_1, d_2, d_3 for the three corners of degree 1 appearing after the opening). Since this sequence contains all the corners of $\mathfrak{n}$, we know that $\mathfrak{n}$ is a unicellular map, and since it has two more vertices than $\mathfrak{m}$, its type is $h(\mathfrak{n}) = h(\mathfrak{m}) - 1$ (by Euler's formula).

Conversely, given a unicellular map $\mathfrak{n}$ with three distinguished leaves d_1, d_2, d_3, the gluing of flavor **B** can be defined by identifying these three vertices to a single vertex v, and then choosing the rotation system and the twisted edges at v appropriately to ensure that the resulting map $\mathfrak{m}$ is unicellular, and that v is an intertwined node of flavor **B** in $\mathfrak{m}$.

The last statement of the Proposition is a consequence of the fact that a precubic unicellular map is orientable if and only if it has left-corners only in its canonical orientation.

3 Main results.

3.1 The number of precubic unicellular maps.

In this section, we present our main results, which rely on two facts. The first one is Proposition 1, which enables us to express the number of precubic unicellular maps of type h carrying a distinguished intertwined node in terms of the number of maps of a smaller type. The second one is the fact that, among maps of type h and fixed size, the average number of intertwined nodes in a map is $2h-1$. This last fact, which is technically the most difficult part of this paper, relies on the existence of an "averaging involution", which will be described in Section 4.

Let $h \geq 1$ be an element of $\frac{1}{2}\mathbb{N}$, and let $m \geq 1$ be an integer. Given m and h, we let $n = 2m + \mathbb{1}_{h \in \mathbb{N}}$, and we let $\mathcal{O}_h(m)$ and $\mathcal{N}_h(m)$, respectively, be the sets of orientable and non-orientable precubic unicellular maps of type h with n edges. We let $\xi_h(m)$ and $\eta_h(m)$, respectively, be the cardinality of $\mathcal{O}_h(m)$ and $\mathcal{N}_h(m)$.

In order to use Proposition 1, we first need the following easy consequence of Euler's formula:

Lemma 3 *Let $l \in \frac{1}{2}\mathbb{N}$ and let $\mathfrak{m}$ be a precubic unicellular map of type l with $n = 2m + \mathbb{1}_{l \in \mathbb{N}}$ edges. Then $\mathfrak{m}$ has $m + (-1)^{2l} - 3\lfloor l \rfloor$ non-root leaves, where $\lfloor l \rfloor = l - \frac{1}{2}\mathbb{1}_{l \notin \mathbb{N}}$ denotes the integer part of l.*

From the lemma and Proposition 1, the number $\eta_h^{\text{inter}}(m)$ of *non-orientable* unicellular precubic maps of type h with n edges carrying a distinguished intertwined node equals:

$$\eta_h^{\text{inter}}(m) = 4\binom{m' - 3\lfloor h-1 \rfloor}{3}\eta_{h-1}(m) + 3\binom{m' - 3\lfloor h-1 \rfloor}{3}\xi_{h-1}(m), \tag{2}$$

where $m' = m + (-1)^{2h}$. Here, the first term accounts for intertwined nodes obtained by gluing three leaves in a *non-orientable* map of type $h-1$ (in which case the flavor of the gluing can be either **A**, **B**, **C** or **D**), and the second term corresponds to the case where the starting map of type $h-1$ is *orientable* (in which case the gluing has to be of flavor **B**, **C** or **D** to destroy the orientability).

The keystone of this paper, which will be discussed in the next section, is the following result:

Proposition 2 *There exists and involution Φ of $\mathcal{N}_h(m)$ such that for all maps $\mathfrak{m} \in \mathcal{N}_h(m)$, the total number of intertwined nodes in the maps $\mathfrak{m}$ and $\Phi(\mathfrak{m})$ is $4h-2$. In particular, the average number of intertwined nodes of elements of $\mathcal{N}_h(m)$ is $(2h-1)$, and one has $\eta_h^{\text{inter}}(m) = (2h-1)\eta_h(m)$.*

It is interesting to compare Proposition 2 with the analogous result in [Cha10]: in the orientable case, *each* map of genus h has exactly $2h$ intertwined nodes, whereas here the quantity $(2h-1)$ is only an *average value*. For example, Figure 5 shows two maps on the Klein bottle ($h = 1$) which are related by the involution Φ: they have respectively 2 and 0 intertwined nodes.

As a direct corollary of Proposition 2 and Equation (2), we can state our main result:

Theorem 1 *The numbers $\eta_h(m)$ of non-orientable precubic unicellular maps of type h with $2m + \mathbb{1}_{h\in\mathbb{N}}$ edges obey the following recursion:*

$$(2h-1) \cdot \eta_h(m) = 4\binom{m' - 3\lfloor h-1 \rfloor}{3}\eta_{h-1}(m) + 3\binom{m' - 3\lfloor h-1 \rfloor}{3}\xi_{h-1}(m), \tag{3}$$

where $m' = m+(-1)^{2h}$, and where $\xi_h(m)$ is the number of orientable precubic unicellular maps of genus h with $2m + \mathbb{1}_{h\in\mathbb{N}}$ edges, which is 0 if $h \notin \mathbb{N}$, and is given by the following formula otherwise [Cha09]:

$$\xi_h(m) = \frac{1}{(2h)!!}\binom{m+1}{3,3,\ldots,3,m+1-3h}\mathrm{Cat}(m) = \frac{(2m)!}{12^h h! m! (m+1-3h)!}. \tag{4}$$

The theorem implies explicit formulas for the numbers $\eta_h(m)$, as shown by the two next corollaries:

Corollary 1 (the case $h \in \mathbb{N}$) *Let $h \in \mathbb{N}$ and $m \in \mathbb{N}$, $m \geq 3h-1$. Then the number of non-orientable precubic unicellular maps of type h with $2m+1$ edges equals:*

$$\eta_h(m) = c_h\binom{m+1}{3,3,\ldots,3,m+1-3h}\mathrm{Cat}(m) = \frac{c_h \cdot (2m)!}{6^h m!(m+1-3h)!} \tag{5}$$

where $c_h = 3 \cdot 2^{3h-2}\dfrac{h!}{(2h)!}\displaystyle\sum_{l=0}^{h-1}\binom{2l}{l}16^{-l}$.

Corollary 2 (the case $h \notin \mathbb{N}$) *Let $h \in \frac{1}{2} + \mathbb{N}$ and $m \in \mathbb{N}$, $m \geq 3\lfloor h \rfloor + 1$. Then the number of non-orientable precubic unicellular maps of type h with $2m$ edges equals:*

$$\begin{aligned}\eta_h(m) &= \frac{4^{\lfloor h \rfloor}}{(2h-1)(2h-3)\ldots 1}\binom{m-1}{3,3,\ldots,3,m-1-3\lfloor h \rfloor} \times \eta_{1/2}(m) \\ &= \frac{4^{m+\lfloor h \rfloor -1}(m-1)!}{6^{\lfloor h \rfloor}(2h-1)!!(m-1-3\lfloor h \rfloor)!}.\end{aligned}$$

Proof of Corollary 1: It follows by induction and Equations (3) and (4) that the statement of Equation (5) holds, with the constant c_h defined by the recurrence $c_0 = 0$ and $c_h = a_{h-1} + b_{h-1}c_{h-1}$, with $a_{h-1} = \frac{3}{2^{h-1}(h-1)!(2h-1)}$ and $b_{h-1} = \frac{4}{2h-1}$. The solution of this recurrence is $c_h = \sum_{l=0}^{h-1} a_l b_{l+1} b_{l+2} \ldots b_{h-1}$. Now, by definition, we have $a_l b_{l+1} b_{l+2} \ldots b_{h-1} = \dfrac{3 \cdot 4^{h-1-l}}{2^l l!(2l+1)(2l+3)(2l+5)\ldots(2h-1)}$. Using the expression $\dfrac{1}{(2l+1)(2l+3)\ldots(2h-1)} = \dfrac{2^h h!(2l)!}{(2h)!2^l l!}$ and reporting it in the sum gives the expression of c_h given in Corollary 1. □

Proof of Corollary 2: Since for non-integer h we have $\xi_{h-1}(m) = 0$, the first equality is a direct consequence of an iteration of the theorem. Therefore the only thing to prove is that $\eta_{1/2}(m) = 4^{m-1}$.

This can be done easily by induction via an adaptation of Rémy's bijection [Rém85], as follows. For $m = 1$, we have $\eta_{1/2}(m) = 1$, since the only precubic projective unicellular map with two edges is "the twisted loop with a hanging leaf". For the induction step, observe that precubic projective unicellular maps with one distinguished non-root leaf are in bijection with precubic projective unicellular maps with one leaf less and a distinguished edge-side: too see that, delete the distinguished leaf, transform the remaining vertex of degree 2 into an edge, and remember the side of that edge on which the original leaf was attached. Since a projective precubic unicellular map with $2k$ edges has $k-1$ non-root leaves and $4k$ edge-sides, we obtain for all $m \geq 1$ that $m\,\eta_{1/2}(m+1) = 4m\,\eta_{1/2}(m)$, and the result follows. $\square$

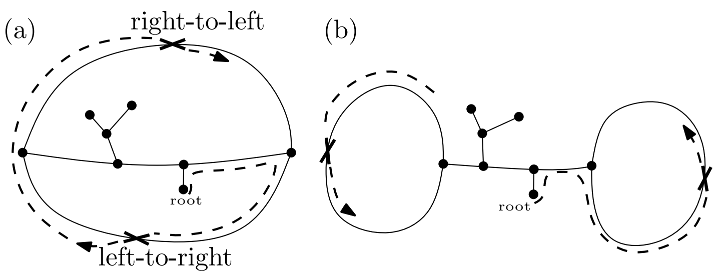


Fig. 5: Two maps on the Klein Bottle. (a) $T_{\mathrm{LR}}(\mathfrak{m}) = 1$, $T_{\mathrm{RL}}(\mathfrak{m}) = 1$; (b) $T_{\mathrm{LR}}(\mathfrak{m}) = 2$, $T_{\mathrm{RL}}(\mathfrak{m}) = 0$.

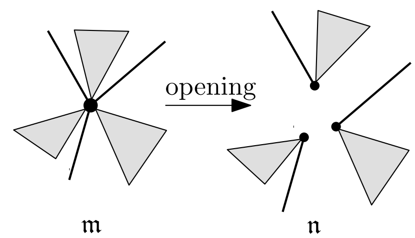


Fig. 6: The opening, in the case of dominant unicellular maps.

3.2 The asymptotic number of rooted unicellular maps.

Though our results do not apply to the general case of *all* unicellular maps of given type (i.e., not necessarily precubic), they do hold, in some sense, *asymptotically*. This is what we explain in this section.

If $\mathfrak{m}$ is a unicellular map, its *core* is the map obtained by deleting recursively all the leaves of $\mathfrak{m}$, until having only vertices of degree 2 or more left. Therefore the core is a unicellular map formed by chains of vertices of degree 2 attached together at vertices of degree at least 3. The *scheme* of $\mathfrak{m}$ is the map obtained by replacing each of these chains by an edge. Hence, in the scheme, all vertices have degree at least 3. We say that a unicellular map is *dominant* if all the vertices of its scheme have degree 3. This terminology, borrowed from [Cha10], comes from the next proposition.

Proposition 3 ([CMS09, BR09]) *Let $h \in \frac{1}{2}\mathbb{N}$. Then, among non-orientable unicellular maps of type h with n edges, the proportion of maps which are dominant tends to 1 when n tends to infinity.*

The idea behind that proposition is the following. Given a scheme $\mathfrak{s}$, one can easily compute the generating series of all unicellular maps of scheme $\mathfrak{s}$, by observing that these maps are obtained by substituting each edge of the scheme with a "branch of tree". From that, it follows that this generating series has a unique principal singularity at $z = \frac{1}{4}$, with dominating term $(1-4z)^{-e(\mathfrak{s})/2-1}$, up to a multiplicative constant. Therefore, the schemes with the greatest contribution are those which have the maximal number of edges, which for a given type, is achieved by schemes whose all vertices have degree 3.

Now, most of the combinatorics defined in this paper still apply to dominant unicellular maps. Given a dominant map $\mathfrak{m}$ of type h and scheme $\mathfrak{s}$, and v an intertwined node of $\mathfrak{s}$, we can define the opening operation of $\mathfrak{m}$ at v by splitting the vertex v in three, and deciding on a convention on the redistribution of the three "subtrees" attached to the scheme at this point (Figure 6): one obtains a dominant map $\mathfrak{n}$ of type $h-1$ with three distinguished vertices. These vertices are not *any* three vertices: they have to be in "general position" in $\mathfrak{n}$ (i.e., they cannot be part of the core, and none can lie on a path from one to another), but again, in the asymptotic case this does not make a big difference: when n tends to infinity, the proportion of triples of vertices which are in "general position" tends to 1. We do not state here the asymptotic estimates that can make the previous claims precise (they can be copied almost verbatim from the orientable case [Cha10]), but rather we state now our asymptotic theorem:

Theorem 2 *Let $\kappa_h(n)$ be the number of non-orientable rooted unicellular maps of type h with n edges. Then one has, when n tends to infinity:*

$$(2h-1)\kappa_h(n) \sim 4\frac{n^3}{3!}\kappa_{h-1}(n) + 3\frac{n^3}{3!}\epsilon_{h-1}(n)$$

where $\epsilon_h(n)$ denotes the number of orientable rooted unicellular maps of genus h with n edges.

Using that $\epsilon_h(n) = 0$ if $h \notin \mathbb{N}$, that $\epsilon_h(n) \sim \frac{1}{12^h h! \sqrt{\pi}} n^{3h-\frac{3}{2}}$ otherwise, and that $\kappa_{1/2}(n) \sim \frac{1}{2}4^n$ [BCR88], we obtain:

Corollary 3 *Let $h \in \frac{1}{2}\mathbb{N}$. Then one has, when n tends to infinity:*

$$\kappa_h(n) \sim \frac{c_h}{\sqrt{\pi}6^h} n^{3h-\frac{3}{2}} 4^n \ \text{ if } h \in \mathbb{N} \quad , \quad \kappa_h(n) \sim \frac{4^{\lfloor h \rfloor}}{2 \cdot 6^{\lfloor h \rfloor}(2h-1)!!} n^{3h-\frac{3}{2}} 4^n \ \text{ if } h \notin \mathbb{N}.$$

where the constant c_h is defined in Corollary 1.

4 The average number of intertwined nodes

In this section we prove Proposition 2, and in particular the key result that the average number of intertwined nodes per map, among precubic unicellular maps of type h and size m is $(2h-1)$:

$$\eta_h^{\text{inter}}(m) = (2h-1)\eta_h(m). \tag{6}$$

Let us emphasize the fact that the number of intertwined nodes is not a constant over the set of unicellular precubic maps of given type and number of edges. For instance among the six maps with 5 edges on the Klein bottle $\mathbb{N}_1$, three maps have 2 intertwined nodes, and three maps have none; see Figure 7. As stated in Proposition 2, our strategy to prove Equation (6) is to exhibit a bijection Φ from the set $\mathcal{N}_h(m)$ to itself, such that for any given map $\mathfrak{m}$, the total number of intertwined nodes in the maps $\mathfrak{m}$, $\Phi(\mathfrak{m})$ is $4h(\mathfrak{m}) - 2$. Observe from Figure 7 that the involution Φ cannot be a simple re-rooting of the map $\mathfrak{m}$.

Before defining the mapping Φ, we relate the number of intertwined nodes of a map to certain properties of its twists. Let $\mathfrak{m}$ be a (canonically oriented) precubic map, and let e be an an edge of $\mathfrak{m}$ which is a twist. Let c be the corner incident to e which appears first in the tour of $\mathfrak{m}$. We say that e is *left-to-right* if c is a left-corner, and that it is *right-to-left* otherwise (see Figure 5). In other words, the twist e is left-to-right if it changes the side of the corners from left, to right, when it is crossed for the first time in the tour of the map (and the converse is true for right-to-left twists). We omit the proof of the next lemma:

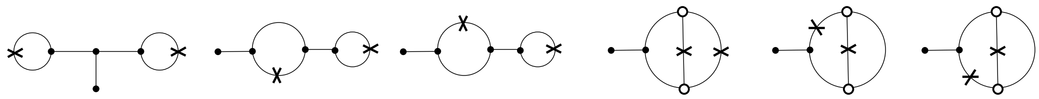


Fig. 7: The precubic unicellular maps with 5 edges on the Klein bottle (the root in the unique leaf corner). Intertwined nodes are indicated as white vertices.

Lemma 4 *Let $\mathfrak{m}$ be a precubic unicellular map of type $h(\mathfrak{m})$, considered in its canonical orientation. Then its numbers $\tau(\mathfrak{m})$ of intertwined nodes, $T_{\mathrm{LR}}(\mathfrak{m})$ of left-to-right twists, and $T_{\mathrm{RL}}(\mathfrak{m})$ of right-to-left twists are related by the formula:*

$$2h(\mathfrak{m}) = \tau(\mathfrak{m}) + T_{\mathrm{LR}}(\mathfrak{m}) - T_{\mathrm{RL}}(\mathfrak{m}). \tag{7}$$

We now define the promised mapping Φ averaging the number of intertwined nodes. Let $\mathfrak{m}$ be a unicellular precubic map on a non-orientable surface. We consider the canonical orientation convention for the map $\mathfrak{m}$, which defines a rotation system and set of twists. The set of twists is non-empty since the map $\mathfrak{m}$ lives on a non-orientable surface. By cutting every twist of $\mathfrak{m}$ at their middle point, one obtains a graph together with a rotation system and some *dangling half-edges* that we call *buds*. The resulting embedded graph with buds, which we denote by $\hat{\mathfrak{m}}$, can have several connected components and each component (which is a map with buds) can have several faces; see Figure 8. We set a convention for the direction in which one *turns around a face* of $\hat{\mathfrak{m}}$: the edges are followed in such a way that every corner is left (this is possible since $\hat{\mathfrak{m}}$ has no twist). For any bud b of $\hat{\mathfrak{m}}$, we let $\sigma(b)$ be the bud following b when turning around the face of $\hat{\mathfrak{m}}$ containing b. Clearly, the mapping σ is a permutation on the set of buds. We now define $\Phi(\mathfrak{m})$ to be the graph with rotation system and twists obtained from $\hat{\mathfrak{m}}$ by gluing together into a twist the buds $\sigma(b)$ and $\sigma(b')$ for every pair of buds b, b' forming a twist of $\mathfrak{m}$. The mapping Φ is represented in Figure 8.

Before proving that $\Phi(\mathfrak{m})$ is a unicellular map, we set some additional notations. We denote by k the number of twists of $\mathfrak{m}$ and we denote by $w(\mathfrak{m}) = w_1 w_2 \cdots w_{2k+1}$ the sequence of corners encountered during the tour of $\mathfrak{m}$, where the subsequences w_i and w_{i+1} are separated by the traversal of a twist for $i = 1 \ldots 2k$. Observe that corners in w_i are left corners of $\mathfrak{m}$ if i is odd, and right corners if i is even (since following a twist leads from a left to a right corner or the converse). Hence, the sequence of corners encountered between two buds around a face of $\hat{\mathfrak{m}}$ are one of the sequences $w'_1, w'_2, \ldots, w'_{2k}$, where $w'_1 = w_{2k+1} w_1$, and for $i > 1$, $w'_i = w_i$ if i is odd and $w'_i = \overline{w}_i$ otherwise (where $\overline{w}_i$ is the *mirror* sequence of w_i obtained by reading w_i backwards). We identify the buds of $\hat{\mathfrak{m}}$ (i.e. the half-twists of $\mathfrak{m}$ or $\hat{\mathfrak{m}}$) with the integers in $\{1, \ldots, 2k\}$ by calling i the bud following the sequence of corners w'_i around the faces of $\hat{\mathfrak{m}}$. The permutation σ can then be considered as a permutation of $\{1, \ldots, 2k\}$ and we denote $r = \sigma^{-1}(1)$. The map in Figure 8 gives $\sigma = (1, 8, 13, 2, 9, 14, 3, 10)(4, 11, 6, 5)(7, 12)$ and $r = 10$.

Lemma 5 *The embedded graph $\Phi(\mathfrak{m})$ is a unicellular map. Moreover, the rotation system and set of twists of $\Phi(m)$ inherited from $\mathfrak{m}$ correspond to the canonical orientation convention of $\Phi(\mathfrak{m})$. Lastly, the sequence of corners encountered during the tour of $\Phi(\mathfrak{m})$ reads $v_1 v_2 \ldots v_{2k+1}$, where the subsequences v_i separated twist traversals are given by $v_i = w_{\sigma(r+1-i)}$ for $i = 1, \ldots, r$, $v_i = w_{\sigma(2n+r+1-i)}$ for $i = r+1, \ldots, 2k$, and $v_{2k+1} = w_{2k+1}$.*

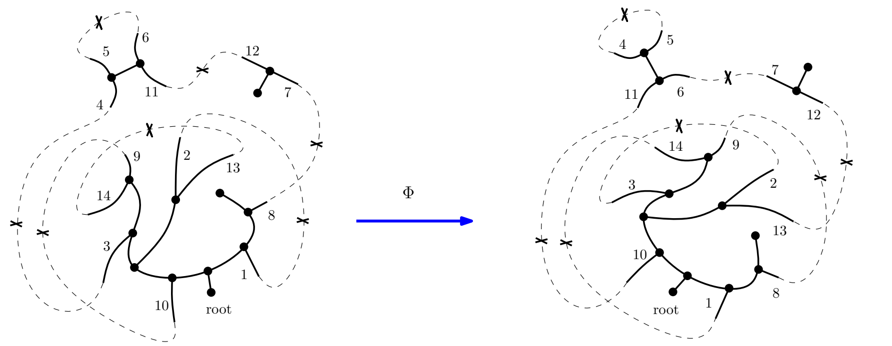


Fig. 8: A unicellular map $\mathfrak{m}$ and its image by the mapping Φ. The twists are indicated by (partially) dotted lines, while the map $\hat{\mathfrak{m}}$ is represented in solid lines.

Lemma 6 *Let m be a positive integer and h be in $\{1/2, 1, 3/2, \ldots\}$. The mapping Φ is a bijection from the set $\mathcal{N}_h(m)$ to itself. Moreover, for every map $\mathfrak{m}$ in $\mathcal{N}_h(m)$, the total number of intertwined nodes in the maps $\mathfrak{m}$ and $\Phi(\mathfrak{m})$ is $4h-2$.*

Proof of Lemma 6: Clearly, the maps $\mathfrak{m}$ and $\Phi(\mathfrak{m})$ have the same number of edges and vertices. Hence, they have the same type by Euler formula. Moreover, they both have $k > 0$ twists (for their canonical convention) hence are non-orientable. Thus, Φ maps the set $\mathcal{N}_h(m)$ to itself. To prove the bijectivity (i.e. injectivity) of Φ, observe that for any map $\mathfrak{m}$, the embedded graphs $\hat{\mathfrak{m}}$ and $\widehat{\Phi(\mathfrak{m})}$ are equal; this is because the canonical rotation system and set of twists of $\mathfrak{m}$ and $\Phi(\mathfrak{m})$ coincide. In particular, the permutation σ on the half-twists of $\mathfrak{m}$ can be read from $\Phi(\mathfrak{m})$. Hence, the twists of $\mathfrak{m}$ are easily recovered from those of $\Phi(\mathfrak{m})$: the buds i and j form a twist of $\mathfrak{m}$ if $\sigma(i)$ and $\sigma(j)$ form a twist of $\Phi(\mathfrak{m})$.

We now proceed to prove that the total number of intertwined nodes in $\mathfrak{m}$ and $\Phi(\mathfrak{m})$ is $4h-2$. By Lemma 4, this amounts to proving that $T_{\text{LR}}(\mathfrak{m}) - T_{\text{RL}}(\mathfrak{m}) + T_{\text{LR}}(\Phi(\mathfrak{m})) - T_{\text{RL}}(\Phi(\mathfrak{m})) = 2$. Since $\mathfrak{m}$ and $\Phi(\mathfrak{m})$ both have k twists, $T_{\text{LR}}(\mathfrak{m}) - T_{\text{RL}}(\mathfrak{m}) + T_{\text{LR}}(\Phi(\mathfrak{m})) - T_{\text{RL}}(\Phi(\mathfrak{m})) = 2(T_{\text{LR}}(\mathfrak{m}) + T_{\text{LR}}(\Phi(\mathfrak{m})) - k)$. Hence, we have to prove $T_{\text{LR}}(\mathfrak{m}) + T_{\text{LR}}(\Phi(\mathfrak{m})) = k+1$.

Let i be a bud of $\hat{\mathfrak{m}}$, let t be the twist of $\mathfrak{m}$ containing i, and let c, c' be the corners preceding and following i in counterclockwise order around the vertex incident to i. By definition, the twist t of $\mathfrak{m}$ is left-to-right if and only if c appears before c' during the tour of $\mathfrak{m}$. Given that the corners c and c' belong respectively to the subsequences w_i and $w_{\sigma(i)}$ (except if $i = r$ in which case $\sigma(i) = 1$ and c' is in w_{2k+1}), the twist t is left-to right if and only if $i < \sigma(i)$ or $i = r$.

Before going on, let us introduce a notation: for an integer i we denote by $\bar{i}$ the representative of i modulo $2k$ belonging to $\{1, \ldots, 2k\}$. Let us now examine under which circumstances the bud $\sigma(i)$ is part of a left-to-right twist of $\Phi(\mathfrak{m})$. The corners d and d' preceding and following the bud $\sigma(i)$ in counterclockwise order around the vertex incident to $\sigma(i)$ belong respectively to $w_{\sigma(i)}$ and $w_{\sigma\sigma(i)}$ (except if $\sigma(i) = r$, in which case $\sigma\sigma(i) = 1$ and c' belongs to w_{2k+1}). By Lemma 5, $w_{\sigma(i)} = v_{\overline{r+1-i}}$ for $i = 1 \ldots 2k$. Therefore, the twist t' of $\Phi(\mathfrak{m})$ containing $\sigma(i)$ is left-to-right if and only if $\overline{r+1-i} < \overline{r+1-\sigma(i)}$ or $\sigma(i) = r$.

The two preceding points gives the number $T_{\text{LR}}(\mathfrak{m}) + T_{\text{LR}}(\Phi(\mathfrak{m}))$ of left-to right twists as

$$T_{\text{LR}}(\mathfrak{m}) + T_{\text{LR}}(\Phi(\mathfrak{m})) = 1 + \tfrac{1}{2} \sum_{i=1}^{2k} \delta(i),$$

where $\delta(i) = \mathbb{1}_{i<\sigma(i)} + \mathbb{1}_{\overline{r+1-i}<\overline{r+1-\sigma(i)}}$ is the sum of two indicator functions (the factor $1/2$ accounts for the fact that a twist has two halves). The contribution $\delta(i)$ is equal to 2 if $i \leq r < \sigma(i)$, 0 if $\sigma(i) \leq r < i$, and 1 otherwise. Finally, there are as many integers i such that $i \leq r < \sigma(i)$ as integers such that $\sigma(i) \leq r < i$ (true for each cycle of σ). Thus, $\sum_{i=1}^{2k} \delta(i) = 2k$, and $T_{\mathrm{LR}}(\mathfrak{m}) + T_{\mathrm{LR}}(\Phi(\mathfrak{m})) = k + 1$. $\square$

The last lemma is sufficient to establish Equation (6), and the enumerative results of Section 3. However, Proposition 2 was saying a little bit more, namely that the bijection Φ can be chosen as an *involution*:

Proof of Proposition 2: Observe that, as we defined it, the bijection Φ is not an involution. But one can easily define an involution from Φ, as the mapping acting as Φ on elements $\mathfrak{m}$ of $\mathcal{N}_h(m)$ such that $\tau(\mathfrak{m}) > 2h - 1$, acting as Φ^{-1} if $\tau(\mathfrak{m}) < 2h - 1$, and as the identity if $\tau(\mathfrak{m}) = 2h - 1$. $\square$

References

[BCR88] E. A. Bender, E. R. Canfield, and R. W. Robinson. The enumeration of maps on the torus and the projective plane. *Canad. Math. Bull.*, 31:257–271, 1988.

[BR09] O. Bernardi and J. Rué. Enumerating simplicial decompositions of surfaces with boundaries. *arXiv:0901.1608*, 2009.

[Cha09] G. Chapuy. A new combinatorial identity for unicellular maps, via a direct bijective approach. *DMTCS (FPSAC '09) – Preliminary long version at www.lix.polytechnique.fr/˜chapuy*, 2009.

[Cha10] G. Chapuy. The structure of unicellular maps, and a connection between maps of positive genus and planar labelled trees. *Probability Theory and Related Fields*, 147(3):415–447, 2010.

[CMS09] G. Chapuy, M. Marcus, and G. Schaeffer. A bijection for rooted maps on orientable surfaces. *SIAM J. Discrete Math.*, 23(3):1587–1611, 2009.

[GJ97] I. Goulden and D.M. Jackson. Maps in locally orientable surfaces and integrals over real symmetric matrices. *Canad. J. Math*, 48, 1997.

[GN05] I. Goulden and A. Nica. A direct bijection for the Harer-Zagier formula. *JCTA*, 2005.

[HZ86] J. Harer and D. Zagier. The Euler characteristic of the moduli space of curves. *Invent. Math.*, 85(3):457–485, 1986.

[Las01] B. Lass. Démonstration combinatoire de la formule de Harer-Zagier. *C. R. Acad. Sci. Paris*, 333:155–160, 2001.

[Led09] M. Ledoux. A recursion formula for the moments of the Gaussian orthogonal ensemble. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):754–769, 2009.

[MT01] B. Mohar and C. Thomassen. *Graphs on surfaces*. J. Hopkins Univ. Press, 2001.

[Rém85] J.-L. Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire. *RAIRO Inform. Théor.*, 19(2):179–195, 1985.

[WL72] T. R. S. Walsh and A. B. Lehman. Counting rooted maps by genus. I. *J. Combinatorial Theory Ser. B*, 13:192–218, 1972.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 47–58

A canonical basis for Garsia-Procesi modules

Jonah Blasiak[1†]

[1]*The University of Chicago, Department of Computer Science, 1100 East 58th Street, Chicago, IL 60637, USA*

Abstract. We identify a subalgebra $\widehat{\mathscr{H}}_n^+$ of the extended affine Hecke algebra $\widehat{\mathscr{H}}_n$ of type A. The subalgebra $\widehat{\mathscr{H}}_n^+$ is a u-analogue of the monoid algebra of $\mathcal{S}_n \ltimes \mathbb{Z}_{\geq 0}^n$ and inherits a canonical basis from that of $\widehat{\mathscr{H}}_n$. We show that its left cells are naturally labeled by tableaux filled with positive integer entries having distinct residues mod n, which we term *positive affine tableaux* (PAT).

We then exhibit a cellular subquotient $\mathscr{R}_{1^n}$ of $\widehat{\mathscr{H}}_n^+$ that is a u-analogue of the ring of coinvariants $\mathbb{C}[y_1, \dots, y_n]/(e_1, \dots, e_n)$ with left cells labeled by PAT that are essentially standard Young tableaux with cocharge labels. Multiplying canonical basis elements by a certain element $\pi \in \widehat{\mathscr{H}}_n^+$ corresponds to rotations of words, and on cells corresponds to cocyclage. We further show that $\mathscr{R}_{1^n}$ has cellular quotients $\mathscr{R}_\lambda$ that are u-analogues of the Garsia-Procesi modules R_λ with left cells labeled by (a PAT version of) the λ-catabolizable tableaux.

Résumé. On définit une sous-algèbre $\widehat{\mathscr{H}}_n^+$ de l'extension affine de l'algèbre de Hecke $\widehat{\mathscr{H}}_n$ de type A. La sous-algèbre $\widehat{\mathscr{H}}_n^+$ est u-analogue à l'algèbre monoïde de $\mathcal{S}_n \ltimes \mathbb{Z}_{\geq 0}^n$ et hérite d'une base canonique de $\widehat{\mathscr{H}}_n$. On montre que ses cellules gauches sont naturellement classées par des tableaux remplis d'entiers naturels ayant chacun des restes différents modulo n, que l'on nomme *Positive Affine Tableaux* (PAT).

On montre ensuite qu'un sous-quotient cellulaire $\mathscr{R}_{1^n}$ de $\widehat{\mathscr{H}}_n^+$ est une u-analogue de l'anneau des co-invariants $\mathbb{C}[y_1, \dots, y_n]/(e_1, \dots, e_n)$ avec des cellules gauches classées PAT qui sont essentiellement des tableaux de Young standards avec des labels cochargés. Multiplier les éléments de la base canonique par un certain élément $\pi \in \widehat{\mathscr{H}}_n^+$ correspond à des rotations de mots, et par rapport aux cellules cela correspond à un cocyclage. Plus loin, on montre que $\mathscr{R}_{1^n}$ a pour quotients cellulaires $\mathscr{R}_\lambda$ qui sont u- analogues aux modules de Garsia-Procesi R_λ avec des cellules gauches définies par (une version PAT) des tableaux λ-catabolisable.

Keywords: Garsia-Procesi modules, affine Hecke algebra, canonical basis, symmetric group, k-atoms

1 Introduction

It is well-known that the ring of coinvariants $R_{1^n} = \mathbb{C}[y_1, \dots, y_n]/(e_1, \dots, e_n)$, thought of as a $\mathbb{C}\mathcal{S}_n$-module with $\mathcal{S}_n$ acting by permuting the variables, is a graded version of the regular representation. However, how a decomposition of this module into irreducibles is compatible with multiplication by the y_i remains a mystery.

A precise question one can ask along these lines goes as follows. Let $E \subseteq R_d$ be an $\mathcal{S}_n$-irreducible, where R_d is the d-th graded part of the polynomial ring $R = \mathbb{C}[y_1, \dots, y_n]$. Suppose that the isotypic

[†]Is currently an NSF postdoctoral fellow. This research was partially conducted during the period the author was employed by the Clay Mathematics Institute as a Liftoff Fellow.

component of R_d containing E is E itself. Then define $I \subseteq R$ to be the sum of all homogeneous ideals $J \subseteq R$ that are left stable under the $\mathcal{S}_n$-action and satisfy $J \cap E = 0$. The quotient R/I contains E as the unique $\mathcal{S}_n$-irreducible of top degree d. It is natural to ask

What is the graded character of R/I?

The most familiar examples of such quotients are the Garsia-Procesi modules R_λ (see [5]), which correspond to the case that E is of shape λ and $d = n(\lambda) = \sum_i (i-1)\lambda_i$; refer to this representation $E \subseteq R_{n(\lambda)}$ as the *Garnir representation of shape* λ or, more briefly, G_λ. Combining the work of Hotta-Springer (see [6]) and Lascoux [10] (see also [15]) gives the Frobenius series

$$\mathscr{F}_{R_\lambda}(t) = \sum_{\substack{T \in SYT \\ \mathrm{ctype}(T) \trianglerighteq \lambda}} t^{\mathrm{cocharge}(T)} s_{\mathrm{sh}(T)}, \tag{1}$$

where $\mathrm{ctype}(T)$ is the catabolizability of T (see §4).

Though this interpretation of the character of R_λ has been known for some time, the only proofs were difficult and indirect. One of the goals of this research, towards which we have been partially successful, was to give a more transparent explanation of the appearance of catabolism in the combinatorics of the coinvariants.

More recent work suggests that there are other combinatorial mysteries hiding in the ring of coinvariants. We strongly suspect that modules with graded characters corresponding to the k-atoms of Lascoux, Lapointe, and Morse [9] and a generalization of k-atoms due to Li-Chung Chen [4] sit inside the coinvariants as subquotients. It is also natural to conjecture that the generalization of catabolism due to Shimozono and Weyman [15] gives a combinatorial description of certain subquotients of the coinvariants which are graded versions of induction products of $\mathcal{S}_n$-irreducibles.

This paper describes an approach to these problems using canonical bases, which has so far been quite successful and will hopefully help solve some of the difficult conjectures in this area. After briefly reviewing Weyl groups, Hecke algebras, and cells (§2), we introduce the central algebraic object of our work, a subalgebra $\widehat{\mathscr{H}}^+$ of the extended affine Hecke algebra which is a u-analogue of the monoid algebra of $\mathcal{S}_n \ltimes \mathbb{Z}^n_{\geq 0}$. In §3, we establish some basic properties of this subalgebra and describe its left cells. It turns out that these cells are naturally labeled by tableaux filled with positive integer entries having distinct residues mod n, which we term *positive affine tableaux* (PAT). Our investigations have convinced us that these are excellent combinatorial objects for describing graded $\mathcal{S}_n$-modules.

After some preparatory combinatorics in §4, we go on to show in §5 that $\widehat{\mathscr{H}}^+$ has a cellular quotient $\mathscr{R}_{1^n}$ that is a u-analogue of R_{1^n}. The module $\mathscr{R}_{1^n}$ has a canonical basis labeled by affine words that are essentially standard words with cocharge labels, with left cells labeled by PAT that are essentially standard tableaux with cocharge labels. Multiplying canonical basis elements by a certain element $\pi \in \widehat{\mathscr{H}}^+$ corresponds to rotations of words, and on left cells corresponds to cocyclage.

In this cellular picture of the coinvariants, G_λ corresponds to a left cell of $\mathscr{R}_{1^n}$ labeled by a PAT of shape λ, termed the *Garnir tableau of shape* λ, again denoted G_λ. In §6, we identify u-analogues $\mathscr{R}_\lambda$ of the R_λ and show that $\mathscr{R}_\lambda$ is cellular and its left cells are labeled by (a PAT version of) the λ-catabolizable tableaux.

Detailed proofs as well as conjectures relating cellular subquotients of $\widehat{\mathscr{H}}^+$ to k-atoms and Chen's atoms and conjectures describing cellular subquotients of $\widehat{\mathscr{H}}^+$ outside of $\mathscr{R}_{1^n}$ are presented in a full version of this extended abstract [1].

2 Hecke algebras and cells

We begin by briefly reviewing Weyl groups and Hecke algebras, referring the reader to [7] for a thorough treatment.

Let W_f, W_a, W_e, Y, Y_+ be the finite Weyl group, affine Weyl group, extended affine Weyl group, weight lattice, and dominant weights associated to the root system specifying the algebraic group $GL_n(\mathbb{C})$ (see [7]). The finite Weyl group W_f is the symmetric group $\mathcal{S}_n$. It acts on the weight lattice $Y = \mathbb{Z}^n = \langle \epsilon_1, \ldots, \epsilon_n \rangle$ by permuting coordinates.

Let $S = \{s_1, \ldots, s_{n-1}\}$ be the simple reflections of W_f and $K = \{s_0, \ldots, s_{n-1}\}$ be those of W_a and W_e. The pairs (W_f, S) and (W_a, K) are Coxeter groups, and (W_e, K) is an extended Coxeter group. The length function ℓ and partial order $\leq$ on W_a extend to $W_e = \Pi \ltimes W_a$: $\ell(\pi v) = \ell(v)$, and $\pi v \leq \pi' v'$ if and only if $\pi = \pi'$ and $v \leq v'$, where $\pi, \pi' \in \Pi$, $v, v' \in W$. For any $J \subseteq K$, the parabolic subgroup $W_{eJ} = W_{aJ}$ is the subgroup of W_e generated by J. Each left (resp. right) coset wW_{eJ} (resp. $W_{eJ}w$) of W_{eJ} contains a unique element of minimal length called a minimal coset representative. The set of all such elements is denoted ${W_e}^J$ (resp. ${}^J W_e$).

We will make use of three descriptions of W_e. First, $W_e = Y \rtimes W_f$; elements of $Y \subseteq W_e$ will be denoted by the multiplicative notation $y^\lambda, \lambda \in Y$ and $y_i := y^{\epsilon_i}$. Second, $W_e = \Pi \ltimes W_a$, where $\Pi \cong \mathbb{Z}$; the element $\pi = y_1 s_1 s_2 \ldots s_{n-1}$ is a generator of Π. This satisfies the relation $\pi s_i = s_{i+1}\pi$, where, here and from now on, the subscripts of the s_i are taken mod n.

The third description of W_e, due to Lusztig, identifies it with the group of permutations $w : \mathbb{Z} \to \mathbb{Z}$ satisfying $w(i + n) = w(i) + n$ and $\sum_{i=1}^{n}(w(i) - i) \equiv 0 \bmod n$. The identification takes s_i to the permutation transposing $i + kn$ and $i + 1 + kn$ for all $k \in \mathbb{Z}$, and takes π to the permutation $k \mapsto k + 1$ for all $k \in \mathbb{Z}$. We take the convention of specifying the permutation of an element $w \in W_e$ by the word

$$n + 1 - w^{-1}(1) \;\; n + 1 - w^{-1}(2) \;\; \ldots \;\; n + 1 - w^{-1}(n).$$

We refer to this as the *affine word* or *word* of w, and it will be written as $w_1 w_2 \cdots w_n$; this is understood to be part of an infinite word so that $w_i = \hat{i} - i + w_{\hat{i}}$, where $\hat{\ } : \mathbb{Z} \to [n]$ is the map sending an integer i to the integer in $[n]$ it is congruent to mod n. For example, if $n = 4$ and $w = \pi^2 s_2 s_0 s_1$, then the word of w is 8 3 5 2, thought of as part of the infinite word . . . 12 7 9 6 8 3 5 2 4 -1 1 -2 . . . . We adopt the convention of writing ab in place of $na + b$ ($a, b \in \mathbb{Z}$). With this convention, the word of w above is written 14 3 11 2.

Let $A = \mathbb{Z}[u, u^{-1}]$ be the ring of Laurent polynomials in the indeterminate u. Let $\mathscr{H}(W)$ be the Hecke algebra of the (extended) Coxeter group W over the ground ring A with standard basis $\{T_w : w \in W\}$. Set $\mathscr{H} = \mathscr{H}(W_f)$, $\widehat{\mathscr{H}} = \mathscr{H}(W_e)$, which will sometimes be decorated with a subscript n to emphasize that they correspond to type A_{n-1} or $\tilde{A}_{n-1}$. The Hecke algebra of an extended Coxeter group has the same relations as the usual Hecke algebra using the length function defined above.

Corresponding to the description $Y \rtimes W_f$ of W_e, there is a presentation of $\widehat{\mathscr{H}}$ due to Bernstein. For any $\lambda \in Y$ there exist $\mu, \nu \in Y_+$ such that $\lambda = \mu - \nu$. Define $Y^\lambda := T_{y^\mu}(T_{y^\nu})^{-1}$, which is independent of the choice of μ and ν. The algebra $\widehat{\mathscr{H}}$ has A-basis

$$\{Y^\lambda T_w : \; w \in W_f, \lambda \in Y\}$$

and is equal to the A-algebra generated by the Y_i and T_{s_i} with relations that are fairly simple to describe.

The *canonical basis* or Kazhdan-Lusztig basis of $\mathscr{H}(W)$ [8] is an A-basis for $\mathscr{H}(W)$, denoted $\{C'_w : w \in W\}$, having nice properties for the action of the Hecke algebra on itself. This action is nice because certain subsets of the canonical basis called cells give rise to representations that are often irreducible. Cells can be defined for any $\mathscr{H}(W)$-module E with a distinguished basis Γ: first, the preorder $\leq_\Gamma$ (also denoted $\leq_E$) on the set Γ is that generated by the relations/edges

$$\delta \underset{\Gamma}{\leftarrow} \gamma \quad \text{if there is an } h \in \mathscr{H}(W) \text{ such that } \delta \text{ appears with non-zero coefficient in the expansion of } h\gamma \text{ in the basis } \Gamma. \tag{2}$$

The *left cells* of Γ (or of E, if $\Gamma \subseteq E$ is understood) are then the equivalence classes of $\leq_\Gamma$. The preorder $\leq_\Gamma$ gives rise to a partial order on left cells, also denoted $\leq_\Gamma$. A *cellular subquotient* of E is a subset Λ of Γ such that there does not exist $\gamma \in \Gamma \backslash \Lambda$ and $\delta, \delta' \in \Lambda$ satisfying $\delta \leq_\Gamma \gamma \leq_\Gamma \delta'$. A cellular subquotient of E is necessarily a union of left cells and gives rise to a subquotient of E. We are most interested in the case where Γ is a W-graph as defined in [8].

3 The positive part of $\widehat{\mathscr{H}}$

Here we introduce a subalgebra $\widehat{\mathscr{H}}^+$ of $\widehat{\mathscr{H}}$ and positive affine tableaux (PAT), which label left cells of $\widehat{\mathscr{H}}^+$. These play a crucial role in our goal of relating subquotients of R to tableau combinatorics.

The subset $Y^+ := \mathbb{Z}^n_{\geq 0}$ of the weight lattice Y is left stable under the action of the Weyl group W_f. Thus $Y^+ \rtimes W_f$ is a submonoid of W_e. We remark that this only works in type A, and this is the main barrier preventing the results of this paper to be generalized to other types.

Proposition-Definition 3.1 *The* positive part *of W_e, denoted W_e^+, has the following three equivalent descriptions:*

(1) $Y^+ \rtimes W_f$,

(2) *The submonoid of W_e generated by π and W_f,*

(3) $\{w \in W_e : w_i > 0$ *for all* $i \in [n]\}$.

The inclusion of monoids $W_e^+ \subseteq W_e$ gives rise to an inclusion of algebras $\widehat{\mathscr{H}}^+ \subseteq \widehat{\mathscr{H}}$:

Proposition-Definition 3.2 *The subalgebra $\widehat{\mathscr{H}}^+$ of $\widehat{\mathscr{H}}$ has the following four equivalent descriptions:*

(i) $A\{Y^\lambda T_w : \lambda \in Y^+, w \in W_f\}$,

(ii) $A\{T_w : w \in W_e^+\}$,

(iii) $A\{C'_w : w \in W_e^+\}$,

(iv) *the subalgebra of $\widehat{\mathscr{H}}$ generated by π and $\mathscr{H}$.*

Write $\leq_{\widehat{\mathscr{H}}^+}$ for the preorder on the canonical basis of $\widehat{\mathscr{H}}^+$ coming from considering $\widehat{\mathscr{H}}^+$ as a left $\widehat{\mathscr{H}}^+$-module. We say that this canonical basis is the W_e^+-graph $\Gamma_{W_e^+}$. The preorder $\leq_{\widehat{\mathscr{H}}^+}$ is difficult to compute, but there are two kinds of easy edges: the edges $C'_{\pi w} \leq_{\widehat{\mathscr{H}}^+} C'_w$, which we refer to as *corotation-edges*; the corresponding edges between cells are *cocyclage-edges* (we will soon see that cocyclage-edges are a generalization of cocyclage for standard Young tableaux). The edges $C'_{sw} \leq_{\widehat{\mathscr{H}}^+} C'_w$ if $sw > w$ and $s \in S$ are *ascent-edges*.

The work of Kazhdan and Lusztig [8] shows that the left cells of $\mathscr{H}$ are in bijection with the set of SYT and the left cell containing C'_w corresponds to the insertion tableau of w under this bijection (keep in mind our unusual convention from §2 for the word of w). The left cell containing those C'_w such that w has insertion tableau P is the left cell *labeled* by P, denoted Γ_P.

Definition 3.3 *A* positive affine tableau *(PAT) of size n is a semistandard Young tableau filled with positive integer entries that have distinct residues mod n.*

For $w \in W_e$, the word $w_1 w_2 \cdots w_n$ may be inserted into a tableau, and the result is a tableau, denoted $P(w)$. It is a positive affine tableau exactly when $w \in W_e^+$. Let Q be a positive affine tableau and let Q_S be the standard tableau obtained from Q by replacing its entries with the numbers $1, \ldots, n$ so that the relative order of the entries in Q and Q_S agree. The set of $w \in W_e$ inserting to Q is $\{vx : v \in W_f \text{ and } P(v) = Q_S\}$, where the word of x is obtained from Q by sorting its entries in decreasing order. For any $x \in {}^S W_e$, define

$$\Gamma_Q := \{C'_{vx} : v \in W_f,\ P(v) = Q_S\} = \{C'_w : w \in W_e,\ P(w) = Q\}. \tag{3}$$

By the following result, Γ_Q is a left cell of $\Gamma_{W_e^+}$, which we refer to as the left cell *labeled by* Q. The following is an easy consequence of results of Roichman on restricting W-graphs that originated in the work of Barbasch and Vogan on primitive ideals (see [14]).

Proposition 3.4 *For any $x \in {}^S W_e^+$, the set $\{C'_{wx} : w \in W_f\}$ is a cellular subquotient of $\widehat{\mathscr{H}}^+$. This set, restricted to be a W_f-graph, is isomorphic to the W_f-graph on $\mathscr{H}$. In particular,*

$$\Gamma_{W_e^+} = \bigsqcup_{Q \in PAT} \Gamma_Q$$

is the decomposition of $\Gamma_{W_e^+}$ into left cells.

4 Cocyclage and catabolism

Before going deeper into the study of the canonical basis of $\widehat{\mathscr{H}}^+$, we introduce combinatorics originating in [10, 11] (see also [15]) that will be used to describe cellular subquotients of $\widehat{\mathscr{H}}^+$.

The *cocharge labeling* of a word v, denoted v^{cc}, is a (non-standard) word of the same length as v, and its numbers are thought of as labels of the numbers of v. It is obtained from v by reading the numbers of v in increasing order, labeling the 1 of v with a 0, and if the i of v is labeled by k, then labeling the $i+1$ of v with a k (resp. $k+1$) if the $i+1$ in v appears to the right (resp. left) of the i in v. For example, the cocharge labeling of 614352 is 302120; also see Example 5.3. Define the cocharge labeling T^{cc} of a tableau T to be the insertion tableau of v^{cc} for any (every) v inserting to inserting to T.

The sum of the numbers in the cocharge labeling of a standard word v (resp. standard tableau T) is the *cocharge* of v (resp. T) or cocharge(v) (resp. cocharge(T)).

For a word w and number $a \neq 1$, aw is a *corotation* of wa. There is a *cocyclage* from the tableau T to the tableau T', written $T \xrightarrow{\mathrm{cc}} T'$, if there exist words u, v such that v is the corotation of u and $P(u) = T$ and $P(v) = T'$. Rephrasing this condition solely in terms of tableaux, $T \xrightarrow{\mathrm{cc}} T'$ if there exists a corner square (r, c) of T and uninserting the square (r, c) from T yields a tableau Q and number a such that T' is the result of column-inserting a into Q.

The *cocyclage poset* $\mathrm{CCP(SYT)}$ is the poset on the set of SYT generated by the relation $\xrightarrow{\mathrm{cc}}$. Similarly, define $\mathrm{CCP}(PAT)$ to be the poset on the set of PAT generated by cocyclage-edges. The covering relations of $\mathrm{CCP(SYT)}$ (resp. $\mathrm{CCP}(PAT)$) are exactly cocyclages (resp. cocyclage-edges). We consider the covering relation $T \xrightarrow{\mathrm{cc}} T'$ to be colored by the following additional datum: the set of outer corners of T that result in a cocyclage to T'. Note that this set can only have more than one element if $\mathrm{sh}(T) = \mathrm{sh}(T')$.

Catabolizability of standard tableaux is a subtle combinatorial statistic, which we will not define in the usual way here. In [3], we show that the *catabolizability* of a standard tableau T, denoted $\mathrm{ctype}(T)$, can be computed from any word v inserting to T using the following *catabolism insertion algorithm*.

Algorithm 4.1 (Catabolism insertion) *Let f be the function below, which takes a pair consisting of a (non-standard) word and a partition to another such pair. Let $x = ya$, y a word and a a number.*

$$f(x,\nu) = \begin{cases} (y, \nu + \epsilon_{a+1}) & \text{if } \nu + \epsilon_{a+1} \text{ is a partition,} \\ (a+1\, y, \nu) & \text{otherwise.} \end{cases} \tag{4}$$

Given the input standard word v, first determine the cocharge labeling z of v. Next, apply f to $(z, \emptyset)$ repeatedly until the word of the pair is empty. Output the partition of this final pair.

Example 4.2 *The sequence of word-partition pairs produced by the algorithm run on $v = 1\,6\,8\,4\,2\,9\,5\,7\,3$ is (reading from left to right and then top to bottom)*

$$\begin{array}{lllll} (023103120,\ \emptyset) & (02310312,\ (1)) & (30231031,\ (1)) & (3023103,\ (1,1)) & (4302310,\ (1,1)) \\ (430231,\ (2,1)) & (43023,\ (2,2)) & (44302,\ (2,2)) & (4430,\ (2,2,1)) & (443,\ (3,2,1)) \\ (44,\ (3,2,1,1)) & (4,\ (3,2,1,1,1)) & (5,\ (3,2,1,1,1)) & (\emptyset,\ (3,2,1,1,1,1)) & \end{array}$$

5 A W_e^+-graph version of the coinvariants

We exhibit a cellular subquotient $\mathscr{R}_{1^n}$ of $\widehat{\mathscr{H}}^+$ which is a W_e^+-graph version of the ring of coinvariants R_{1^n}. Under a natural identification of the left cells of $\mathscr{R}_{1^n}$ with SYT, the subposet of $\leq_{\mathscr{R}_{1^n}}$ consisting of the cocyclage-edges is exactly the cocyclage poset on SYT.

There are two important theorems that give the canonical basis of $\widehat{\mathscr{H}}$ a more explicit description.

The dominant weights Y_+ are weakly decreasing n-tuples of integers; put $Y_+^+ = Y^+ \cap Y_+$. As is customary, let w_0 denote the longest element of W_f. If $\lambda \in Y_+$, then $w_0 y^\lambda$ is maximal in its double coset $W_f y^\lambda W_f$. For $\lambda \in Y_+^+$, let $s_\lambda(Y) \in \widehat{\mathscr{H}}$ denote the Schur function of shape λ in the Bernstein generators Y_i.

Theorem 5.1 (Lusztig [12, Proposition 8.6]) *For any $\lambda \in Y_+^+$, the canonical basis element $C'_{w_0 y^\lambda}$ can be expressed in terms of the Bernstein generators as*

$$C'_{w_0 y^\lambda} = s_\lambda(Y) C'_{w_0} = C'_{w_0} s_\lambda(Y).$$

Let v be an element of W_f, thought of as a standard word. Thinking of the cocharge labeling v^{cc} as an element of Y^+, let $D \subset Y^+$ denote the set of cocharge labelings, which is in bijection with W_f. The set $\{y^\beta : \beta \in D\}$ are the *descent monomials*. Next, put

$$\begin{aligned} D^S &:= \{y^\beta v : \beta \in D \text{ and } v \in W_f \text{ such that } y^\beta v \text{ is minimal in } y^\beta W_f\}, \\ D^S w_0 &:= \{y^\beta v : \beta \in D \text{ and } v \in W_f \text{ such that } y^\beta v \text{ is maximal in } y^\beta W_f\}. \end{aligned} \tag{5}$$

The set $D^S w_0$ will index a canonical basis of the coinvariants.

Proposition 5.2 *There is a bijection $W_f \to D^S w_0$, $v \mapsto w$, defined by setting the word of w to be $w_i = nv_i^{cc} + v_i$. Its inverse has the two descriptions*

$$w_S \leftarrow\!\dashv\, w \tag{6}$$

$$\hat{w}_1\hat{w}_2\ldots\hat{w}_n \leftarrow\!\dashv\, w \tag{7}$$

where $\hat{w}_i$ is the residue of w_i as defined in §2 and w_S is the standard word such that the relative order of the entries in w_S and w agree.

Example 5.3 *For the $v \in \mathcal{S}_9$ given by its word below, the corresponding v^{cc} and w follow.*

$$\begin{array}{rcccccccccc} v & = & 1 & 6 & 8 & 4 & 2 & 9 & 5 & 7 & 3, \\ v^{cc} & = & 0 & 2 & 3 & 1 & 0 & 3 & 1 & 2 & 0, \\ w = nv^{cc} + v & = & 1 & 26 & 38 & 14 & 2 & 39 & 15 & 27 & 3. \end{array}$$

As preparation for the next theorem, we have a proposition giving the factorization of any $w \in W_e^+$ with w maximal in wW_f in terms of descent monomials.

Proposition 5.4 ([2, Proposition 3.7]) *For any $w \in W_e^+$ such that w is maximal in wW_f, there is a unique expression for w of the form*

$$w = u \cdot w_0 y^\lambda,$$

where $u \in D^S$ and $\lambda \in Y_+^+$.

The next theorem simplifying the canonical basis of $\widehat{\mathscr{H}}^+$ is a special case of a result of Xi ([17, Corollary 2.11]), also found independently by the author. We state here a combination of Lusztig's theorem (Theorem 5.1) and Xi's theorem.

Theorem 5.5 *For $w = u \cdot w_0 y^\lambda$ as in Proposition 5.4, we have the factorization*

$$C'_w = s_\lambda(Y) C'_{uw_0}.$$

Remark 5.6 *The general version of this theorem holds for the entire lowest two-sided cell of W_e and in arbitrary type. The general version for type A is used in the full version of this paper [1] to prove analogues of the results below for $\widehat{\mathscr{H}}^+$, rather than just $\widehat{\mathscr{H}}^+ e^+$.*

Let $e^+ = C'_{w_0}$. Then Ae^+ is the one-dimensional trivial left-module of $\mathscr{H}$ in which the T_i act by u for $i \in [n-1]$. The $\widehat{\mathscr{H}}^+$-module $\widehat{\mathscr{H}}^+ e^+ = \widehat{\mathscr{H}}^+ \otimes_{\mathscr{H}} e^+$ is a u-analogue of the polynomial ring R. It can be identified with the cellular submodule of $\widehat{\mathscr{H}}^+$ spanned by $\{C'_w : w \text{ maximal in } wW_f\}$.

Let $\mathscr{R}$ denote the subalgebra of $\widehat{\mathscr{H}}^+$ generated by the Bernstein generators Y_i. It is known that $\mathscr{R} \cong R$ as algebras. Write $(Y^+)_{\geq d}^{W_f} \subseteq \mathscr{R}$ for the set of W_f-invariant polynomials of degree at least d. Now Theorem 5.5 applied to the canonical basis of $\widehat{\mathscr{H}}^+ e^+$ yields the following corollary, which gives a u-analogue of the ring of coinvariants.

Corollary 5.7 *The $\widehat{\mathscr{H}}_n^+$-module $\widehat{\mathscr{H}}_n^+ e^+$ has a cellular quotient equal to*

$$\mathscr{R}_{1^n} := \widehat{\mathscr{H}}_n^+ e^+ / \widehat{\mathscr{H}}_n^+ (Y^+)_{\geq 1}^{\mathcal{S}_n} e^+$$

with canonical basis $\{C'_w : w \in D^S w_0\}$.

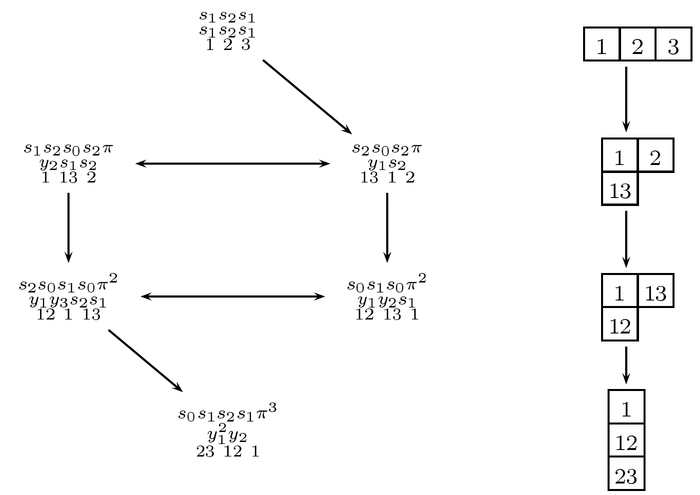


Fig. 1: On the left is the W_e^+-graph of $\mathscr{R}_{1^3}$ with three labels for each canonical basis element. The bottom labels are affine words. On the right are the corresponding left cells and the partial order $\leq_{\mathscr{R}_{1^3}}$ on left cells.

Example 5.8 *The W_e^+-graph $\mathscr{R}_{1^3}$ is drawn in Figure 1. Arrows indicate relations in the preorder $\leq_{\mathscr{R}_{1^n}}$ and those with a downward component are exactly the corotation-edges. Figure 2 depicts the left cells of the W_e^+-graph on $\mathscr{R}_{1^5}$ and the partial order $\leq_{\mathscr{R}_{1^5}}$ on left cells.*

These examples and the next proposition show that the partial order $\leq_{\mathscr{R}_{1^n}}$ contains strictly more information than the cocyclage poset on SYT.

Let $\mathrm{CCP}(\mathscr{R}_{1^n})$ be the subposet of $\mathrm{CCP}(PAT)$ on the set of tableaux corresponding to the cells of $\mathscr{R}_{1^n}$ and let $T + T'$ denote the entry-wise sum of two tableau T, T' of the same shape. Using Proposition 5.2, we deduce the following

Proposition 5.9 *The map $\mathrm{CCP}(SYT) \to \mathrm{CCP}(\mathscr{R}_{1^n}), T \mapsto nT^{cc}+T$ is a color-preserving isomorphism of cocyclage posets.*

For a PAT P labeling a cell of $\mathscr{R}_{1^n}$, let $\mathrm{ctype}(P)$ be $\mathrm{ctype}(T)$, where T is the SYT corresponding to P in the bijection above.

6 A W_e^+-graph version of the Garsia-Procesi modules

The Garsia-Procesi approach to understanding the $R_\lambda = R/I_\lambda$ realizes I_λ as the ideal of leading forms of functions vanishing on an orbit $\mathcal{S}_n\mathbf{a}$, for certain $\mathbf{a} \in \mathbb{C}^n = \operatorname{Spec} R$. We adapt this approach to the Hecke algebra setting using certain representations of $\widehat{\mathscr{H}}$ studied by Bernstein and Zelevinsky in order to prove our main result, Theorem 6.6, which shows that the u-analogues $\mathscr{R}_\lambda$ of the R_λ are actually cellular.

Let $\mathcal{C}_n^{\mathbb{Z}}$ (resp. $\mathcal{C}_n^{+\mathbb{Z}}$) be the category of finite-dimensional $\widehat{\mathscr{H}_n}$-modules (resp. $\widehat{\mathscr{H}_n^+}$-modules) in which the Y_i's have their eigenvalues in $u^{2\mathbb{Z}}$. See [16] for many of the known results about the category $\mathcal{C}_n^{\mathbb{Z}}$.

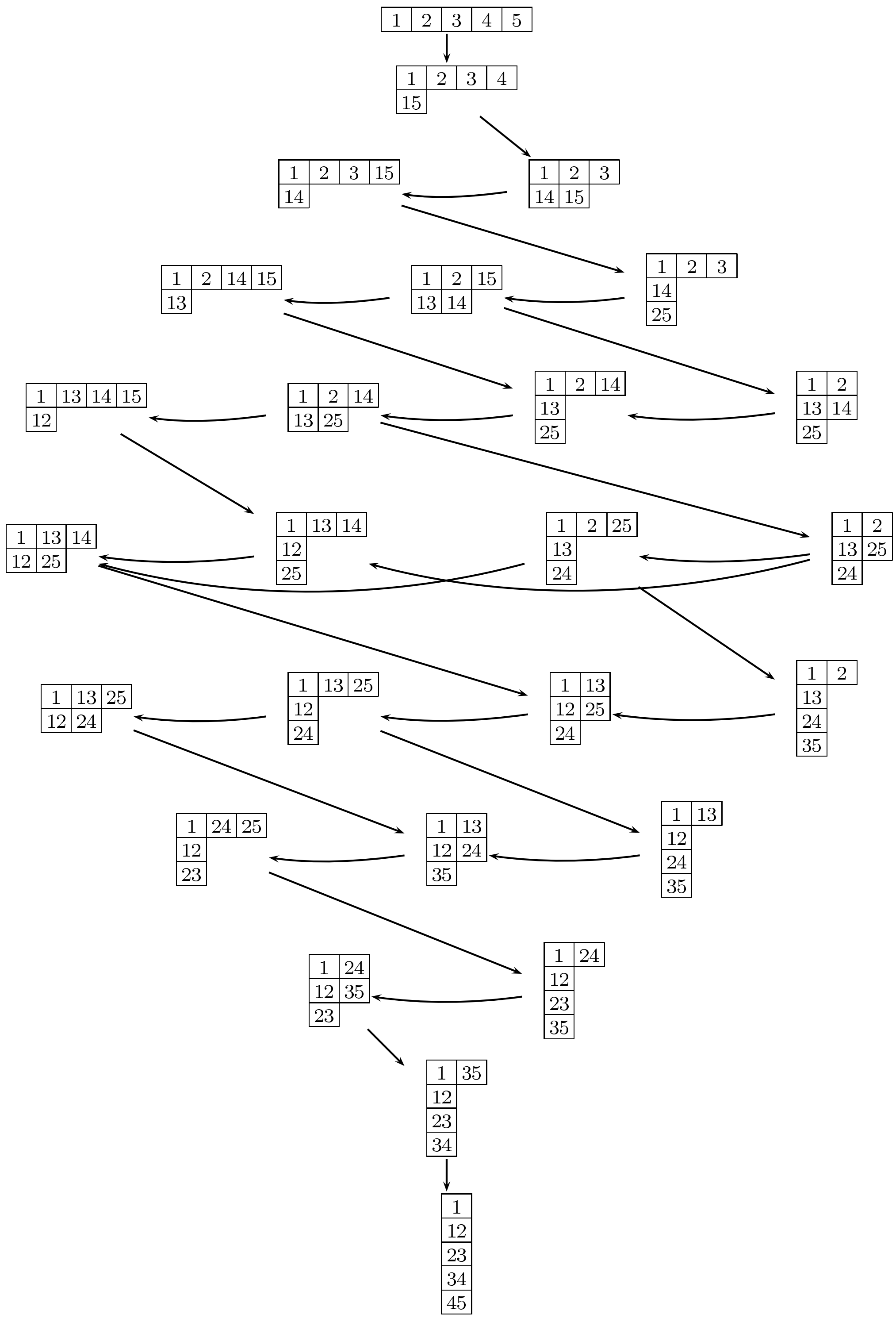


Fig. 2: The cells of the W_e^+-graph on $\mathscr{R}_{1^5}$. Edges are the covering relations of the partial order on cells.

For $\eta = (\eta_1, \eta_2, \ldots, \eta_r)$ an r-composition of n, write $l_j = \sum_{i=1}^{j-1} \eta_i$, $j \in [r+1]$ for the partial sums of η (where the empty sum is defined to be 0). Let B_j be the interval $[l_j + 1, l_{j+1}]$, $j \in [r]$, and define

$$J_\eta = \{s_i : \{i, i+1\} \subseteq B_j \text{ for some } j\} \tag{8}$$

so that $\mathcal{S}_{n\,J_\eta} \cong \mathcal{S}_{\eta_1} \times \cdots \times \mathcal{S}_{\eta_r}$.

Let $\widehat{\mathscr{H}^+_\eta} \cong \widehat{\mathscr{H}^+_{\eta_1}} \times \cdots \times \widehat{\mathscr{H}^+_{\eta_r}}$ be the subalgebra of $\widehat{\mathscr{H}^+}$ generated by $\mathscr{H}_{J_\eta}$ and Y_i, $i \in [n]$. For $\mathbf{a} = (a_1, \ldots, a_r) \in \mathbb{Z}^r$, let $\mathbf{C}_{\eta,\mathbf{a}}$ be the 1-dimensional representation of $\widehat{\mathscr{H}^+_\eta}$ on which $\mathscr{H}_{J_\eta} \subseteq \widehat{\mathscr{H}^+_\eta}$ acts trivially (T_i acts by u for $s_i \in J_\eta$) and Y_{l_i+1} acts by u^{2a_i}, $i \in [r]$. The relations in $\widehat{\mathscr{H}^+_\eta}$ demand that Y_{l_i+k} acts by $u^{2(a_i-k+1)}$ for $l_i + k \in B_i$.

Next define $M_{\eta,\mathbf{a}}$ to be the induced module

$$M_{\eta,\mathbf{a}} = \widehat{\mathscr{H}^+_n} \otimes_{\widehat{\mathscr{H}^+_\eta}} \mathbf{C}_{\eta,\mathbf{a}}. \tag{9}$$

For M in $\mathcal{C}_n^{+\mathbb{Z}}$, the *points* of M are the joint generalized eigenspaces for the action of the Y_i. The *coordinates* of a point v of M is the tuple $(c_1, \ldots, c_n)$ of generalized eigenvalues, also identified with the word $c_1\, c_2 \cdots c_n$. The next proposition follows from a special case of well-known results about $\mathcal{C}_n^{\mathbb{Z}}$.

Proposition 6.1 *If the intervals $[a_i - \eta_i, a_i]$ are disjoint, then the points of $M_{\eta,\mathbf{a}}$ are 1-dimensional and are the shuffles of the words*

$$u^{2a_1}\, u^{2(a_1-1)} \cdots u^{2(a_1-\eta_1)},\ \ u^{2a_2} \cdots u^{2(a_2-\eta_2)},\ \ \ldots, u^{2a_r}\, u^{2(a_r-1)} \cdots u^{2(a_r-\eta_r)}.$$

An essential part of the Garsia-Procesi approach is that the ideal of leading forms of functions vanishing on $\mathcal{S}_n\mathbf{a}$ affords the same $\mathcal{S}_n$-representation as the ideal of functions vanishing on $\mathcal{S}_n\mathbf{a}$. The analogous fact in this setting is

Proposition 6.2 *Let $M_{\eta,\mathbf{a}}$ be as above. If $M_{\eta,\mathbf{a}}$ is irreducible, then it contains an element v^+ such that, setting $N = \operatorname{Ann} v^+$, $\widehat{\mathscr{H}^+}e^+/Ne^+ \cong M_{\eta,\mathbf{a}}$ as $\widehat{\mathscr{H}^+}$-modules. It follows that $\widehat{\mathscr{H}^+}e^+/\mathrm{gr}(N)e^+ \cong \widehat{\mathscr{H}^+}e^+/Ne^+ \cong M_{\eta,\mathbf{a}}$ as $\mathscr{H}$-modules.*

The ideals I_λ are generated by certain elementary symmetric functions in subsets variables, also known as Tanisaki generators (see [5, 6]). By the next theorem, certain $C'_w \in \mathscr{R}_{1^n}$ are essentially these generators. This will relate the ideals $\mathrm{gr}(\operatorname{Ann} M_{\eta,\mathbf{a}})e^+$ to the canonical basis of $\widehat{\mathscr{H}^+}e^+$.

Theorem 6.3 *For $k, d \in [n]$ such that $d \le k$, let w be the maximal element of $y_{k-d+1}y_{k-d+2} \cdots y_k W_f$. Then*

$$C'_w = u^{d(k-n)} s_{1^d}(Y_1, \ldots, Y_k) C'_{w_0}. \tag{10}$$

Suppose $d, k \in [n]$, $d \le k$. Consider the following property of a partition $\lambda \vdash n$:

$$d > k - n + \lambda'_1 + \cdots + \lambda'_{n-k}, \tag{11}$$

where λ' is the partition conjugate to λ.

A result of Garsia-Procesi ([5, Proposition 3.1]) carries over to this setting virtually unchanged. For a composition η, let η_+ denote the partition obtained from η by sorting its parts in decreasing order.

Proposition 6.4 *Suppose η is an r-composition of n with $\lambda := \eta_+$, and $k, d \in [n]$, $d \leq k$, such that (11) holds. If $M_{\eta,\boldsymbol{a}}$ satisfies the hypotheses of Proposition 6.1, then*

$$s_{1^d}(Y_1, \ldots, Y_k) \in \operatorname{gr}(\operatorname{Ann} M_{\eta,\boldsymbol{a}}).$$

For $h \in \widehat{\mathscr{H}}^+$, write $[C'_w]h$ for the coefficient of C'_w of h written as an A-linear combination of $\{C'_w : w \in W_e^+\}$. Define $\langle , \rangle_\lambda : \widehat{\mathscr{H}}^+ \times \widehat{\mathscr{H}}^+ e^+ \to A$ by

$$\langle h_1, h_2 \rangle_\lambda = [C'_{g_\lambda}]h_1 h_2, \tag{12}$$

where g_λ is the row reading word of G_λ.

Through the work of Kazhdan-Lusztig and Beilinson-Bernstein-Deligne-Gabber we know (see, for instance, [13]) that the structure coefficients of the C''s are nonnegative. Using this, we prove

Corollary 6.5 *If $\gamma \in \mathcal{I}_\lambda^{pair}$, $\gamma \in \Gamma_{W_e^+}$, then $\delta \leq_{\widehat{\mathscr{H}}^+} \gamma$ ($\delta \in \Gamma_{W_e^+}$) implies $\delta \in \mathcal{I}_\lambda^{pair}$, i.e., the cellular submodule generated by γ is contained in $\mathcal{I}_\lambda^{pair}$.*

We now come to our main result.

Theorem 6.6 *Suppose $M_{\eta,\boldsymbol{a}}$ satisfies the hypotheses of Propositions 6.1 and 6.2 and maintain the notation of Proposition 6.2. Then the following submodules of $\widehat{\mathscr{H}}^+ e^+$ are equal.*

(i) $\mathcal{I}_\lambda^o := \operatorname{gr}(\operatorname{Ann} v^+)e^+$,

(ii) $\mathcal{I}_\lambda^T := \widehat{\mathscr{H}}^+\{s_{1^d}(Y_1, \ldots, Y_k) : d, k, \lambda$ *satisfy (11)*$\}e^+$,

(iii) $\mathcal{I}_\lambda^{pair} := \{v \in \widehat{\mathscr{H}}^+ e^+ : \langle \widehat{\mathscr{H}}^+, v\rangle_\lambda = 0\}$,

(iv) $\mathcal{I}_\lambda^{cell}$ *:= The maximal cellular submodule of* $\widehat{\mathscr{H}}^+ e^+$ *not containing* Γ_{G_λ},

(v) $\mathcal{I}_\lambda^{cat} := A\{C'_w : \operatorname{ctype}(P(w)) \trianglerighteq \lambda\}$.

The abbreviations o, T, pair, are shorthand for orbit, Tanisaki, and pairing. Also note that modules $M_{\eta,\mathbf{a}}$ satisfying the hypotheses of Propositions 6.1 and 6.2 exist by the general theory. For instance, if $|a_i - a_j| >> 0$ for all $i \neq j$, then these hypotheses are satisfied.

Proof sketch: The inclusion $\mathcal{I}_\lambda^{\mathrm{T}} \subseteq \mathcal{I}_\lambda^{\mathrm{o}}$ follows from Proposition 6.4. An argument of a similar flavor to Proposition 6.4 together with Proposition 6.2 yields $\mathcal{I}_\lambda^{\mathrm{o}} \subseteq \mathcal{I}_\lambda^{\mathrm{pair}}$. Next, the inclusion $\mathcal{I}_\lambda^{\mathrm{T}} \subseteq \mathcal{I}_\lambda^{\mathrm{pair}}$ together with Theorem 6.3 and Corollary 6.5 show that $\mathcal{I}_\lambda^{\mathrm{T}}$ is cellular, implying $\mathcal{I}_\lambda^{\mathrm{T}} \subseteq \mathcal{I}_\lambda^{\mathrm{cell}}$. It follows from the catabolism insertion algorithm (Algorithm 4.1) for any w satisfying $\operatorname{ctype}(P(w)) \trianglerighteq \lambda$, there is a sequence of ascent-edges and corotation-edges from w to g_λ . This proves $\mathcal{I}_\lambda^{\mathrm{cell}} \subseteq \mathcal{I}_\lambda^{\mathrm{cat}}$. Finally, a dimension counting argument using the $u = 1$ results of Garsia-Procesi and Bergeron-Garsia (see [6]) and the standardization map of Lascoux (see [15]) completes the proof. □

Given the theorem, define $\mathscr{R}_\lambda$ to be $\widehat{\mathscr{H}}^+ e^+/\mathcal{I}_\lambda$ for $\mathcal{I}_\lambda$ equal to any (all) of the submodules above. By description (iv), $\mathscr{R}_\lambda$ is the minimal cellular quotient of $\widehat{\mathscr{H}}^+ e^+$ containing Γ_{G_λ}. By description (iii) and the description of R_λ from the introduction, $\mathscr{R}_\lambda$ is a u-analogue of R_λ.

Acknowledgements

I am deeply grateful to my advisor Mark Haiman for his wonderful ideas, his generous advice, and our many detailed discussions that made this paper possible. I am also grateful to John Wood for help typing and typesetting figures.

References

[1] Jonah Blasiak. Cyclage, catabolism, and the affine Hecke algebra. *In preparation*, 2009.

[2] Jonah Blasiak. A factorization theorem for affine Kazhdan-Lusztig basis elements. *Preprint, arXiv:0908.0340v1*, 2009.

[3] Jonah Blasiak. An insertion algorithm for catabolizability. *Preprint, arXiv:0908.1967v1*, 2009.

[4] Li-Chung Chen. Private communication.

[5] A. M. Garsia and C. Procesi. On certain graded S_n-modules and the q-Kostka polynomials. *Adv. Math.*, 94(1):82–138, 1992.

[6] Mark Haiman. Combinatorics, symmetric functions, and Hilbert schemes. In *Current developments in mathematics, 2002*, pages 39–111. Int. Press, Somerville, MA, 2003.

[7] Mark Haiman. Cherednik algebras, Macdonald polynomials and combinatorics. In *International Congress of Mathematicians. Vol. III*, pages 843–872. Eur. Math. Soc., Zürich, 2006.

[8] David Kazhdan and George Lusztig. Representations of Coxeter groups and Hecke algebras. *Invent. Math.*, 53(2):165–184, 1979.

[9] L. Lapointe, A. Lascoux, and J. Morse. Tableau atoms and a new Macdonald positivity conjecture. *Duke Math. J.*, 116(1):103–146, 2003.

[10] Alain Lascoux. Cyclic permutations on words, tableaux and harmonic polynomials. In *Proceedings of the Hyderabad Conference on Algebraic Groups (Hyderabad, 1989)*, pages 323–347, Madras, 1991. Manoj Prakashan.

[11] Alain Lascoux and Marcel-P. Schützenberger. Le monoïde plaxique. In *Noncommutative structures in algebra and geometric combinatorics (Naples, 1978)*, volume 109 of *Quad. "Ricerca Sci."*, pages 129–156. CNR, Rome, 1981.

[12] George Lusztig. Singularities, character formulas, and a q-analog of weight multiplicities. In *Analysis and topology on singular spaces, II, III (Luminy, 1981)*, volume 101 of *Astérisque*, pages 208–229. Soc. Math. France, Paris, 1983.

[13] George Lusztig. Cells in affine Weyl groups. In *Algebraic groups and related topics (Kyoto/Nagoya, 1983)*, volume 6 of *Adv. Stud. Pure Math.*, pages 255–287. North-Holland, Amsterdam, 1985.

[14] Yuval Roichman. Induction and restriction of Kazhdan-Lusztig cells. *Adv. Math.*, 134(2):384–398, 1998.

[15] Mark Shimozono and Jerzy Weyman. Graded characters of modules supported in the closure of a nilpotent conjugacy class. *European J. Combin.*, 21(2):257–288, 2000.

[16] M. Vazirani. Parameterizing Hecke algebra modules: Bernstein-Zelevinsky multisegments, Kleshchev multipartitions, and crystal graphs. *Transform. Groups*, 7(3):267–303, 2002.

[17] Nan Hua Xi. The based ring of the lowest two-sided cell of an affine Weyl group. *J. Algebra*, 134(2):356–368, 1990.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 59–70

Computing Node Polynomials for Plane Curves

Florian Block †

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA.

Abstract. According to the Göttsche conjecture (now a theorem), the degree $N^{d,\delta}$ of the Severi variety of plane curves of degree d with δ nodes is given by a polynomial in d, provided d is large enough. These "node polynomials" $N_\delta(d)$ were determined by Vainsencher and Kleiman–Piene for $\delta \le 6$ and $\delta \le 8$, respectively. Building on ideas of Fomin and Mikhalkin, we develop an explicit algorithm for computing all node polynomials, and use it to compute $N_\delta(d)$ for $\delta \le 14$. Furthermore, we improve the threshold of polynomiality and verify Göttsche's conjecture on the optimal threshold up to $\delta \le 14$. We also determine the first 9 coefficients of $N_\delta(d)$, for general δ, settling and extending a 1994 conjecture of Di Francesco and Itzykson.

Résumé. Selon la Conjecture de Göttsche (maintenant un Théorème), le degré $N^{d,\delta}$ de la variété de Severi des courbes planes de degré d avec δ noeuds est donné par un polynôme en d, pour d assez grand. Ces *polynômes de noeuds* $N_\delta(d)$ ont été déterminés par Vainsencher et Kleiman–Piene pour $\delta \le 6$ et $\delta \le 8$, respectivement. S'appuyant sur les idées de Fomin et Mikhalkin, nous développons un algorithme explicite permettant de calculer tous les polynômes de noeuds, et l'utilisons pour calculer $N_\delta(d)$, pour $\delta \le 14$. De plus, nous améliorons le seuil de polynomialité et vérifions la Conjecture de Göttsche sur le seuil optimal jusqu'à $\delta \le 14$. Nous déterminons aussi les 9 premiers coéfficients de $N_\delta(d)$, pour un δ quelconque, confirmant et étendant la Conjecture de Di Francesco et Itzykson de 1994.

Keywords: Severi degree, curve enumeration, plane curve, node polynomial, labeled floor diagram.

1 Introduction and Main Results

Node Polynomials

Counting algebraic plane curves is a very old problem. In 1848, J. Steiner determined that the number of curves of degree d with 1 node through $\frac{d(d+3)}{2} - 1$ generic points in the complex projective plane $\mathbb{P}^2$ is $3(d-1)^2$. Much effort has since been put forth towards answering the following question:

How many (possibly reducible) degree d nodal curves with
δ nodes pass through $\frac{d(d+3)}{2} - \delta$ generic points in $\mathbb{P}^2$?

The answer to this question is the *Severi degree* $N^{d,\delta}$, the degree of the corresponding Severi variety. In 1994, P. Di Francesco and C. Itzykson [DFI95] conjectured that $N^{d,\delta}$ is given by a polynomial in d (assuming δ is fixed and d is sufficiently large). It is not hard to see that, if such a polynomial exists, it has to be of degree 2δ.

†The author was partially supported by the NSF grant DMS-055588.

Recently, S. Fomin and G. Mikhalkin [FM, Theorem 5.1] established the polynomiality of $N^{d,\delta}$ using tropical geometry and floor decompositions. More precisely, they showed that there exists, for every $\delta \geq 1$, a *node polynomial* $N_\delta(d)$ which satisfies $N^{d,\delta} = N_\delta(d)$ for all $d \geq 2\delta$. (The $\delta = 0$ case is trivial as $N^{d,0} = 1$ for all $d \geq 1$.)

For $\delta = 1, 2, 3$, the polynomiality of the Severi degrees and the formulas for $N_\delta(d)$ were determined in the 19th century. For $\delta = 4, 5, 6$, this was only achieved by I. Vainsencher [Vai95] in 1995. In 2001, S. Kleiman and R. Piene [KP04] settled the cases $\delta = 7, 8$. Earlier, L. Göttsche [Göt98] conjectured a more detailed (still not entirely explicit) description of these polynomials for counting curves on arbitrary projective algebraic surfaces.

Main Results

In this paper we develop, building on ideas of S. Fomin and G. Mikhalkin [FM], an explicit algorithm for computing the node polynomials $N_\delta(d)$ for an arbitrary δ. This algorithm is then used to calculate the node polynomials for all $\delta \leq 14$.

Theorem 1.1 *The node polynomials $N_\delta(d)$, for $\delta \leq 14$, are as listed in [Blo10, Appendix A].*

A list of all $N_\delta(d)$ for $\delta \leq 14$ is implicitly given in Theorem 3.1 of this paper using generating functions. P. Di Francesco and C. Itzykson [DFI95] conjectured the first seven terms of the node polynomial $N_\delta(d)$, for arbitrary δ. We confirm and extend their assertion. The first two terms already appeared in [KP04].

Theorem 1.2 *The first nine coefficients of $N_\delta(d)$ are given by*

$$\begin{aligned} N_\delta(d) = \frac{3^\delta}{\delta!}\Bigg[& d^{2\delta} - 2\delta d^{2\delta-1} - \frac{\delta(\delta-4)}{3} d^{2\delta-2} + \frac{\delta(\delta-1)(20\delta-13)}{6} d^{2\delta-3} + \\ & - \frac{\delta(\delta-1)(69\delta^2-85\delta+92)}{54} d^{2\delta-4} - \frac{\delta(\delta-1)(\delta-2)(702\delta^2-629\delta-286)}{270} d^{2\delta-5} + \\ & + \frac{\delta(\delta-1)(\delta-2)(6028\delta^3-15476\delta^2+117701\delta+4425)}{3240} d^{2\delta-6} + \\ & + \frac{\delta(\delta-1)(\delta-2)(\delta-3)(13628\delta^3-6089\delta^2-295720\delta-24485)}{11340} d^{2\delta-7} + \\ & - \frac{\delta(\delta-1)(\delta-2)(\delta-3)(282855\delta^4-931146\delta^3+417490\delta^2+425202\delta+1141616)}{204120} d^{2\delta-8} + \cdots \Bigg]. \end{aligned} \tag{1.1}$$

Let $d^*(\delta)$ denote the *polynomiality threshold* for Severi degrees, i.e., the smallest positive integer $d^* = d^*(\delta)$ such that $N_\delta(d) = N^{d,\delta}$ for $d \geq d^*$. As mentioned above S. Fomin and G. Mikhalkin showed that $d^* \leq 2\delta$. We improve this as follows:

Theorem 1.3 *For $\delta \geq 1$, we have $d^*(\delta) \leq \delta$.*

In other words, $N^{d,\delta} = N_\delta(d)$ provided $d \geq \delta \geq 1$. L. Göttsche [Göt98, Conjecture 4.1] conjectured that $d^* \leq \lceil \frac{\delta}{2} \rceil + 1$ for $\delta \geq 1$. This was verified for $\delta \leq 8$ by S. Kleiman and R. Piene [KP04]. By direct computation we can push it further.

Proposition 1.4 *For $3 \leq \delta \leq 14$, we have $d^*(\delta) = \lceil \frac{\delta}{2} \rceil + 1$.*

That is, Göttsche's threshold is correct and sharp for $3 \leq \delta \leq 14$. For $\delta = 1, 2$ it is easy to see that $d^*(1) = 1$ and $d^*(2) = 1$.

P. Di Francesco and C. Itzykson [DFI95] hypothesized that $d^*(\delta) \leq \left\lceil \frac{3}{2} + \sqrt{2\delta + \frac{1}{4}} \right\rceil$ (which is equivalent to $\delta \leq \frac{(d^*-1)(d^*-2)}{2}$). However, our computations show that this fails for $\delta = 13$ as $d^*(13) = 8$.

The main techniques of this paper are combinatorial. By the celebrated Correspondence Theorem of G. Mikhalkin [Mik05, Theorem 1] one can replace the algebraic curve count by an enumeration of certain *tropical curves*. E. Brugallé and G. Mikhalkin [BM07, BM09] introduced some purely combinatorial gadgets, called *(marked) labeled floor diagrams* (see Section 2), which, if counted correctly, are equinumerous to these tropical curves. Recently, S. Fomin and G. Mikhalkin [FM] enhanced Brugallé's and Mikhalkin's definition and introduced a *template decomposition* of labeled floor diagrams which is crucial in the proofs of all results in this paper, as is the reformulation of algebraic plane curve counts in terms of labeled floor diagrams (see Theorem 2.5).

This paper is organized as follows: In Section 2 we review labeled floor diagrams, their markings, and their relationship with the enumeration of plane algebraic curves. The proofs of Theorems 1.1 and 1.2 are algorithmic in nature and involve a computer computation. We describe both algorithms in detail in Sections 3 and 5, respectively. The first algorithm computes the node polynomials $N_\delta(d)$ for arbitrary δ, the second determines a prescribed number of leading terms of $N_\delta(d)$. The latter algorithm relies on the polynomiality of solutions of certain polynomial difference equations: This polynomiality has been verified for pertinent values of δ (see Section 5). Proposition 1.4 is proved by comparison of the numerical values of $N_\delta(d)$ and $N^{d,\delta}$ for various d and δ (see Appendices A and B of [Blo]). Theorem 1.3 is discussed in Section 4. For complete proofs of all statements see [Blo].

Additional Comments

In principle, once polynomiality of the Severi degrees $N^{d,\delta}$ is established with some threshold, one could use the Caporaso-Harris recursion [CH98] to compute the node polynomials using simple interpolation. This method, together with the threshold proved in Section 4 of this paper, can in principle be used to compute $N_\delta(d)$ for larger values of δ, and also to increase the upper bound in Proposition 1.4.

The *Gromov-Witten invariant* $N_{d,g}$ enumerates irreducible plane curves of degree d and genus g through $3d + g - 1$ generic points in $\mathbb{P}^2$. Algorithm 1 (with minor adjustments, cf. Theorem 2.5(2)) can be used to directly compute $N_{d,g}$, without resorting to a recursion involving relative Gromov-Witten invariants à la Caporaso–Harris [CH98].

By extending ideas of S. Fomin and G. Mikhalkin [FM] and of the present paper, we can obtain polynomiality results for *relative Severi degrees*, associated with counting curves satisfying given tangency conditions to a fixed line. This will be discussed in the forthcoming paper [Blo10].

A. Gathmann, H. Markwig and the author [BGM] define Psi-floor diagrams which enumerate plane curves which satisfy point and tangency conditions, and conditions given by *Psi-classes*. We prove a Caporaso-Harris type recursion for Psi-floor diagrams, and show that relative descendant Gromov-Witten invariants equal their tropical counterparts.

Acknowledgements

I am thankful to Sergey Fomin for suggesting this problem and fruitful guidance. I also thank Erwan Brugallé and Grigory Mikhalkin for valuable comments and suggestions, and Gregg Musiker and the two referees for helpful comments on an earlier version of this paper. Part of this work was accomplished at the MSRI (Mathematical Sciences Research Institute) in Berkeley, CA, USA, during the semester program on tropical geometry. I thank MSRI for hospitality.

2 Labeled Floor Diagrams

Labeled floor diagrams are combinatorial gadgets which, if counted correctly, enumerate plane curves with certain prescribed properties. E. Brugallé and G. Mikhalkin introduced them in [BM07] (in slightly different notation) and studied them further in [BM09]. To keep this paper self-contained and to fix notation we review them and their markings following [FM] where the framework that best suits our purposes was introduced.

Definition 2.1 *A* labeled floor diagram $\mathcal{D}$ *on a vertex set* $\{1, \dots, d\}$ *is a directed graph (possibly with multiple edges) with positive integer edge weights* $w(e)$ *satisfying:*

1. *The edge directions respect the order of the vertices, i.e., for each edge* $i \to j$ *of* $\mathcal{D}$ *we have* $i < j$.
2. *(Divergence Condition) For each vertex* j *of* $\mathcal{D}$, *we have*

$$\mathit{div}(j) \stackrel{\mathit{def}}{=} \sum_{\substack{\text{edges } e \\ j \xrightarrow{e} k}} w(e) - \sum_{\substack{\text{edges } e \\ i \xrightarrow{e} j}} w(e) \leq 1. \tag{2.1}$$

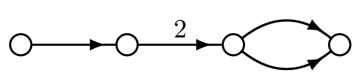
This means that at every vertex of $\mathcal{D}$ *the total weight of the outgoing edges is larger by at most 1 than the total weight of the incoming edges.*

The *degree* of a labeled floor diagram $\mathcal{D}$ is the number of its vertices. It is *connected* if its underlying graph is. Note that in [FM] labeled floor diagrams are required to be connected. If $\mathcal{D}$ is connected its *genus* is the genus of the underlying graph (or the first Betti number of the underlying topological space). The *cogenus* of a connected labeled floor diagram $\mathcal{D}$ of degree d and genus g is given by $\delta(\mathcal{D}) = \frac{(d-1)(d-2)}{2} - g$. If $\mathcal{D}$ is not connected, let $d_1, d_2, \dots$ and $\delta_1, \delta_2, \dots$ be the degrees and cogenera, respectively, of its connected components. Then the *cogenus* of $\mathcal{D}$ is $\sum_j \delta_j + \sum_{j<j'} d_j d_{j'}$. Via the correspondence between algebraic curves and labeled floor diagrams ([FM, Theorem 3.9]) these notions correspond literally to the respective analogues for algebraic curves. Connectedness corresponds to irreducibility. Lastly, a labeled floor diagram $\mathcal{D}$ has *multiplicity*[(i)]

$$\mu(\mathcal{D}) = \prod_{\text{edges } e} w(e)^2. \tag{2.2}$$

We draw labeled floor diagrams using the convention that vertices in increasing order are arranged left to right. Edge weights of 1 are omitted.

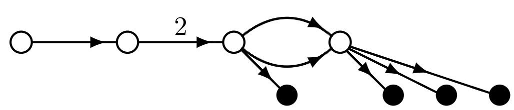
Example 2.2 *An example of a labeled floor diagram of degree* $d = 4$, *genus* $g = 1$, *cogenus* $\delta = 2$, *divergences* $1, 1, 0, -2$, *and multiplicity* $\mu = 4$ *is drawn below.*



To enumerate algebraic curves via labeled floor diagrams we need the notion of markings of such diagrams.

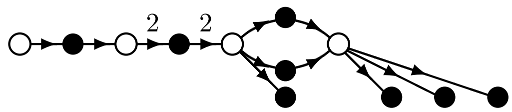
Definition 2.3 *A* marking *of a labeled floor diagram* $\mathcal{D}$ *is defined by the following three step process which we illustrate in the case of Example 2.2.*

Step 1: *For each vertex* j *of* $\mathcal{D}$ *create* $1 - \mathit{div}(j)$ *many new vertices and connect them to* j *with new edges directed away from* j.

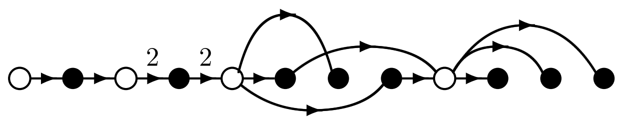


(i) If floor diagrams are viewed as floor contractions of tropical plane curves this corresponds to the notion of multiplicity of tropical plane curves.

Step 2: *Subdivide each edge of the original labeled floor diagram $\mathcal{D}$ into two directed edges by introducing a new vertex for each edge. The new edges inherit their weights and orientations. Call the resulting graph $\tilde{\mathcal{D}}$.*



Step 3: *Linearly order the vertices of $\tilde{\mathcal{D}}$ extending the order of the vertices of the original labeled floor diagram $\mathcal{D}$ such that, as before, each edge is directed from a smaller vertex to a larger vertex.*



The extended graph $\tilde{\mathcal{D}}$ together with the linear order on its vertices is called a marked floor diagram*, or a* marking *of the original labeled floor diagram $\mathcal{D}$.*

We want to count marked floor diagrams up to equivalence. Two markings $\tilde{\mathcal{D}}_1$, $\tilde{\mathcal{D}}_2$ of a labeled floor diagram $\mathcal{D}$ are *equivalent* if there exists an automorphism of weighted graphs which preserves the vertices of $\mathcal{D}$ and maps $\tilde{\mathcal{D}}_1$ to $\tilde{\mathcal{D}}_2$. The *number of markings* $\nu(\mathcal{D})$ is the number of marked floor diagrams $\tilde{\mathcal{D}}$ up to equivalence.

Example 2.4 *The labeled floor diagram $\mathcal{D}$ of Example 2.2 has $\nu(\mathcal{D}) = 7$ markings (up to equivalence): In step 3 the extra 1-valent vertex connected to the third white vertex from the left can be inserted in three ways between the third and fourth white vertex (up to equivalence) and in four ways right of the fourth white vertex (again up to equivalence).*

Now we can make precise how to rephrase the initial question of this paper in terms of combinatorics of labeled floor diagrams.

Theorem 2.5 (Corollary 1.9 of [FM]) *The Severi degree $N^{d,\delta}$, i.e., the number of (possibly reducible) nodal curves in $\mathbb{P}^2$ of degree d with δ nodes through $\frac{d(d+3)}{2} - \delta$ generic points, is equal to*

$$N^{d,\delta} = \sum_{\mathcal{D}} \mu(\mathcal{D})\nu(\mathcal{D}), \tag{2.3}$$

where $\mathcal{D}$ runs over all (possibly disconnected) labeled floor diagrams of degree d and cogenus δ.

3 Computing Node Polynomials

In this section we give an explicit algorithm that symbolically computes the node polynomials $N_\delta(d)$, for given $\delta \geq 1$. (As $N^{d,0} = 1$ for $d \geq 1$, we put $N_0(d) = 1$.) An implementation of this algorithm was used to prove Theorem 1.1 and Proposition 1.4. We mostly follow the notation in [FM, Section 5]. First, we rephrase Theorem 1.1 in more compact notation. For $\delta \leq 8$ one recovers [KP04, Theorem 3.1].

Theorem 3.1 *The node polynomials $N_\delta(d)$, for $\delta \leq 14$, are given by the generating function $\sum_{\delta \geq 0} N_\delta(d)x^\delta$ via the transformation*

$$\sum_{\delta \geq 0} N_\delta(d)x^\delta = \exp\left(\sum_{\delta \geq 0} Q_\delta(d)x^\delta\right), \tag{3.1}$$

where

$$
\begin{aligned}
Q_0(d) &= 1,\\
Q_1(d) &= 3(d-1)^2,\\
Q_2(d) &= \tfrac{-3}{2}(d-1)(14d-25),\\
Q_3(d) &= \tfrac{1}{3}(690d^2-2364d+1899),\\
Q_4(d) &= \tfrac{1}{4}(-12060d^2+47835d-45207),\\
Q_5(d) &= \tfrac{1}{5}(217728d^2-965646d+1031823),\\
Q_6(d) &= \tfrac{1}{6}(-4010328d^2+19451628d-22907925),\\
Q_7(d) &= \tfrac{1}{7}(74884932d^2-391230216d+499072374),\\
Q_8(d) &= \tfrac{1}{8}(-1412380980d^2+7860785643d-10727554959),\\
Q_9(d) &= \tfrac{1}{9}(26842726680d^2-157836614730d+228307435911),\\
Q_{10}(d) &= \tfrac{1}{10}(-513240952752d^2+3167809665372d-4822190211285),\\
Q_{11}(d) &= \tfrac{1}{11}(9861407170992d^2-63560584231524d+101248067530602),\\
Q_{12}(d) &= \tfrac{1}{12}(-190244562607008d^2+1275088266948600d-2115732543025293),\\
Q_{13}(d) &= \tfrac{1}{13}(3682665360521280d^2-25576895657724768d+44039919476860362),\\
Q_{14}(d) &= \tfrac{1}{14}(-71494333556133600d^2+513017995615177680d-913759995239314452).
\end{aligned}
$$

In particular, all $Q_\delta(d)$, for $1 \le \delta \le 14$, are quadratic.

L. Göttsche [Göt98] conjectured that all $Q_\delta(d)$ are quadratic. This theorem proves his conjecture for $\delta \le 14$.

The basic idea of the algorithm (see [FM, Section 5]) is to decompose labeled floor diagrams into smaller building blocks. These gadgets will be crucial in the proofs of all theorems in this paper.

Definition 3.2 *A* template Γ *is a directed graph (with possibly multiple edges) on vertices $\{0,\dots,l\}$, for $l \ge 1$, and edge weights $w(e) \in \mathbb{Z}_{>0}$, satisfying:*

1. *If $i \to j$ is an edge then $i < j$.*
2. *Every edge $i \xrightarrow{e} i+1$ has weight $w(e) \ge 2$. (No "short edges.")*
3. *For each vertex j, $1 \le j \le l-1$, there is an edge "covering" it, i.e., there exists an edge $i \to k$ with $i < j < k$.*

Every template Γ comes with some numerical data associated with it. Its *length* $l(\Gamma)$ is the number of vertices minus 1. The product of squares of the edge weights is its *multiplicity* $\mu(\Gamma)$. Its *cogenus* $\delta(\Gamma)$ is

$$\delta(\Gamma) = \sum_{i \xrightarrow{e} j} \Big[(j-i)w(e) - 1\Big]. \tag{3.2}$$

For $1 \le j \le l(\Gamma)$ let $\varkappa_j = \varkappa_j(\Gamma)$ denote the sum of the weights of edges $i \to k$ with $i < j \le k$ and define

$$k_{\min}(\Gamma) = \max_{1 \le j \le l}(\varkappa_j - j + 1). \tag{3.3}$$

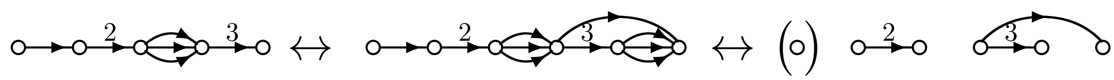
This makes $k_{\min}(\Gamma)$ the smallest positive integer k such that Γ can appear in a floor diagram on $\{1, 2, \dots\}$ with left-most vertex k. Lastly, set

$$\varepsilon(\Gamma) = \begin{cases} 1 & \text{if all edges arriving at } l \text{ have weight } 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.4}$$

For a list of all templates with $\delta \leq 2$ see [FM, Figure 10].

A labeled floor diagram $\mathcal{D}$ with d vertices decomposes into an ordered collection $(\Gamma_1, \dots, \Gamma_m)$ of templates as follows: First, add an additional vertex $d + 1$ ($> d$) to $\mathcal{D}$ along with, for every vertex j of $\mathcal{D}$, $1 - div(j)$ new edges of weight 1 from j to the new vertex $d + 1$. The resulting floor diagram $\mathcal{D}'$ has divergence 1 at every vertex coming from $\mathcal{D}$. Now remove all *short edges* from $\mathcal{D}'$, that is, all edges of weight 1 between consecutive vertices. The result is an ordered collection of templates $(\Gamma_1, \dots, \Gamma_m)$, listed left to right, and it is not hard to see that $\sum \delta(\Gamma_i) = \delta(\mathcal{D})$. This process is reversible once we record the smallest vertex k_i of each template Γ_i (see Example 3.3).

Example 3.3 *An example of the decomposition of a labeled floor diagram into templates is illustrated below. Here, $k_1 = 2$ and $k_2 = 4$.*



To each template Γ we associate a polynomial that records the number of "markings of Γ:" For $k \in \mathbb{Z}_{>0}$ let $\Gamma_{(k)}$ denote the graph obtained from Γ by first adding $k + i - 1 - \varkappa_i$ short edges connecting $i - 1$ to i, for $1 \leq i \leq l(\Gamma)$, and then subdividing each edge of the resulting graph by introducing one new vertex for each edge. By [FM, Lemma 5.6] the number of linear extensions (up to equivalence) of the vertex poset of the graph $\Gamma_{(k)}$ extending the vertex order of Γ is a polynomial in k, if $k \geq k_{\min}(\Gamma)$, which we denote by $P(\Gamma, k)$ (see [FM, Figure 10]). The number of markings of a labeled floor diagram $\mathcal{D}$ decomposing into templates $(\Gamma_1, \dots, \Gamma_m)$ is then

$$\nu(\mathcal{D}) = \prod_{i=1}^{m} P(\Gamma_i, k_i), \tag{3.5}$$

where k_i is the smallest vertex of Γ_i in $\mathcal{D}$. The algorithm is based on

Theorem 3.4 ([FM], (5.13)) *The Severi degree $N^{d,\delta}$, for $d, \delta \geq 1$, is given by the template decomposition formula*

$$\sum_{(\Gamma_1,\dots,\Gamma_m)} \prod_{i=1}^{m} \mu(\Gamma_i) \sum_{k_m = k_{\min}(\Gamma_m)}^{d - l(\Gamma_m) + \varepsilon(\Gamma_m)} P(\Gamma_m, k_m) \cdots \sum_{k_1 = k_{\min}(\Gamma_1)}^{k_2 - l(\Gamma_1)} P(\Gamma_1, k_1), \tag{3.6}$$

where the first sum is over all ordered collections of templates $(\Gamma_1, \dots, \Gamma_m)$, for all $m \geq 1$, with $\sum_{i=1}^{m} \delta(\Gamma_i) = \delta$, and the sums indexed by k_i, for $1 \leq i < m$, are over $k_{\min}(\Gamma_i) \leq k_i \leq k_{i+1} - l(\Gamma_i)$,

Expression (3.6) can be evaluated symbolically, using the following two lemmata. The first is Faulhaber's formula [Knu93] from 1631 for discrete integration of polynomials. The second treats lower limits of iterated discrete integrals and its proof is straightforward. Here B_j denotes the jth Bernoulli number with the convention that $B_1 = +\frac{1}{2}$.

Lemma 3.5 ([Knu93]) *Let $f(k) = \sum_{i=0}^{d} c_i k^i$ be a polynomial in k. Then, for $n \geq 0$,*

$$F(n) \stackrel{def}{=} \sum_{k=0}^{n} f(k) = \sum_{s=0}^{d} \frac{c_s}{s+1} \sum_{j=0}^{s} \binom{s+1}{j} B_j n^{s+1-j}. \tag{3.7}$$

Data: The cogenus δ.
Result: The node polynomial $N_\delta(d)$.
begin
  Generate all templates Γ with $\delta(\Gamma) \leq \delta$;
  $N_\delta(d) \leftarrow 0$;
  forall the *ordered collections of templates* $\tilde{\Gamma} = (\Gamma_1, \ldots, \Gamma_m)$ *with* $\sum_{i=1}^{m} \delta(\Gamma_i) = \delta$ **do**
    $i \leftarrow 1$;
    $Q_1 \leftarrow 1$;
    while $i \leq m$ **do**
      $a_i \leftarrow \max\big(k_{\min}(\Gamma_i), k_{\min}(\Gamma_{i-1}) + l(\Gamma_{i-1}), \ldots, k_{\min}(\Gamma_1) + l(\Gamma_1) + \cdots + l(\Gamma_{i-1})\big)$;
    end
    while $i \leq m-1$ **do**
      $Q_{i+1}(k_{i+1}) \leftarrow \sum_{k_i=a_i}^{k_{i+1}-l(\Gamma_i)} P(\Gamma_i, k_i) Q_i(k_i)$;
      $i \leftarrow i+1$;
    end
    $Q^{\tilde{\Gamma}}(d) \leftarrow \sum_{k_m=a_m}^{d-l(\Gamma_m)+\varepsilon(\Gamma_m)} P(\Gamma_m, k_m) Q_m(k_m)$;
    $Q^{\tilde{\Gamma}}(d) \leftarrow \prod_{i=1}^{m} \mu(\Gamma_i) \cdot Q^{\tilde{\Gamma}}(d)$;
    $N_\delta(d) \leftarrow N_\delta(d) + Q^{\tilde{\Gamma}}(d)$;
  end
end

Algorithm 1: Algorithm to compute node polynomials.

In particular, $deg(F) = deg(f) + 1$.

Lemma 3.6 *Let* $f(k_1)$ *and* $g(k_2)$ *be polynomials in* k_1 *and* k_2*, respectively, and let* $a_1, b_1, a_2, b_2 \in \mathbb{Z}_{\geq 0}$*. Furthermore, let* $F(k_2) = \sum_{k_1=a_1}^{k_2-b_1} f(k_1)$ *be a discrete anti-derivative of* $f(k_1)$*, where* $k_2 \geq a_1 + b_1$*. Then, for* $n \geq \max(a_1 + b_1 + b_2, a_2 + b_2)$*,*

$$\sum_{k_2=a_2}^{n-b_2} g(k_2) \sum_{k_1=a_1}^{k_2-b_1} f(k_1) = \sum_{k_2=\max(a_1+b_1,a_2)}^{n-b_2} g(k_2) F(k_2). \tag{3.8}$$

Using these results Algorithm 1 can be used to compute node polynomials $N_\delta(d)$ for an arbitrary number of nodes δ. The first step, the template enumeration, is explained in [Blo, Section 3].

Proof of Correctness of Algorithm 1.: The algorithm is a direct implementation of Theorem 3.4. The m-fold discrete integral is evaluated symbolically, one sum at a time, using Faulhaber's formula (Lemma 3.5). The lower limit a_i of the ith sum is given by an iterated application of Lemma 3.6. □

As Algorithm 1 is stated its termination in reasonable time is hopeless for $\delta \geq 8$ or 9. The novelty of this section, together with an explicit formulation, is how to implement the algorithm efficiently. This is explained in Remark 3.7.

Remark 3.7 *The running time of the algorithm can be improved vastly as follows: As the limits of summation in (3.6) only depend on* $k_{\min}(\Gamma_i)$*,* $l(\Gamma_i)$ *and* $\varepsilon(\Gamma_m)$*, we can replace the template polynomials* $P(\Gamma_i, k_i)$ *by* $\sum P(\Gamma_i, k_i)$*, where the sum is over all templates* Γ_i *with prescribed* $(k_{\min}, l, \varepsilon)$*. After this transformation the first sum in (3.6) is over all combinations of those tuples. This reduces the computation* *drastically as, for example, the* 167885753 *templates of cogenus* 14 *make up only* 343 *equivalence classes. Also, in (3.6) we can distribute the template multiplicities* $\mu(\Gamma_i)$ *and replace* $P(\Gamma_i, k_i)$ *by* $\mu(\Gamma_i)P(\Gamma_i, k_i)$ *and thereby eliminate* $\prod \mu(\Gamma_i)$*. Another speed-up is to compute all discrete integrals of monomials using Lemma 3.5 in advance.*

The generation of the templates is the bottleneck of the algorithm. Their number grows rapidly with δ as can be seen from Figure 1. However, their generation can be parallelized easily (see [Blo]).

Algorithm 1 has been implemented in Maple. Computing $N_{14}(d)$ on a machine with two quad-core Intel(R) Xeon(R) CPU L5420 @ 2.50GHz, 6144 KB cache, and 24 GB RAM took about 70 days.

Remark 3.8 *We can use Algorithm 1 to compute the values of the Severi degrees* $N^{d,\delta}$ *for prescribed values of* d *and* δ*. After we specify a degree* d *and a number of nodes* δ *all sums in our algorithm become finite and can be evaluated numerically. See [Blo, Appendix B] for all values of* $N^{d,\delta}$ *for* $0 \leq \delta \leq 14$ *and* $1 \leq d \leq 13$*.*

δ	# of templates	δ	# of templates	δ	# of templates
1	2	6	1711	11	2233572
2	7	7	7135	12	9423100
3	26	8	29913	13	39769731
4	102	9	125775	14	167885753
5	414	10	529755		

Fig. 1: The number of templates with cogenera $\delta \leq 14$.

4 Threshold Values

S. Fomin and G. Mikhalkin [FM, Theorem 5.1] proved polynomiality of Severi degrees $N^{d,\delta}$ in d, for fixed δ, if d is sufficiently large. More precisely, they showed that $N_\delta(d) = N^{d,\delta}$ for $d \geq 2\delta$. Here we show that their threshold can be improved to $d \geq \delta$ (Theorem 1.3).

We need the following elementary observation about robustness of discrete anti-derivatives of polynomials whose continuous counterpart is the well known fact that $\int_{a-1}^{a-1} f(x)dx = 0$.

Lemma 4.1 *For a polynomial* $f(k)$ *and* $a \in \mathbb{Z}_{>0}$ *let* $F(n) = \sum_{k=a}^{n} f(k)$ *be the polynomial in* n *uniquely determined by large enough values of* n*. (*$F(n)$ *is a polynomial by Lemma 3.5.) Then* $F(a-1) = 0$*. In particular,* $\sum_{k=a}^{n} f(k)$ *is a polynomial in* n*, for* $n \geq a-1$*.*

The lemma is non-trivial as, in general, $F(a-2) \neq 0$.

Proof of Theorem 1.3 (Sketch): This follows from Equation (3.6) and repeated application of Lemma 3.6 and Lemma 4.1 as $d \geq \delta$ simultaneously implies

$$\begin{aligned} d \geq & l(\Gamma_m) - \varepsilon(\Gamma_m) + k_{\min}(\Gamma_m) - 1, \\ d \geq & l(\Gamma_m) - \varepsilon(\Gamma_m) + l(\Gamma_{m-1}) + k_{\min}(\Gamma_{m-1}) - 2, \\ & \vdots \\ d \geq & l(\Gamma_m) - \varepsilon(\Gamma_m) + l(\Gamma_{m-1}) + \cdots + l(\Gamma_1) + k_{\min}(\Gamma_1) - m, \end{aligned} \tag{4.1}$$

for all collections of templates $(\Gamma_1, \ldots, \Gamma_m)$ with $\sum_{i=1}^{m} \delta(\Gamma_i) = \delta$. For details see [Blo]. $\square$

5 Coefficients of Node Polynomials

The goal of this section is to present an algorithm for the computation of the coefficients of $N_\delta(d)$, for general δ. The algorithm can be used to prove Theorem 1.2 and thereby confirm and extend a conjecture of P. Di Francesco and C. Itzykson in [DFI95] where they conjectured the 7 terms of $N_\delta(d)$ of largest degree.

Our algorithm should be able to find formulas for arbitrarily many coefficients of $N_\delta(d)$. We prove correctness of our algorithm in this section. The algorithm rests on the polynomiality of solutions of certain polynomial difference equations (see [Blo, (5.7)]).

First, we fix some notation building on terminology of Section 3. By Remark 3.7 we can replace the polynomials $P(\Gamma, k)$ in (3.6) by the product $\mu(\Gamma)P(\Gamma, k)$, thereby removing the product $\prod \mu(\Gamma_i)$ of the template multiplicities. In this section we write $P^*(\Gamma, k)$ for $\mu(\Gamma)P(\Gamma, k)$. For integers $i \geq 0$ and $a \geq 0$ let $M_i(a)$ denote the matrix of the linear map

$$f(k) \mapsto \sum_{\Gamma:\delta(\Gamma)=i} \sum_{k=k_{\min}(\Gamma)}^{n-l(\Gamma)} P^*(\Gamma, k) \cdot f(k), \tag{5.1}$$

where $f(k) = c_0k^a + c_1k^{a-1} + \cdots$, a polynomial of degree a, is mapped to the polynomial $M_i(a)(f(k)) = d_0n^{a+i+1} + d_1n^{a+i} + \cdots$ in n. (By Lemma 3.5 and the proof of Lemma 5.1 the image has degree $a+i+1$.) Hence $M_i(a)\mathbf{c} = \mathbf{d}$. Similarly, define $M_i^{\text{end}}(a)$ to be the matrix of the linear map

$$f(k) \mapsto \sum_{\Gamma:\delta(\Gamma)=i} \sum_{k=k_{\min}(\Gamma)}^{n-l(\Gamma)+\varepsilon(\Gamma)} P^*(\Gamma, k) \cdot f(k). \tag{5.2}$$

Later we will consider square sub-matrices of $M_i(a)$ and $M_i^{\text{end}}(a)$ by restriction to the first few rows and columns which will be denoted $M_i(a)$ and $M_i^{\text{end}}(a)$ as well. Note that $M_i(a)$ and $M_i^{\text{end}}(a)$ are lower triangular. The following observation is key to our algorithm.

Lemma 5.1 *The first $a + i$ rows of $M_i(a)$ and $M_i^{end}(a)$ are independent of the lower limits of summation in (5.1) and (5.2), respectively.*

The basic idea of the algorithm is that templates with higher cogenera do not contribute to higher degree terms of the node polynomial. With this in mind we define, for each finite collection $(\Gamma_1, \dots, \Gamma_m)$ of templates, its *type* $\tau = (\tau_2, \tau_3, \dots)$, where τ_i is the number of templates in $(\Gamma_1, \dots, \Gamma_m)$ with cogenus equal to i, for $i \geq 2$. Note that we do not record the number of templates with cogenus equal to 1.

To collect the contributions of all collections of templates with a given type τ, let $\tau = (\tau_2, \tau_3, \dots)$ and fix $\delta \geq \sum_{j\geq 2} \tau_j$ (so that there exist template collections $(\Gamma_1, \dots, \Gamma_m)$ of type τ with $\sum \delta(\Gamma_j) = \delta$). We define two (column) vectors $C_\tau(\delta)$ and $C_\tau^{\text{end}}(\delta)$ as the coefficient vectors, listed in decreasing order, of the polynomials

$$\sum_{(\Gamma_1,\dots,\Gamma_m)} \sum_{k_m=k_{\min}(\Gamma_m)}^{n-l(\Gamma_m)} P^*(\Gamma_m, k_m) \cdots \sum_{k_1=k_{\min}(\Gamma_1)}^{k_2-l(\Gamma_1)} P^*(\Gamma_1, k_1) \tag{5.3}$$

and

$$\sum_{(\Gamma_1,\dots,\Gamma_m)} \sum_{k_m=k_{\min}(\Gamma_m)}^{n-l(\Gamma_m)+\varepsilon(\Gamma)} P^*(\Gamma_m, k_m) \sum_{k_{m-1}=k_{\min}(\Gamma_{m-1})}^{k_m-l(\Gamma_{m-1})} \cdots \sum_{k_1=k_{\min}(\Gamma_1)}^{k_2-l(\Gamma_1)} P^*(\Gamma_1, k_1) \tag{5.4}$$

Data: A positive integer N.
Result: The coefficient vector C of the first N coefficients of $N_\delta(d)$.
begin
  Compute all templates Γ with $\delta(\Gamma) \leq N$;
  forall the *types* τ *with* $\operatorname{def}(\tau) < N$ **do**
    Compute initial values $C_\tau(\delta_0(\tau))$ using (5.3), with $\delta_0(\tau)$ as in Proposition 5.3;
    Solve recursion (5.5) for first $N - \operatorname{def}(\tau)$ coordinates of $C_\tau(\delta)$;
    Set

$$C_\tau^{\text{end}}(\delta) \leftarrow \sum_{i:\tau_i \neq 0} M_i^{\text{end}}\big(2\delta - i - 1 - \operatorname{def}(\tau)\big) C_{\tau\downarrow i}(\delta - i) + M_1^{\text{end}}\big(2\delta - 2 - \operatorname{def}(\tau)\big) C_\tau(\delta - 1);$$

  end
  $C \leftarrow 0$;
  forall the *types* τ *with* $\operatorname{def}(\tau) < N$ **do**
    Shift the entries of $C_\tau^{\text{end}}(\delta)$ down by $\operatorname{def}(\tau)$;
    $C \leftarrow C +$ shifted $C_\tau^{\text{end}}(\delta)$;
  end
end

Algorithm 2: Computation of the leading coefficients of the node polynomial.

in the indeterminate n, where the respective first sums are over all ordered collections of templates of type τ.

Before we can state the main recursion we need two more notations. For a type $\tau = (\tau_2, \tau_3, \dots)$ and $i \geq 2$ with $\tau_i > 0$ define a new type $\tau{\downarrow_i}$ via $(\tau{\downarrow_i})_i = \tau_i - 1$ and $(\tau{\downarrow_i})_j = \tau_j$ for $j \neq i$. Furthermore, let $\operatorname{def}(\tau) = \sum_{j \geq 2}(j-1)\tau_j$ be the *defect* of τ. The following lemma justifies this terminology. Its proof is elementary and can be found in [Blo].

Lemma 5.2 *The polynomials (5.3) and (5.4) are of degree* $2\delta - \operatorname{def}(\tau)$.

The last lemma makes precise which collections of templates contribute to which coefficients of $N_\delta(d)$. Namely, the first N coefficients of $N_\delta(d)$ of largest degree depend only on collections of templates with types τ such that $\operatorname{def}(\tau) < N$. The following recursion is the heart of the algorithm.

Proposition 5.3 *For every type* τ *and integer* δ *large enough, it holds that*

$$C_\tau(\delta) = \sum_{i:\tau_i \neq 0} M_i\big(2\delta - i - 1 - \operatorname{def}(\tau)\big) C_{\tau\downarrow i}(\delta - i) + M_1\big(2\delta - 2 - \operatorname{def}(\tau)\big) C_\tau(\delta - 1). \tag{5.5}$$

More precisely, if we restrict all matrices M_i *to be square of size* $N - \operatorname{def}(\tau)$ *and all* C_τ *to be vectors of length* $N - \operatorname{def}(\tau)$*, then recursion (5.5) holds for*

$$\delta \geq \max\left(\left\lceil \frac{N+1}{2} \right\rceil, \sum_{j \geq 2} j\tau_j \right). \tag{5.6}$$

We propose Algorithm 2 for the computation of the coefficients of the node polynomial $N_\delta(d)$. Due to spacial constrains we explain the step which requires a solution of recursion (5.5) in [Blo].

As in Section 3 (Remark 3.7), Algorithm 2 can be improved significantly by summing the template polynomials $P(\Gamma,k)$ for templates Γ with fixed $\big(k_{\min}(\Gamma), l(\Gamma), \varepsilon(\Gamma)\big)$ in advance. Algorithm 2 has been implemented in Maple. Once the templates are known the bottleneck of the algorithm is the initial value computation. With an improved implementation this should become faster than the template enumeration. Hence we expect Algorithm 2 to be able to compute the first 14 terms of $N_\delta(d)$ in reasonable time.

References

[BGM] F. Block, A. Gathmann, and H. Markwig. Psi-floor diagrams and a Caporaso-Harris type recursion. Preprint, arXiv:1003.2067, submitted.

[Blo] F. Block. Computing node polynomials for plane curves. Preprint, arXiv:1006.0218.

[Blo10] F. Block. Relative node polynomials. in preparation, May 2010.

[BM07] E. Brugallé and G. Mikhalkin. Enumeration of curves via floor diagrams. *C. R. Math. Acad. Sci. Paris*, 345(6):329–334, 2007.

[BM09] E. Brugallé and G. Mikhalkin. Floor decompositions of tropical curves: the planar case. In *Proceedings of Gökova Geometry-Topology Conference 2008*, pages 64–90. Gökova Geometry/Topology Conference (GGT), Gökova, 2009.

[CH98] L. Caporaso and J. Harris. Counting plane curves of any genus. *Invent. Math.*, 131(2):345–392, 1998.

[DFI95] P. Di Francesco and C. Itzykson. Quantum intersection rings. In *The moduli space of curves (Texel Island, 1994)*, volume 129 of *Progr. Math.*, pages 81–148. Birkhäuser Boston, Boston, MA, 1995.

[FM] S. Fomin and G. Mikhalkin. Labeled floor diagrams for plane curves. J. Eur. Math. Soc. (JEMS) (to appear), arXiv: math.AG/0906.3828.

[Göt98] L. Göttsche. A conjectural generating function for numbers of curves on surfaces. *Comm. Math. Phys.*, 196(3):523–533, 1998.

[Knu93] D. Knuth. Johann Faulhaber and sums of powers. *Math. Comp.*, 61(203):277–294, 1993.

[KP04] S. Kleiman and R. Piene. Node polynomials for families: methods and applications. *Math. Nachr.*, 271:69–90, 2004.

[Mik05] G. Mikhalkin. Enumerative tropical geometry in $\mathbb{R}^2$. *J. Amer. Math. Soc.*, 18:313–377, 2005.

[Vai95] I. Vainsencher. Enumeration of n-fold tangent hyperplanes to a surface. *J. Algebraic Geom.*, 4(3):503–526, 1995.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 71–82

Random Walks in the Plane

Jonathan M. Borwein[1] and Dirk Nuyens[2†]and Armin Straub[3] and James Wan[4]

[1]*University of Newcastle, Australia. Email:* `jonathan.borwein@newcastle.edu.au`
[2]*K.U.Leuven, Belgium. Email:* `dirk.nuyens@cs.kuleuven.be`
[3]*Tulane University, New Orleans, USA. Email:* `astraub@tulane.edu`
[4]*University of Newcastle, Australia. Email:* `james.wan@newcastle.edu.au`

Abstract. We study the expected distance of a two-dimensional walk in the plane with unit steps in random directions. A series evaluation and recursions are obtained making it possible to explicitly formulate this distance for small number of steps. Formulae for all the moments of a 2-step and a 3-step walk are given, and an expression is conjectured for the 4-step walk. The paper makes use of the combinatorical features exhibited by the even moments which, for instance, lead to analytic continuations of the underlying integral.

Résumé. Nous étudions la distance espérée d'une marche aléatoire à deux dimensions et à pas unité dans des directions aléatoires. Nous obtenons une évaluation des séries et des récurrences qui permettent de formuler explicitement cette distance pour un petit nombre de pas. Nous donnons des formules pour tous les moments d'une marche aléatoire à 2 et à 3 pas et nous formulons une conjecture pour l'expression d'une marche à 4 pas. Pour les moments pairs, nous utilisons des relations combinatoires qui, par example, permettent le prolongement analytique des intégrales.

Resumen. Se estudia la expectación de la distancia recorrida por una marcha aleatoria en dimensión 2 con pasos de longitud 1. Se presenta una expresión en forma de series y recursiones que permiten encontrar formulas explícitas para la distancia mencionada para un número pequeño de pasos. Fórmulas para todos los momentos en dimensiones 2 y 3 son dadas y se conjectura una expresión analítica para el caso de dimensión 4. Este artículo emplea aspectos de la combinatoria que aparecen en los momentos de order par, para producir una continuación analítica de la integral asociada con este proceso.

Keywords: random walks, hypergeometric functions, high-dimensional integration, analytic continuation

1 Introduction and Preliminaries

This is an extended abstract of (BNSW09) which contains the exposition given here complemented with much more detail. In particular, we often refer to (BNSW09) for full proofs of statements that we present.

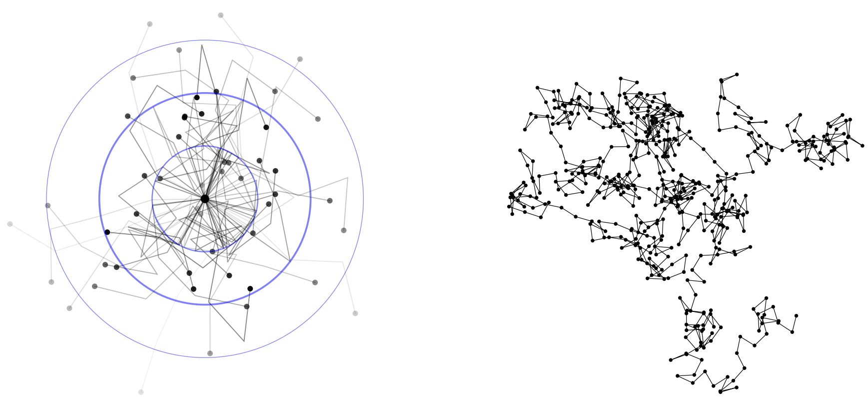
Throughout, we consider the n-dimensional integral

$$W_n(s) := \int_{[0,1]^n} \left| \sum_{k=1}^{n} e^{2\pi x_k i} \right|^s \mathrm{d}\boldsymbol{x} \tag{1}$$

†part of this work done while a research associate at the Department of Mathematics and Statistics, University of New South Wales, Australia

which occurs in the theory of uniform random walk integrals in the plane, where at each step a unit-step is taken in a random direction, see Figure 1. As such, the integral (1) expresses the sth moment of the distance to the origin after n steps. Particularly interesting is the special case of the expected distance $W_n(1)$ after n steps.

A lot is known about the one-dimensional random walk. E.g., its expected distance after n unit-steps is $(n-1)!!/(n-2)!!$ when n is even and $n!!/(n-1)!!$ when n is odd (and asymptotically this distance is $\sqrt{2n/\pi}$). For the two-dimensional walk no such explicit expressions were known, although the term random walk first appears in a (related) question by Karl Pearson in Nature in 1905 (Pea1905) for explicitly this two-dimensional walk under consideration. Pearson triggered answers by Lord Rayleigh (Ray1905) on the asymptotic behaviour of the probability for n very large and by Benett (referred to in (Pea1905b)) for the case $n = 2$, after which he concluded that there still was a large interest for the unresolved case of small n which is dramatically different from the case of large n. Note that the expected value for the root-mean-square distance is well known to be just $\sqrt{n}$ (in that case the implicit square root in (1) disappears which greatly simplifies the problem).



(a) Several 4-step walks (b) A 500-step walk

Fig. 1: Random walks in the plane.

We picked up the special case $s = 1$ of (1) from the whiteboard in the common room at UNSW where it was written as a generalization of a discrete problem in a cryptographic context by Peter Donovan, discussed in (Don09). However, the problem in itself appears in numerous applications, e.g., in problems involving Brownian motion in physics. Numerical values of W_n evaluated at integers can be seen in Tables 1 and 2. One immediately notices the apparent integer sequences for the even moments—which are the moments of the squared expected distance (thus the square root for $s = 2$ gives the root-mean-square distance $\sqrt{n}$). By experimentation and some sketchy arguments we quickly conjectured and believed that, for k a nonnegative integer,

$$W_3(k) = \operatorname{Re}\, {}_3F_2\left(\left.\begin{matrix}\frac{1}{2}, -\frac{k}{2}, -\frac{k}{2} \\ 1, 1\end{matrix}\right| 4\right). \tag{2}$$

(In fact, (2) also holds for negative odd integers.) This was for long a mystery, but it will be proven in the final section of the paper.

n	$s=2$	$s=4$	$s=6$	$s=8$	$s=10$	(Slo09)
2	2	6	20	70	252	`A000984`
3	3	15	93	639	4653	`A002893`
4	4	28	256	2716	31504	`A002895`
5	5	45	545	7885	127905	
6	6	66	996	18306	384156	

Tab. 1: $W_n(s)$ at even integers.

n	$s=1$	$s=3$	$s=5$	$s=7$	$s=9$
2	1.27324	3.39531	10.8650	37.2514	132.449
3	1.57460	6.45168	36.7052	241.544	1714.62
4	1.79909	10.1207	82.6515	822.273	9169.62
5	2.00816	14.2896	152.316	2037.14	31393.1
6	2.19386	18.9133	248.759	4186.19	82718.9

Tab. 2: $W_n(s)$ at odd integers.

In Section 2 we develop an infinite series expression for $W_n(s)$ which holds for all real $s > 0$, see Theorem 2.1. From this it then follows in Corollary 2.2 that the even moments of $W_n(s)$ are given by integer sequences. The combinatorial features of $f_n(k) := W_n(2k)$, k a nonnegative integer, are studied in Section 3. We show that there is a recurrence relation for the numbers $f_n(k)$ and confirm that indeed, an observation from Table 1, the last digit in the column for $s = 10$ is always n modulo 10.

In Section 4 some analytic and numerical results for $n = 1, 2, 3$ are given and we lift the recursion for $f_n(k)$ to $W_n(s)$ by the use of Carlson's theorem. The recursions for $n = 2, 3, 4$ are given explicitly as an example. These recursions then give further information on the poles of the analytic continuations of W_n (graphs of W_n for $n = 3, 4, 5, 6$ and their analytic continuations are shown in Figure 2). From here we conjecture the recursion

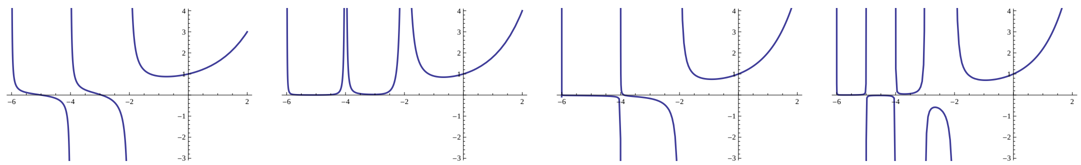
$$W_{2n}(s) \overset{?}{=} \sum_{j\geqslant 0} \binom{s/2}{j}^2 W_{2n-1}(s-2j),$$

based on analytic continuations, and the explicit form, related to (2),

$$W_4(k) \overset{?}{=} \mathrm{Re} \sum_{j\geqslant 0} \binom{s/2}{j}^2 {}_3F_2\left(\left. \begin{matrix} \frac{1}{2}, -\frac{k}{2}+j, -\frac{k}{2}+j \\ 1, 1 \end{matrix} \right| 4 \right).$$

for k a positive integer. High precision numerical evaluations for $n = 3$ and $n = 4$ are given.

In the final section we explore the underlying probability model more closely, starting with another answer to Pearson, this time by Kluyver (Klu1906). Finally, considering conditional densities, we are able to give an alternative form for $W_3(s)$ which eventually leads to a proof of (2).



(a) W_3 (b) W_4 (c) W_5 (d) W_6

Fig. 2: Various W_n and their analytic continuations.

2 A Series Evaluation of $W_n(s)$

Theorem 2.1 *For* $\operatorname{Re} s \geqslant 0$,

$$W_n(s) = n^s \sum_{m\geqslant 0} (-1)^m \binom{s/2}{m} \sum_{k=0}^{m} (-1)^k \binom{m}{k} \left\{ n^{-2k} \sum_{a_1+\cdots+a_n=k} \binom{k}{a_1,\ldots,a_n}^2 \right\}. \tag{3}$$

Proof: We first exploit the binomial theorem to show that

$$W_n(s) = n^s \sum_{m\geqslant 0} (-1)^m \binom{s/2}{m} n^{-2m} \int_{[0,1]^n} \left(4 \sum_{1\leqslant i<j\leqslant n} \sin^2(\pi(x_j - x_i)) \right)^m \mathrm{d}\boldsymbol{x}. \tag{4}$$

Next we evaluate the trigonometric integral in (4). To this end, we show that it is the constant term of

$$(n^2 - (x_1 + \cdots + x_n)(1/x_1 + \cdots + 1/x_n))^m.$$

The details appear in (BNSW09). Alternatively, one may start with the observation that $W_n(s)$ is the constant term of

$$((x_1 + \cdots + x_n)(1/x_1 + \cdots + 1/x_n))^{s/2} \tag{5}$$

which follows directly from the integral definition. □

From Theorem 2.1 and the fact that the binomial transform is an involution we additionally learn that the even moments are integer sequences as detailed by the following corollary.

Corollary 2.2 *For nonnegative integers* k,

$$W_n(2k) = \sum_{a_1+\cdots+a_n=k} \binom{k}{a_1,\ldots,a_n}^2. \tag{6}$$

An outline of the genesis of these evaluations is also given in (BNSW09).

3 Further Combinatorial Features

In light of Corollary 2.2, we consider the combinatorial sums $f_n(k) := W_n(2k)$ of multinomial coefficients squared. These numbers also appear in (RS09) in the following way: $f_n(k)$ counts the number of *abelian squares* of length $2k$ over an alphabet with n letters (that is strings xx' of length $2k$ from an alphabet with n letters such that x' is a permutation of x). It is not hard to see that, (Bar64),

$$f_{n_1+n_2}(k) = \sum_{j=0}^{k} \binom{k}{j}^2 f_{n_1}(j)\, f_{n_2}(k-j), \tag{7}$$

for two non-overlapping alphabets with n_1 and n_2 letters. In particular, we may use (7) to obtain $f_1(k) = 1$, $f_2(k) = \binom{2k}{k}$, as well as

$$f_3(k) = \sum_{j=0}^{k} \binom{k}{j}^2 \binom{2j}{j} = {}_3F_2\left(\left.\begin{matrix}\frac{1}{2}, -k, -k\\ 1, 1\end{matrix}\right| 4\right) = \binom{2k}{k} {}_3F_2\left(\left.\begin{matrix}-k, -k, -k\\ 1, -k+\frac{1}{2}\end{matrix}\right| \frac{1}{4}\right), \tag{8}$$

$$f_4(k) = \sum_{j=0}^{k} \binom{k}{j}^2 \binom{2j}{j}\binom{2(k-j)}{k-j} = \binom{2k}{k} {}_4F_3\left(\left.\begin{matrix}\frac{1}{2}, -k, -k, -k\\ 1, 1, -k+\frac{1}{2}\end{matrix}\right| 1\right). \tag{9}$$

Here and below ${}_pF_q$ denotes the hypergeometric function.

The following result is established in (Bar64) with the recursions for $n \leqslant 6$ given explicitly.

Theorem 3.1 *For fixed $n \geqslant 2$, the sequence $f_n(k)$ satisfies a recurrence of order $\lambda := \lceil n/2 \rceil$ with polynomial coefficients of degree $n-1$:*

$$c_{n,0}(k) f_n(k) + \cdots + c_{n,\lambda}(k) f_n(k+\lambda) = 0.$$

Remark 3.2 For fixed k, the map $n \mapsto f_n(k)$ is a polynomial of degree k. This follows from

$$f_n(k) = \sum_{j=0}^{k} \binom{n}{j} \sum_{\substack{a_1+\cdots+a_j=k\\ a_i>0}} \binom{k}{a_1, \ldots, a_j}^2, \tag{10}$$

because the right-hand side is a linear combination (with positive coefficients only depending on k) of the polynomials $\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{k}$ of respective degrees $0, 1, \ldots, k$. From (10) the coefficient of $\binom{n}{k}$ is seen to be $(k!)^2$. We therefore obtain the first-order approximation $W_n(s) \approx_n n^{s/2}\Gamma(s/2+1)$ for n approaching infinity, see also (Klu1906). In particular, $W_n(1) \approx_n \sqrt{\pi n}/2$. Similarly, the coefficient of $\binom{n}{k-1}$ is $\frac{k-1}{4}(k!)^2$ which gives rise to the second-order approximation

$$(k!)^2 \binom{n}{k} + \frac{k-1}{4}(k!)^2 \binom{n}{k-1} = k! n^k - \frac{k(k-1)}{4}\, k! n^{k-1} + O(n^{k-2}).$$

of $f_n(k)$. We therefore obtain

$$W_n(s) \approx_n n^{s/2-1} \left\{ \left(n - \frac{1}{2}\right) \Gamma\left(\frac{s}{2}+1\right) + \Gamma\left(\frac{s}{2}+2\right) - \frac{1}{4}\Gamma\left(\frac{s}{2}+3\right) \right\},$$

which is exact for $s = 0, 2, 4$. In particular, $W_n(1) \approx_n \sqrt{\pi n}/2 + \sqrt{\pi/n}/32$. More general approximations are given in (Cra09). ◇

Remark 3.3 It follows straight from (6) that, for primes p, $f_n(p) \equiv n$ modulo p. Further, for $k \geqslant 1$, $f_n(k) \equiv n$ modulo 2. This may be derived inductively from the recurrence (7) since, assuming that $f_n(k) \equiv n$ modulo 2 for some n and all $k \geqslant 1$,

$$f_{n+1}(k) = \sum_{j=0}^{k} \binom{k}{j}^2 f_n(j) \equiv 1 + \sum_{j=1}^{k} \binom{k}{j} n = 1 + n(2^k - 1) \equiv n + 1 \pmod 2.$$

Hence for odd primes p,

$$f_n(p) \equiv n \pmod{2p}. \tag{11}$$

The congruence (11) also holds for $p = 2$ since $f_n(2) = (2n-1)n$, compare (10). ◇

Remark 3.4 The integers $f_3(k)$ (respectively $f_4(k)$), the first of which are given in Table 1, also arise in physics, see for instance (BBBG08), and are referred to as *hexagonal* (respectively *diamond*) *lattice integers*. The following formulae (BBBG08, (186)–(188)) relate these sequences in non-obvious ways:

$$\begin{aligned} \left(\sum_{k \geqslant 0} f_3(k)(-x)^k \right)^2 &= \sum_{k \geqslant 0} f_2(k)^3 \frac{x^{3k}}{((1+x)^3(1+9x))^{k+\frac{1}{2}}} \\ &= \sum_{k \geqslant 0} f_2(k) f_3(k) \frac{(-x(1+x)(1+9x))^k}{((1-3x)(1+3x))^{2k+1}} \\ &= \sum_{k \geqslant 0} f_4(k) \frac{x^k}{((1+x)(1+9x))^{k+1}}. \end{aligned}$$

It would be instructive to similarly engage $f_5(k)$. ◇

4 Analytic and Numerical Results

We start with investigating the analyticity of $W_n(s)$ for a given n. In (BNSW09), we show that $W_n(s)$, as defined in (1), is analytic at least for $\operatorname{Re} s > 0$. It is then shown (based on the results of Section 4.2) that (1) is indeed finite and analytic for $\operatorname{Re} s > -2$ for each $n \geqslant 3$ (compare Figure 2).

4.1 Small number of steps

The case $n = 1$ is trivial: it follows straight from the integral definition (1) that $W_1(s) = 1$.

In the case $n = 2$, direct integration of (18) with $n = 2$ yields

$$W_2(s) = 2^{s+1} \int_0^{1/2} \cos(\pi t)^s \, \mathrm{d}t = \binom{s}{s/2}, \tag{12}$$

which may also be obtained using (3).

For $n = 3$, based on (8) we define

$$V_3(s) := {}_3F_2\left(\left.{\frac{1}{2}, -\frac{s}{2}, -\frac{s}{2} \atop 1, 1}\right| 4\right), \tag{13}$$

so that by Corollary 2.2 and (8), $W_3(2k) = V_3(2k)$ for nonnegative integers k. This led us to explore $V_3(s)$ more generally numerically and so to conjecture the following which we prove in the penultimate section:

Theorem 4.1 *For nonnegative even integers and all odd integers k:*

$$W_3(k) = \operatorname{Re} V_3(k). \tag{14}$$

From here, we derive the following equivalent expressions for $W_3(1)$:

$$\begin{aligned} W_3(1) &= \frac{4\sqrt{3}}{3}\left({}_3F_2\left(\left.{-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \atop 1, 1}\right| \frac{1}{4}\right) - \frac{1}{\pi}\right) + \frac{\sqrt{3}}{24}\,{}_3F_2\left(\left.{\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \atop 2, 2}\right| \frac{1}{4}\right) \\ &= 2\sqrt{3}\,\frac{K^2(k_3)}{\pi^2} + \sqrt{3}\,\frac{1}{K^2(k_3)} \\ &= \frac{3}{16}\frac{2^{1/3}}{\pi^4}\Gamma^6\left(\frac{1}{3}\right) + \frac{27}{4}\frac{2^{2/3}}{\pi^4}\Gamma^6\left(\frac{2}{3}\right). \end{aligned}$$

These rely on using Legendre's identity and several Clausen-like product formulae, plus Legendre's evaluation of $K(k_3)$ where $k_3 := \frac{\sqrt{3}-1}{2\sqrt{2}}$ is the *third singular value* as in (BB87). Similar expressions can be given for W_3 evaluated at any odd integer.

4.2 Carlson's Theorem

We may lift the recursive structure of f_n, defined in Section 3, to W_n to a fair degree on appealing to Carlson's theorem (Tit39, 5.81):

Theorem 4.2 (Carlson) *Let f be analytic in the right half-plane $\operatorname{Re} z \geqslant 0$ and of exponential type (meaning that $|f(z)| \leqslant Me^{c|z|}$ for some M and c), with the additional requirement that*

$$|f(z)| \leqslant Me^{d|z|}$$

for some $d < \pi$ on the imaginary axis $\operatorname{Re} z = 0$. If $f(k) = 0$ for $k = 0, 1, 2, \ldots$ then $f(z) = 0$ identically.

By verifying that Carlson's theorem applies, we get:

Theorem 4.3 *Given that $f_n(k)$ satisfies a recurrence*

$$c_{n,0}(k)f_n(k) + \cdots + c_{n,\lambda}(k)f_n(k+\lambda) = 0$$

with polynomial coefficients $c_{n,j}(k)$ (see Theorem 3.1) then $W_n(s)$ satisfies the corresponding functional equation

$$c_{n,0}(s/2)W_n(s) + \cdots + c_{n,\lambda}(s/2)W_n(s+2\lambda) = 0,$$

for $\operatorname{Re} s \geqslant 0$.

Example 4.4 For $n = 2, 3, 4$ we find

$$\begin{aligned}(s+2)W_2(s+2) - 4(s+1)W_2(s) &= 0,\\ (s+4)^2W_3(s+4) - 2(5s^2+30s+46)W_3(s+2) + 9(s+2)^2W_3(s) &= 0,\\ (s+4)^3W_4(s+4) - 4(s+3)(5s^2+30s+48)W_4(s+2) + 64(s+2)^3W_4(s) &= 0.\end{aligned}$$

Note that for all complex s, the function $V_3(s)$ defined in (13) also satisfies the recursion given above for $W_3(s)$—as is routine to prove symbolically. ◇

We note that in each case the recursion lets us determine significant information about the nature and position of any poles of W_n. Details appear in (BNSW09). In particular, for $n \geqslant 3$, the recursion guaranteed by Theorem 4.3 provides an analytic continuation of W_n to all of the complex plane with poles at certain negative integers. Here, we confine ourselves to show the continuations of W_3, W_4, W_5, and W_6 on the negative real axis in Figure 2. These illustrate the fact that, e.g., W_3 and W_5 have simple poles at $-2, -4, -6, \ldots$ whereas W_4 has double poles at these integers. It is further shown in (BNSW09) that, for instance, $\operatorname{Res}_{-2}(W_3) = \frac{2}{\sqrt{3}\pi}$.

Our next somewhat audacious conjecture is (it is now much less audacious as David Broadhurst (Bro09) using a new expression for W_n, namely (20), has been able to verify the statement for $n = 2, 3, 4, 5$ and odd $s < 50$ to a precision of 50 digits):

Conjecture 4.5 *For positive integers n and complex s,*

$$W_{2n}(s) \stackrel{?}{=} \sum_{j \geqslant 0} \binom{s/2}{j}^2 W_{2n-1}(s-2j). \tag{15}$$

It is understood that the right-hand side of (15) refers to the analytic continuation of W_n. By (7) Conjecture 4.5 clearly holds for s an even positive integer. Further, it follows from (12) that the conjecture holds for $n = 1$.

Recall that the real part of $V_3(k)$ as defined in (13) gives $W_3(k)$ for nonnegative integers k. Define

$$V_4(s) := \sum_{j \geqslant 0} \binom{s/2}{j}^2 V_3(s-2j) = \sum_{j \geqslant 0} \binom{s/2}{j}^2 {}_3F_2\left(\left. \begin{matrix} \frac{1}{2}, -\frac{s}{2}+j, -\frac{s}{2}+j \\ 1, 1 \end{matrix} \right| 4 \right) \tag{16}$$

This combines with the much better substantiated special case $n = 2$ of Conjecture 4.5 to provide:

Conjecture 4.6 *For all integers k,*

$$W_4(k) \stackrel{?}{=} \operatorname{Re} V_4(k). \tag{17}$$

4.3 Numerical Evaluations

Note that the following one-dimensional reduction of the integral may be achieved by taking periodicity into account.

$$W_n(s) = \int_{[0,1]^{n-1}} \left| 1 + \sum_{k=1}^{n-1} e^{2\pi x_k i} \right|^s \mathrm{d}(x_1, \ldots, x_{n-1}). \tag{18}$$

$n = 3$ Using this reduction, David Bailey (running tanh-sinh integration on a 256-core LBNL system for roughly 15 minutes) has confirmed that the first 175 digits of $W_3(1)$ are given by

$$\begin{aligned} W_3(1) \quad \approx \quad & 1.57459723755189365749469218307651969022166618075851917019369309830 \\ & 18311805944543821310853133622419530649842236115540882056173012611 \\ & 0810313314994381434429751157865275210084244 58. \end{aligned}$$

This agreed with the evaluation $W_3(1) = \operatorname{Re} V_3(1)$ originally conjectured in (14). He has also confirmed 175 digits for $W_3(s) = \operatorname{Re} V_3(s)$ for $s = 2, \ldots, 7$.

$n = 4$ Using Conjecture 4.6 we provide the approximation

$$\begin{aligned} W_4(1) \quad \approx \quad & 1.79909247984285103353260284584610891006628200329162045662664177735 \\ & 98854266932120575241161930573474828056017014444517983687288 5. \end{aligned}$$

It is worthwhile observing that this level of approximation is made possible by the fact that, roughly, one correct digit is added by each term of the sum.

5 More Probability

As noted such problems have a long lineage. For example, in response to the question posed by Pearson in *Nature*, Kluyver (Klu1906) makes a lovely analysis of the cumulative distribution function of the distance traveled by a "rambler" in the plane for various step lengths. In particular, for our uniform walk Kluyver provides the Bessel representation

$$P_n(t) = t \int_0^\infty J_1(xt)\, J_0^n(x)\,\mathrm{d}x. \tag{19}$$

Thus, $W_n(s) = \int_0^n t^s\, p_n(t)\,\mathrm{d}t$, where $p_n = P_n'$. From here, David Broadhurst (Bro09) obtains

$$W_n(s) = 2^{s+1-k} \frac{\Gamma(1+\frac{s}{2})}{\Gamma(k-\frac{s}{2})} \int_0^\infty x^{2k-s-1} \left(-\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^k J_0^n(x)\,\mathrm{d}x \tag{20}$$

for real s with $2k > s > \max(-2, -\frac{n}{2})$. (20) enables Broadhurst (Bro09) to verify Conjecture 4.5 for $n = 2, 3, 4, 5$ and odd $s < 50$ to a precision of 50 digits.

Remark 5.1 For $n = 3, 4$, symbolic integration in *Mathematica* of (20) leads to interesting analytic continuations (Cra09) such as

$$W_3(s) = \frac{1}{2^{2s+1}} \tan\left(\frac{\pi s}{2}\right) \binom{s}{\frac{s-1}{2}}^2 {}_3F_2\left(\left. {\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \atop \frac{s+3}{2}, \frac{s+3}{2}} \right| \frac{1}{4} \right) + \binom{s}{\frac{s}{2}} {}_3F_2\left(\left. {-\frac{s}{2}, -\frac{s}{2}, -\frac{s}{2} \atop 1, -\frac{s-1}{2}} \right| \frac{1}{4} \right), \tag{21}$$

and

$$W_4(s) = \frac{1}{2^{2s}} \tan\left(\frac{\pi s}{2}\right) \binom{s}{\frac{s-1}{2}}^3 {}_4F_3\left(\left. {\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{s}{2}+1 \atop \frac{s+3}{2}, \frac{s+3}{2}, \frac{s+3}{2}} \right| 1 \right) + \binom{s}{\frac{s}{2}} {}_4F_3\left(\left. {\frac{1}{2}, -\frac{s}{2}, -\frac{s}{2}, -\frac{s}{2} \atop 1, 1, -\frac{s-1}{2}} \right| 1 \right). \tag{22}$$

We note that for s a positive even integer the first term in (21) (resp. (22)) is zero and the second is a term also appearing in (8) (resp. (9)). $\diamond$

Herein, we will take a related probabilistic approach so as to be able to express our quantities of interest in terms of special functions which allows us to explicitly resolve $W_3(2k+1)$ for all odd values.

It is elementary to express the distance y of an $(n+1)$-step walk conditioned on a given distance x of an n-step walk. Since, by a simple application of the cosine rule we find

$$y^2 = x^2 + 1 + 2x\cos(\theta), \tag{23}$$

where θ is the outside angle of the triangle with sides $x, 1, y$. It follows, for details see (BNSW09), that the conditional density for the distance y of an $(n+1)$-step walk as an extension of an n-step walk with distance x is

$$h_x(y) = \frac{2y}{\pi\sqrt{4x^2 - (y^2 - x^2 - 1)^2}} \tag{24}$$

which, of course, is independent of n.

We therefore have the following trivial evaluation

$$W_{n+1}(s) = \mathbb{E}(y^s) = \mathbb{E}\left(\mathbb{E}\left(y^s \mid x\right)\right) = \int_0^n \left(\int_{|x-1|}^{x+1} y^s\, h_x(y)\,\mathrm{d}y\right) p_n(x)\,\mathrm{d}x, \tag{25}$$

under the assumption that the probability density p_n for the n-step walk is known. Clearly, for the 1-step walk we have $p_1(x) = \delta_1(x)$, a Dirac delta at $x = 1$. It then follows immediately that the probability density for a 2-step walk is given by $p_2(x) = \frac{2}{\pi\sqrt{4-x^2}}$ for $0 \leqslant x \leqslant 1$ and 0 otherwise.

5.1 Applications to W_3

The explicit form of $p_2(x)$ leads to some alternative probabilistically inspired expressions for $W_3(s)$. The inner integral in (25) is in fact expressible in terms of the hypergeometric function with details appearing in (BNSW09). For instance, in the case $s = 1$ we find

$$\int_{|x-1|}^{x+1} y\, h_x(y)\,\mathrm{d}y = \frac{2(x+1)}{\pi} E\left(\frac{2\sqrt{x}}{x+1}\right), \tag{26}$$

(for $x > 0$ and $x \neq 1$) where $E(k) = \frac{\pi}{2}\,{}_2F_1(\frac{1}{2}, -\frac{1}{2}; 1; k^2)$ denotes the complete elliptic integral of the second kind with parameter k. This leads to the following expression for the 3-step walk:

$$W_3(1) = \int_0^2 \frac{4(x+1)}{\pi^2\sqrt{4-x^2}}\, E\left(\frac{2\sqrt{x}}{x+1}\right)\,\mathrm{d}x. \tag{27}$$

We are now in a position to prove Theorem 4.1.

Proof of Theorem 4.1: It remains to prove the result for odd integers. Since, as noted, for all complex s, the function $V_3(s)$ defined in (13) also satisfies the recursion given for $W_3(s)$ in Example 4.4, it suffices to show that the values given for $s = 1$ and $s = -1$ are correct. First, (BB87, Exercise 1c), p. 16) allows us to write

$$(x+1)\, E\left(2\,\frac{\sqrt{x}}{x+1}\right) = \operatorname{Re}\left(2\,E(x) - \left(1 - x^2\right) K(x)\right)$$

for $0 < x < \infty$ where we have used Jacobi's imaginary transformations (BB87, Exercises 7a) & 8b), p. 73) to introduce the real part for $x > 1$. Thus, from (27),

$$\begin{aligned} W_3(1) &= \frac{4}{\pi^2}\,\mathrm{Re}\int_0^{\pi/2}\Big(2\,E(2\sin(t)) - (1-4\sin^2(t))K(2\sin(t))\Big)\,\mathrm{d}t \\ &= \frac{4}{\pi^2}\,\mathrm{Re}\int_0^{\pi/2}\int_0^{\pi/2} 2\sqrt{1-4\sin^2(t)\sin^2(r)}\,\mathrm{d}t\,\mathrm{d}r \\ &\quad -\frac{4}{\pi^2}\,\mathrm{Re}\int_0^{\pi/2}\int_0^{\pi/2}\frac{1-4\sin^2(t)}{\sqrt{1-4\sin^2(t)\sin^2(r)}}\,\mathrm{d}t\,\mathrm{d}r. \end{aligned}$$

Joining up the two last integrals and parameterizing, we consider

$$Q(a) := \frac{4}{\pi^2}\int_0^{\pi/2}\int_0^{\pi/2}\frac{1+a^2\sin^2(t)-2\,a^2\sin^2(t)\sin^2(r)}{\sqrt{1-a^2\sin^2(t)\sin^2(r)}}\,\mathrm{d}t\,\mathrm{d}r. \tag{28}$$

We now use the binomial theorem to integrate (28) term-by-term for $|a| < 1$ and substitute

$$\frac{2}{\pi}\int_0^{\pi/2}\sin^{2m}(t)\,\mathrm{d}t = (-1)^m\binom{-1/2}{m}$$

throughout. Moreover, $(-1)^m\binom{-\alpha}{m} = (\alpha)_m/m!$ where the later denoted the *Pochhammer* symbol. Evaluation of the consequent infinite sum produces:

$$\begin{aligned} Q(a) &= \sum_{k\geqslant 0}(-1)^k\binom{-1/2}{k}\left(a^{2k}\binom{-1/2}{k}^2 - a^{2k+2}\binom{-1/2}{k}\binom{-1/2}{k+1} - 2a^{2k+2}\binom{-1/2}{k+1}^2\right) \\ &= \sum_{k\geqslant 0}(-1)^k a^{2k}\binom{-1/2}{k}^3\frac{1}{(1-2k)^2} \\ &= {}_3F_2\left(\left.{-\frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \atop 1, 1}\right| a^2\right). \end{aligned}$$

Analytic continuation to $a = 2$ yields the claimed result as per formula (13) for $s = 1$. The case $s = -1$ is similar, see (BNSW09). $\square$

6 Conclusion

The behaviour of these two-dimensional walks provides a fascinating blend of probabilistic, analytic, algebraic and combinatorial challenges. Work on understanding Conjectures 4.5 and 4.6 is in progress.

Acknowledgements We are grateful to David Bailey for his substantial computational help, to Richard Crandall for helpful suggestions, to Bruno Salvy for reminding us of the existence of (Bar64), Michael Mossinghoff for showing us (Klu1906), and to Peter Donovan for stimulating this research and for much subsequent useful discussion.

References

[BBBG08] D. H. Bailey, J. M. Borwein, D. J. Broadhurst, and M. L. Glasser. "Elliptic integral evaluations of Bessel moments and applications." *J. Phys. A: Math. Theor.*, **41** (2008), 5203–5231.

[Bar64] P. Barrucand. "Sur la somme des puissances des coefficients multinomiaux et les puissances successives d'une fonction de Bessel." *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, **258** (1964), 5318–5320.

[BB87] J. M. Borwein and P. B. Borwein. *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*. Wiley, 1987.

[BNSW09] J. M. Borwein, D. Nuyens, A. Straub, and J. Wan. "Random walk integrals." Preprint, October 2009.

[Bro09] D. J. Broadhurst. "Bessel moments, random walks and Calabi-Yau equations." Preprint, November 2009.

[Cra09] R. E. Crandall, "Analytic representations for random walk moments," Preprint, September 2009.

[Don09] P. Donovan. "The flaw in the JN-25 series of ciphers, II." *Cryptologia*, (2010), in press.

[Klu1906] J. C. Kluyver. "A local probability problem." *Nederl. Acad. Wetensch. Proc.*, **8** (1906), 341–350.

[Pea1905] K. Pearson. "The random walk." *Nature*, **72** (1905), 294.

[Pea1905b] K. Pearson. "The problem of the random walk." *Nature*, **72** (1905), 342.

[Ray1905] Lord Rayleigh. "The problem of the random walk." *Nature*, **72** (1905), 318.

[RS09] L. B. Richmond and J. Shallit. "Counting abelian squares." *The Electronic Journal of Combinatorics*, **16**, 2009.

[Slo09] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences, 2009. Published electronically at `http://www.research.att.com/sequences`.

[Tit39] E. Titchmarsh. *The Theory of Functions*. Oxford University Press, 2nd edition, 1939.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 83–94

Three notions of tropical rank for symmetric matrices

Dustin Cartwright [1] and Melody Chan[1†]

[1]*Dept. of Mathematics, University of California, Berkeley, CA 94720, USA*

Abstract. We introduce and study three different notions of tropical rank for symmetric matrices and dissimilarity matrices in terms of minimal decompositions into rank 1 symmetric matrices, star tree matrices, and tree matrices. Our results provide a close study of the tropical secant sets of certain nice tropical varieties, including the tropical Grassmannian. In particular, we determine the dimension of each secant set, the convex hull of the variety, and in most cases, the smallest secant set which is equal to the convex hull.

Résumé. Nous introduisons et étudions trois notions différentes de rang tropical pour des matrices symétriques et des matrices de dissimilarité, en utilisant des décompositions minimales en matrices symétriques de rang 1, en matrices d'arbres étoiles, et en matrices d'arbres. Nos résultats donnent lieu à une étude détaillée des ensembles des sécantes tropicales de certaines jolies variétés tropicales, y compris la grassmannienne tropicale. En particulier, nous déterminons la dimension de chaque ensemble des sécantes, l'enveloppe convexe de la variété, ainsi que, dans la plupart des cas, le plus petit ensemble des sécantes qui est égal à l'enveloppe convexe.

Resumen. Introducimos y estudiamos tres nociones diferentes de rango tropical para matrices simétricas y matrices de disimilaridad, utilizando las decomposiciones minimales en matrices simétricas de rango 1 en matrices de árboles estrella y en matrices de árboles. Nuestros resultados brindan un estudio detallado de conjuntos de secantes de ciertas variedades tropicales clásicas, incluyendo la grassmanniana tropical. En particular, determinamos la dimensión de cada conjunto de dichas secantes, la cápsula convexa de la variedad, y también, en la mayoridad de los casos, el conjunto más pequeño de secantes que coincide con la cápsula convexa.

Keywords: tropical geometry, tropical convexity, secant varieties, rank, symmetric matrices, hypergraph coloring

---

1 Introduction

In this paper, we study tropical secant sets and rank for symmetric matrices. Our setting is the *tropical semiring* $(\mathbb{R} \cup \{\infty\}, \oplus, \odot)$, where tropical addition is given by $x \oplus y = \min(x, y)$ and tropical multiplication is given by $x \odot y = x + y$. The kth *tropical secant set* of a subset V of $\mathbb{R}^N$ is defined to be the set of points

$$\{x \in \mathbb{R}^N : x = v_1 \oplus \cdots \oplus v_k,\ v_i \in V\},$$

where $\oplus$ denotes coordinate-wise minimum. This set is typically not a tropical variety and thus we prefer the term "secant set" to "secant variety," which has appeared previously in the literature. The *rank* of a

---

[†]Supported by the Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program.

point $x \in \mathbb{R}^N$ with respect to V is the smallest integer k such that x lies in the kth tropical secant set of V, or ∞ if there is no such k.

In [3], Develin, Santos, and Sturmfels define the Barvinok rank of a matrix, not necessarily symmetric, to be the rank with respect to the subset of $n \times n$ rank 1 matrices, and their definition serves as a model for ours. In addition, they define two other notions of rank, Kapranov rank and tropical rank, for which there are no analogues in this paper. Further examination of min-plus ranks of matrices can be found in the review article [1].

We give a careful examination of secant sets and rank with respect to three families of tropical varieties in the space of symmetric matrices and the space of dissimilarity matrices. By a $n \times n$ *dissimilarity matrix* we simply mean a function from $\binom{[n]}{2}$ to $\mathbb{R}$, which we will write as a symmetric matrix with the placeholder symbol $*$ along the diagonal. There is a natural projection from $n \times n$ symmetric matrices to $n \times n$ dissimilarity matrices which we denote by π. For example,

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \pi(M) = \begin{bmatrix} * & 1 & 0 & 0 \\ 1 & * & 0 & 0 \\ 0 & 0 & * & 1 \\ 0 & 0 & 1 & * \end{bmatrix} \tag{1}$$

are a symmetric matrix and dissimilarity matrix respectively.

Our first family is the *tropical Veronese* of degree 2, which is the tropicalization of the classical space of symmetric matrices of rank 1. It is a classical linear subspace of the space of symmetric matrices consisting of those matrices which can be written as $v^T \odot v$ for some row vector v. The rank of a matrix with respect to the tropical Veronese is called *symmetric Barvinok rank*, because it is the symmetric analogue of Barvinok rank.

Second, we consider the *space of star trees*, which is the image of the tropical Veronese under the projection π. Equivalently, it can be obtained by first projecting the classical Veronese onto its off-diagonal entries and then tropicalizing. The classical variety and its secant varieties were studied in a statistical context in [5]. The tropical variety is a classical linear subspace of the space of dissimilarity matrices, and we call the corresponding notion of rank *star tree rank*. The name reflects the fact that the matrices with star tree rank 1 are precisely those points of the tropical Grassmannian which correspond to trees with no internal edges, i.e. star trees, in in the identification below.

Third, we consider the *tropical Grassmannian* $G_{2,n}$, which is the tropicalization of the Grassmannian of 2-dimensional subspaces in an n-dimensional vector space, and was first studied in [9]. It consists of exactly those dissimilarity matrices arising as the distance matrix of a weighted tree with n leaves in which internal edges have negative weights. Therefore, we call the points in the tropical Grassmannian *tree matrices*, and call rank with respect to the tropical Grassmannian the *tree rank*. Note that our definition of tree rank differs from that in [7, Ch. 3], which uses a different notion of mixtures.

We use our examples of M and $\pi(M)$ from (1) to illustrate our three notions of rank. Proposition 4 tells us that the symmetric Barvinok rank of M is 4. Theorem 8 tells us that the star tree rank of $\pi(M)$ is 2. Explicitly, we have

$$\pi(M) = \begin{bmatrix} * & 1 & 0 & 0 \\ 1 & * & 2 & 2 \\ 0 & 2 & * & 1 \\ 0 & 2 & 1 & * \end{bmatrix} \oplus \begin{bmatrix} * & 1 & 2 & 2 \\ 1 & * & 0 & 0 \\ 2 & 0 & * & 1 \\ 2 & 0 & 1 & * \end{bmatrix}.$$

Finally, the tree rank of $\pi(M)$ is 1 by Proposition 13. This example shows that all three of our notions of rank can be different.

However, for any $n \times n$ symmetric matrix M, we have

$$\text{symmetric Barvinok rank}(M) \geq \text{star tree rank}(\pi(M)) \geq \text{tree rank}(\pi(M)). \tag{2}$$

The first inequality follows from the fact that the set of dissimilarity matrices of star tree rank 1 is the projection of the set of matrices of symmetric Barvinok rank 1. The second inequality follows from the fact that the space of star trees is contained in the tropical Grassmannian.

The rest of the paper is organized as follows. In Section 2, we present a technique for proving lower bounds on rank. We introduce a hypergraph whose chromatic number is a lower bound on rank. We examine symmetric Barvinok rank, star tree rank, and tree rank in Sections 3, 4, and 5 respectively. We prove upper bounds on the rank in each case, and with the exception of tree rank, our upper bounds are sharp. We show that the symmetric Barvinok rank of an $n \times n$ symmetric matrix can be infinite, but even when the rank is finite it can exceed n, and in fact can grow quadratically in n (Theorem 5). For each notion of rank, the set of matrices with rank at most k is a union of polyhedral cones, and we compute the dimension of these sets, defined as the dimension of the largest cone. In each case, the dimension of the tropical secant set equals the dimension of the clasical secant variety, confirming Draisma's observation that tropical geometry provides useful lower bounds for the dimensions of classical secant varieties [4]. We also give a combinatorial characterization of each notion of rank for a $0/1$ matrix in terms of graph covers. Finally, in Section 6, we explicitly characterize the stratification of the 5×5 dissimilarity matrices by star tree rank and tree rank respectively, and show that the lower bounds from the chromatic number in Section 2 are exact in these cases.

2 Lower bounds on rank via hypergraph coloring

We begin by giving a general combinatorial construction: a hypergraph whose chromatic number yields a lower bound on rank.

Recall that a *hypergraph* consists of a ground set, called vertices, and a set of subsets of the ground set, called hyperedges. The *chromatic number* of a hypergraph H, denoted $\chi(H)$, is the smallest number r such that the vertices of H can be partitioned into r color classes with no hyperedge of H monochromatic. In particular, if H contains a hyperedge of size 1, then $\chi(H)$ is ∞.

Now, suppose we have a tropical prevariety $V \subseteq \mathbb{R}^N$. Recall that a tropical polynomial

$$p(x_1, \ldots, x_N) = \bigoplus_{i=1}^{t} a_i \odot x_1^{c_{i1}} \odot \cdots \odot x_N^{c_{iN}} \tag{3}$$

defines a tropical hypersurface consisting of those vectors $x \in \mathbb{R}^N$ such that the minimum in evaluating $p(x)$ is achieved at least twice. A *tropical prevariety* is the intersection of finitely many tropical hypersurfaces, and we call the set of tropical polynomials defining the prevariety V a *defining set*.

Now, given a point $w \in \mathbb{R}^N$ and a defining set S for V, we construct a hypergraph on ground set $[N]$ as follows. Let p from (3) be a tropical polynomial in S, with all exponents $c_{ij} \geq 0$. If the minimum is achieved uniquely when p is evaluated at w, then we add a hyperedge E whose elements correspond to the coordinates that appear with non-zero exponent in the unique minimal term. The *deficiency hypergraph* of w with respect to V and S consists of hyperedges coming from all polynomials in S with a unique minimum at w. In particular, the deficiency hypergraph has no hyperedges (and thus has chromatic number 1) if and only if w is in V.

Proposition 1 *If H is the deficiency hypergraph constructed above, then the rank of $w \in \mathbb{R}^N$ with respect to $V \subseteq \mathbb{R}^N$ is at least $\chi(H)$.*

Corollary 2 *If the deficiency hypergraph H has a hyperedge of size 1, then the rank of w with respect to V is infinite.*

The lower bound in Proposition 1 may be strict, such as with $S = \{xz \oplus y^2, xw \oplus yz\}$ and $w = (0,0,0,1)$. The rank of w with respect to the variety defined by S is infinite, but the deficiency hypergraph is 2-colorable.

For the varieties considered in this paper, we will take quadratic tropical bases as our defining tropical polynomials, and thus the deficiency hypergraph will always be a graph (possibly with loops). Accordingly, we will call it the *deficiency graph*.

3 Symmetric Barvinok rank

The symmetric Barvinok rank of a symmetric matrix M is the smallest number r such that M can be written as the sum of r rank 1 symmetric matrices. The 2×2 minors $x_{ij}x_{kl} \oplus x_{il}x_{kj}$ of M for $i \neq k$ and $l \neq j$ form a tropical basis for the variety of rank 1 symmetric matrices. We will always construct our deficiency graph with respect to this tropical basis.

Our first observation is that the symmetric Barvinok rank of a matrix can be infinite. More precisely,

Proposition 3 *If M is a symmetric matrix and $2M_{ij} < M_{ii} + M_{jj}$ for some i and j, then the symmetric Barvinok rank of M is infinite.*

Proof: The tropical polynomial $x_{ij}^2 \oplus x_{ii}x_{jj}$ is in the tropical basis, so if $2M_{ij} < M_{ii} + M_{jj}$ for some i and j, then the deficiency graph for M has a loop at the node ij. Therefore, M has infinite rank by Corollary 2. □

In fact, the converse to Proposition 3 is also true; see Theorem 5.

Next, we give a graph-theoretic characterization of the symmetric Barvinok rank of $0/1$-matrices. We define a *clique cover* of a simple graph G to be a collection of r complete subgraphs such that every edge and every vertex of G is in some element of the collection. Given an $n \times n$ symmetric $0/1$ matrix M with zeroes on the diagonal, define G_M to be the graph whose vertices are the integers $[n]$, and which has an edge between i and j if and only if $M_{ij} = 0$.

Proposition 4 *Suppose M is a symmetric $0/1$ matrix with zeroes on the diagonal. Then the symmetric Barvinok rank of M is the size of a smallest clique cover of G_M.*

On the other hand, suppose that M is a symmetric $0/1$ matrix with at least one entry of 1 on the diagonal. If there exist i and j such that $M_{ii} = 1$ and $M_{ij} = 0$, then the symmetric Barvinok rank of M is infinite. Otherwise, let M' be the maximal principal submatrix with zeroes on the diagonal. The symmetric Barvinok rank of M is one greater than the symmetric Barvionk rank of M'.

This characterization gives us two families of matrices which have rank n and $\lfloor n^2/4 \rfloor$ respectively, namely those corresponding to the trivial graph with n isolated vertices and the complete bipartite graph

$K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. In the latter case, $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is triangle-free, so no clique can consist of more than one edge. On the other hand, there are $\lfloor n^2/4 \rfloor$ edges in the graph, so $\lfloor n^2/4 \rfloor$ cliques are needed in a cover. In fact, these two examples have the maximum possible rank for $n \times n$ matrices.

Theorem 5 *Suppose that M is a symmetric $n \times n$ matrix with $M_{ii} + M_{jj} \leq 2M_{ij}$ for all i and j. Then the symmetric Barvinok rank of M is at most $\max\{n, \lfloor n^2/4 \rfloor\}$, and this bound is tight. Thus, every matrix with finite rank has rank at most $\max\{n, \lfloor n^2/4 \rfloor\}$.*

The next theorem shows that the dimensions of the tropical secant sets and their classical secant varieties agree.

Theorem 6 *The dimension of the space of symmetric $n \times n$ matrices of symmetric Barvinok rank at most r is $\binom{n+1}{2} - \binom{n-r+1}{2}$, which is the dimension of the classical secant variety, i.e. the space of classical symmetric matrices of classical rank at most r.*

Proof: Let $D = \binom{n+1}{2} - \binom{n-r+1}{2}$. The tropical secant set is contained in the tropicalization of the classical secant variety, so the dimension is at most D, by the Bieri-Groves Theorem [2, Thm. A]. Thus, it is sufficient to find an open neighborhood in which the tropical secant set has dimension D. For i from 1 to r, let $v_i = (C, \dots, C, v_{i,i}, \dots, v_{i,n})$ be a vector with C for the first $i-1$ entries. Choose the coordinates $v_{i+1,j}$ to be smaller than all the $v_{i,j}$ and let C be very large. Then,

$$v_1^T \odot v_1 \oplus \cdots \oplus v_r^T \odot v_r = \begin{bmatrix} 2v_{11} & v_{11} + v_{12} & \cdots & v_{11} + v_{1n} \\ v_{11} + v_{12} & 2v_{22} & \cdots & v_{22} + v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{11} + v_{1n} & v_{22} + v_{2n} & \dots & 2v_{rn} \end{bmatrix}$$

This matrix is an injective function of the vector entries v_{ij} for $i \leq r$ and $j \geq i$. Thus, it defines a neighborhood of the rth secant set of the desired dimension

$$n + (n-1) + \cdots + (n-r+1) = \binom{n+1}{2} - \binom{n-r+1}{2} = D.$$

□

In the case $n = 3$, we can explicitly describe the stratification of symmetric matrices by symmetric Barvinok rank. By Theorem 5, the symmetric Barvinok rank of a 3×3 matrix is at most 3, and the locus of rank 1 matrices is the tropical variety defined by the 2×2 minors, so it suffices to characterize the matrices of rank at most 2. Following [3], we call a square matrix *tropically singular* if it lies in the tropical variety of the determinant.

Proposition 7 *Let M be a symmetric 3×3 matrix. Then the following are equivalent:*

1. *M has symmetric Barvinok rank at most 2;*
2. *The deficiency graph of M is 2-colorable;*
3. *M is tropically singular and $M_{ii} + M_{jj} \leq 2M_{ij}$ for all $1 \leq i, j \leq 3$.*

We remark that for larger matrices, the symmetric Barvinok rank does not have as simple a characterization as the third condition in Proposition 7. A necessary condition for a symmetric $n \times n$ matrix to have rank at most r is that $M_{ii} + M_{jj} \leq 2M_{ij}$ and all the $(r+1) \times (r+1)$ submatrices are tropically singular, but one can show that this condition is not sufficient for $n \geq 4$.

4 Star tree rank

A star tree matrix is a dissimilarity matrix which can be written as $\pi(v^T \odot v)$ for $v \in \mathbb{R}^n$ a row vector. The star tree matrices form a classical linear space in the space of $n \times n$ dissimilarity matrices defined by the tropical polynomials

$$x_{ij}x_{kl} \oplus x_{ik}x_{jl} \qquad \text{for } i,\, j,\, k, \text{ and } l \text{ distinct integers.} \tag{4}$$

In this section, the deficiency graph will always be taken with respect to this tropical basis.

Unlike the case of symmetric Barvinok rank, the star tree rank is always finite.

Theorem 8 *For n at least 3, the star tree rank of a $n \times n$ dissimilarity matrix M is at most $n-2$, and this bound is sharp. In particular, the dissimilarity matrix defined by $M_{ij} = \min\{i, j\}$ has star tree rank $n-2$.*

Note that the matrix with maximal star tree rank in Theorem 8 is in fact in the Grassmannian, i.e. it has tree rank 1. Indeed, one may check that the four-point condition (5) holds.

We can also give a graph-theoretic characterization of the star tree rank of $0/1$-matrices. For M a $0/1$ dissimilarity matrix, we define G_M to be the graph whose edges correspond to the zeroes of M. As in the case of symmetric Barvinok rank, we can characterize the star tree rank of M in terms of covers of G_M, this time by both cliques and star trees. We will also say that a cover of G_M by cliques and star trees is a *solid cover* if for every pair of distinct vertices i and j either:

1. there is an edge between i and j,
2. either i or j belongs to a clique in the cover,
3. either i or j is the center of a star tree in the cover, or
4. both i and j are leaves of the same star tree.

Proposition 9 *Let M be a $0/1$ dissimilarity matrix. Let r be the minimal number of graphs in a cover of G_M by cliques and star trees, such that every edge (but not necessarily every vertex) is in some element of the cover. Then M has star tree rank either r or $r+1$.*

Moreover, if G_M has a solid cover by r graphs, then M has star tree rank r.

In contrast to symmetric Barvinok rank, the upper bound of $n-2$ on the star tree rank of an $n \times n$ dissimilarity matrix cannot be achieved by a $0/1$ matrix for large n. Recall that the *Ramsey number* $R(k,k)$ is the smallest integer such that any graph on at least $R(k,k)$ vertices has either a clique or a independent set of size k. Then we have the following stronger bound on the star tree rank of a $0/1$ matrix.

Proposition 10 *For $n \geq R(k,k)$, any $n \times n$ $0/1$ dissimilarity matrix has star tree rank at most $n-k+1$.*

Proof: By the assumption on n, the graph G_M has either a clique of size k or an independent set of size k. In the former case, we can cover G_M by a star tree centered at each vertex not part of the clique, together with the clique itself. This gives a solid cover by $n-k+1$ subgraphs. In the latter case, we can just take the star trees centered at the vertices not in the independent set, giving a cover of G_M by $n-k$ subgraphs. In either case, Proposition 9 shows that M has rank at most $n-k+1$. $\square$

Corollary 11 *For $n \geq 18$, every $n \times n$ $0/1$ dissimilarity matrix has star tree rank at most $n-3$.*

Proof: The Ramsey number $R(4,4)$ is 18 [8]. □

In [5, Theorem 2], Drton, Sturmfels, and Sullivant prove a dimension theorem for the secant varieties of the classical Veronese projected to off-diagonal entries. Here, we prove the tropical analogue of their result.

Theorem 12 *Let r and n be positive integers. Then the dimension of the space of dissimilarity $n \times n$ matrices of star tree rank at most r is*

$$\min\left\{\binom{n+1}{2} - \binom{n-r+1}{2}, \binom{n}{2}\right\}.$$

In fact, the difficult part of Theorem 2 in [5] is proving the lower bound on the dimension of the classical secant variety. Our computation of the dimension of the tropical secant set, combined with the Bieri-Groves Theorem [2, Theorem A], provides an alternative proof of this lower bound.

5 Tree rank

The tropical Grassmannian $G_{2,n}$ is the tropical variety defined by the 3-term Plücker relations:

$$x_{ij}x_{k\ell} \oplus x_{ik}x_{j\ell} \oplus x_{i\ell}x_{jk} \qquad \text{for all } i < j < k < \ell. \tag{5}$$

This condition is equivalent to coming from the distances along a weighted tree which has negative weights along the internal edges [9, Sec. 4]. In this section, we will always take the deficiency graph to be with respect to the Plücker relations in (5).

As with the previous notions of rank, the tree rank of a $0/1$ matrix can be characterized in terms of covers of graphs. For any disjoint subsets $I_1, \ldots, I_k \subset [n]$ (not necessarily a partition), the *complete k-partite graph* is the graph which has an edge between the elements of I_i and I_j for all $i \neq j$. Complete k-partite graphs are characterized by the property that among vertices which are incident to some edge, the relation of having a non-edge is a transitive relation.

Note that the complete k-partite graphs defined above are exactly those graphs whose edge set forms the set of bases of a rank 2 matroid on n elements. The transitivity of being a non-edge is equivalent to the basis exchange axiom. Alternatively, each of the sets $I_1, \ldots, I_k$ partition the set of non-loops in the matroid into parallel classes. See [6] for definitions of these terms. In the following proposition, we will see that the Plücker relations imply the basis exchange axiom for the 0 entries of a non-negative tree matrix.

Proposition 13 *Let M be an $n \times n$ $0/1$ dissimilarity matrix and let r be smallest size of a cover of G_M by complete k-partite subgraphs. As in Proposition 9, we only require every edge to be in the cover, not necessarily every vertex. If G_M has at most one isolated vertex then M has tree rank r. Otherwise, M has tree rank $r+1$.*

Note that by taking the I_i in the definition of k-partite graph to be singletons, we get complete graphs, and by taking $k = 2$ with I_1 a singleton and I_2 any set disjoint from I_1, we get star trees. Together with Propositions 9 and 13, this confirms, for $0/1$-matrices, the second inequality in (2).

Again, we can show that the tropical secant sets and classical secant varieties agree in dimension:

n	maximum tree rank	example
3	1	
4	2	
5	3	$0/1$ matrix corresponding to 5-cycle
6	3	
7	4	
8	5	
9	6	M in (6)
10	6 or 7	Any extension of M in (6)
$9k$	between $6k$ and $9k-3$	M_k from discussion following (6)

Tab. 1: Maximum possible tree rank of an $n \times n$ dissimilarity matrix, to the best of our knowledge. The upper bounds come from Theorems 8 and 15. The examples have the largest tree ranks that are known to us. The omitted examples can be provided by taking a principal submatrix of a larger example, by Lemma 16.

Proposition 14 *The dimension of the set of dissimilarity $n \times n$ matrices of tree rank at most r is the dimension of the classical secant variety,*

$$\begin{aligned} \binom{n}{2} - \binom{n-2r}{2} \quad & \text{if } r \leq \frac{n}{2}, \\ \binom{n}{2} \quad & \text{if } r \geq \frac{n-1}{2}. \end{aligned}$$

Unlike the cases of symmetric Barvinok rank and star tree rank, we do not know the maximum tree rank of a $n \times n$ dissimilarity matrix for large n. We have an upper bound of $n-2$ by Theorem 8, and we can improve on this slightly:

Theorem 15 *For $n \geq 6$, an $n \times n$ dissimilarity matrix M has tree rank at most $n-3$.*

Beginning with $n = 10$, we don't know whether or not the bound in Theorem 15 is sharp. For the following 9×9 matrix, found by random search, the deficiency graph was computed to have chromatic number 6:

$$M = \begin{bmatrix} * & 1 & 6 & 7 & 2 & 3 & 8 & 9 & 6 \\ 1 & * & 2 & 7 & 9 & 7 & 5 & 7 & 1 \\ 6 & 2 & * & 6 & 0 & 6 & 1 & 7 & 1 \\ 7 & 7 & 6 & * & 3 & 3 & 8 & 5 & 3 \\ 2 & 9 & 0 & 3 & * & 5 & 7 & 5 & 7 \\ 3 & 7 & 6 & 3 & 5 & * & 9 & 3 & 9 \\ 8 & 5 & 1 & 8 & 7 & 9 & * & 2 & 3 \\ 9 & 7 & 7 & 5 & 5 & 3 & 2 & * & 8 \\ 6 & 1 & 1 & 3 & 7 & 9 & 3 & 8 & * \end{bmatrix} \tag{6}$$

Together with Theorem 15, this computation shows that M has tree rank 6. For any $k \geq 1$, we can form an $9k \times 9k$ matrix M_k by putting M in blocks along the diagonal and setting all other entries to 10. The deficiency graph of M_k includes k copies of the deficiency graph of M, and all edges between distinct copies. Therefore, the chromatic number, and thus the tree rank, are at least $6k$.

On the other hand, in order to provide examples of an $n \times n$ matrix with tree rank $n-3$ for all $n \leq 9$, we have the following lemma.

Lemma 16 *Let M be an $n \times n$ matrix. If any $(n-m) \times (n-m)$ principal submatrix has tree rank r, then M has tree rank at most $r+m$.*

Proof: Fix a decomposition of the $(n-m) \times (n-m)$ principal submatrix into r tree matrices. We can extend each tree matrix to an $n \times n$ tree matrix by adding leaf edges with large positive weights. For each index i not in the principal submatrix, define v_i to be the vector which is $C + M_{ij}$ in the jth entry and $-C$ in the ith entry, where C is a large real number. Then, the extended tree matrices, together with $\pi(v_i^T \odot v_i)$ for all i not in the principal submatrix, give a decomposition of M into $r+m$ tree matrices, as desired. $\square$

These results on the maximum tree rank are summarized in Table 1.

6 Star tree rank and tree rank for $n = 5$

In this section, we characterize the secant sets of the space of star trees and of the tropical Grassmannian in the case $n = 5$. Both give examples where the lower bound of Proposition 1 is actually an equality.

6.1 Star tree rank for $n = 5$

From Theorem 8, we know that the maximum star tree rank of a 5×5 matrix is 3. On the other hand, the set of dissimilarity matrices of star tree rank 1 is defined by the 2×2 minors. Thus, our task is to describe the second secant set of the space of star trees, i.e. the set of dissimilarity matrices of star tree rank 2.

First, we recall the defining ideal of the classical secant variety. The space of star trees is the tropicalization of the projection of the rank 1 symmetric matrices onto their off-diagonal entries. Its second secant variety is a hypersurface in $\mathbb{C}^{10}$ defined by the following 12-term quintic, known as the pentad [5]:

$$\begin{aligned} &x_{12}x_{13}x_{24}x_{35}x_{45} - x_{12}x_{13}x_{25}x_{34}x_{45} - x_{12}x_{14}x_{23}x_{35}x_{45} + x_{12}x_{14}x_{25}x_{34}x_{35} \\ &+ x_{12}x_{15}x_{23}x_{34}x_{45} - x_{12}x_{15}x_{24}x_{34}x_{35} + x_{13}x_{14}x_{23}x_{25}x_{45} - x_{13}x_{14}x_{24}x_{25}x_{35} \\ &- x_{13}x_{15}x_{23}x_{24}x_{45} + x_{13}x_{15}x_{24}x_{25}x_{34} - x_{14}x_{15}x_{23}x_{25}x_{34} + x_{14}x_{15}x_{23}x_{24}x_{35} \end{aligned}$$

Note that the 12 terms of the pentad correspond to the 12 different cycles on 5 vertices. The second secant set of the space of star trees is contained in the tropicalization of the pentad, but the containment is proper. Nonetheless, the terms of the pentad play a fundamental role in characterizing matrices of rank at most 2.

Theorem 17 *Let M be a 5×5 dissimilarity matrix. The following are equivalent:*

1. *M has star tree rank at most 2;*
2. *The deficiency graph of M is 2-colorable;*
3. *The minimum of the terms of the pentad is achieved at two terms which satisfy the following:*
   (a) *The terms differ by a transposition;*
   (b) *Assuming, without loss of generality, that the minimized terms are $x_{12}x_{23}x_{34}x_{45}x_{15}$ and $x_{13}x_{23}x_{24}x_{45}x_{15}$, then we have that $M_{14} + M_{23} \leq M_{12} + M_{34} = M_{13} + M_{24}$.*

The proof is similar in spirit to the proof of Theorem 18 below, and we omit it.

6.2 Tree rank for $n = 5$

We now turn our attention to tree rank of 5×5 dissimilarity matrices. As in the case of star tree rank, the maximum tree rank is 3, and so it suffices to characterize 5×5 dissimilarity matrices of tree rank at most 2. Unlike the previous case, the second classical secant variety is already all of $\mathbb{C}^{10}$, so there is no classical polynomial whose tropicalization gives us a clue to the tropical secant set. However, the tropical pentad again shows up in our characterization.

First, here is a simple example of a 5×5 $0/1$ dissimilarity matrix with tree rank 3. Consider the $0/1$ matrix corresponding to the 5-cycle C_5. Now, C_5 cannot be covered by fewer than 3 k-partite graphs, and so the matrix has tree rank at least 3 by Proposition 13. On the other hand, it has tree rank at most 3 by Theorem 8 and the inequality in (2). We will see in Remark 1 that this matrix is, in a certain sense, the only such example.

Let P be the tropical polynomial in variables $\{x_{ij} : 1 \leq i < j \leq 5\}$ which is the tropical sum of the 22 tropical monomials of degree 5 in which each $i \in \{1, \ldots, 5\}$ appears in a subscript exactly twice. Thus P has 12 monomials of the form $x_{12}x_{23}x_{34}x_{45}x_{15}$, forming the terms of the pentad, and 10 new monomials of the form $x_{12}x_{23}x_{31}x_{45}^2$. Let us call terms of the former kind *pentagons*, and terms of the latter kind *triangles*.

Theorem 18 *Let M be a 5×5 dissimilarity matrix. Then the following are equivalent:*

1. *M has tree rank at most 2;*
2. *The deficiency graph is 2-colorable;*
3. *The tropical polynomial P achieves its minimum at a triangle.*

Proof: First, (1) implies (2) by Proposition 1.

For (2) implies (3), we prove the contrapositive. Suppose the minimal terms of P are all pentagons; without loss of generality, we assume that $x_{12}x_{23}x_{34}x_{45}x_{15}$ is a minimal term. Since $x_{14}x_{45}x_{15}x_{23}^2$ is not minimal, we have $M_{12} + M_{34} < M_{14} + M_{23}$. Similarly, we have,

$$M_{12} + M_{23} + M_{34} + M_{45} + M_{15} < 2M_{15} + M_{23} + M_{34} + M_{24}, \text{ and}$$
$$M_{12} + M_{23} + M_{34} + M_{45} + M_{15} < 2M_{45} + M_{12} + M_{23} + M_{13}.$$

Adding these together and cancelling, we get $M_{12} + M_{34} < M_{13} + M_{24}$. Thus, 12 and 34 are adjacent in the deficiency graph. By similar reasoning, we have adjacencies $12 - 34 - 15 - 23 - 45 - 12$ in the deficiency graph, so it has a five cycle and is not 2-colorable.

Finally, we prove that (3) implies (1). Assume without loss of generality that $x_{34}x_{35}x_{45}x_{12}^2$ is among the terms minimizing P. This implies that $x_{12}x_{34}$, $x_{12}x_{35}$, and $x_{12}x_{45}$ are each minimal terms in their respective Plücker equations. Then we can use Lemmas 19 and 20, whose proofs we omit, to obtain a decomposition of M into two tree matrices.

Lemma 19 *For any 5×5 dissimilarity matrix M such that $x_{12}x_{34}$, $x_{12}x_{35}$, and $x_{12}x_{45}$ are each minimal terms in their respective Plücker equations, there exists some 5×5 tree matrix T, such that for every $ij \in \binom{[5]}{2}$, we have $T_{ij} \geq M_{ij}$, with equality if $ij \in \{12, 13, 14, 15, 23, 24, 25\}$.*

Lemma 20 *For any 5×5 dissimilarity matrix M, there exists some 5×5 tree matrix T' such that for every pair of indices i and j, we have $T'_{ij} \geq M_{ij}$, with equality if $ij \in \{34, 35, 45\}$.*

We can now finish the proof of Theorem 18. Let T be as given in Lemma 19 and let T' be as given in Lemma 20. Then $M = T \oplus T'$ and so M has tree rank at most 2. □

In fact, we can describe precisely which subgraphs of the Petersen graph arise as deficiency graphs Δ_M for $n = 5$. There are 5 tropical Plücker relations on a 5×5 matrix, each containing 3 terms. Each term is the tropical product of terms with disjoint entries. Thus, Δ_M is a subgraph with at most 5 edges of the Petersen graph.

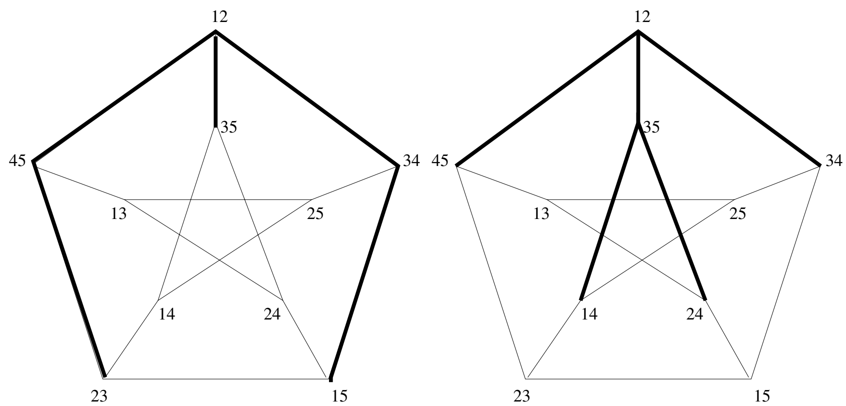


Fig. 1: The two 2-colorable possibilities for Δ_M.

Theorem 21 *Let M be a 5×5 dissimilarity matrix. Then the deficiency graph Δ_M is precisely one of the following:*

1. *The trivial graph, in which case M has tree rank* 1.
2. *A non-trivial graph with fewer than* 5 *edges, in which case M has tree rank* 2.
3. *Up to relabeling, either of the two graphs in Figure 1, in which case M has tree rank* 2.
4. *A* 5*-cycle, in which case M has tree rank* 3.

Proof: The matrix M is a tree matrix if and only if the four-point condition holds for all 4-tuples, i.e. if and only if Δ_M is trivial. This is the first case.

Now suppose that Δ_M is a non-trivial graph with at most 4 edges. Then, at least one four-point condition holds, so Lemma 16 implies that M has tree rank at most 2. However, at least one four-point condition is violated, so M must have tree rank exactly 2. We omit the case analysis that shows that, up to relabeling, the only two possibilities for Δ_M, assuming that it is 2-colorable, are those in Figure 1.

Finally, if Δ_M is not 2-colorable, then it must have an odd cycle. The Petersen graph has no 3-cycles, so Δ_M must be a 5-cycle, since it has at most 5 edges. □

Remark 1 *If M is the $0/1$ matrix corresponding to the 5-cycle C_5, then Δ_M is also a 5-cycle by Theorem 21. Explicitly, Δ_M has an edge for each non-adjacent pair of edges in the graph C_5. Moreover, Theorem 21 tells us that any other matrix N with tree rank 3 must have the same deficiency graph (up to relabeling). In this sense, M is the only example of a 5×5 dissimilarity matrix with tree rank 3.*

Acknowledgements

We thank Bernd Sturmfels for his guidance and a close reading of the text. Maria Angélica Cueto helped us understand the connection to phylogenetics.

References

[1] M. Akian, S. Gaubert, A. Guterman, Linear independence over tropical semirings and beyond, preprint, arXiv:0812.3496, 2008.

[2] R. Bieri, J. Groves, The geometry of the set of characters induced by valuations, J. Reine. Angew. Math. 347:168–195, 1984.

[3] M. Develin, F. Santos, B. Sturmfels, On the rank of a tropical matrix, in "Discrete and Computational Geometry" (E. Goodman, J. Pach and E. Welzl, eds), MSRI Publications, Cambridge University Press, 2005.

[4] J. Draisma, A tropical approach to secant dimensions, Journal of Pure and Applied Algebra, 212(2):349–363, 2008.

[5] M. Drton, B. Sturmfels, S. Sullivant, Algebraic factor analysis: tetrads, pentads and beyond, Probability Theory and Related Fields 138(3/4): 463–493, 2007.

[6] J. Oxley, Matroid Theory, Oxford Univ. Press, New York, 1992.

[7] L. Pachter, B. Sturmfels, Algebraic Statistics for Computational Biology, Cambridge University Press, Cambridge, 2005.

[8] S. Radziszowski, Small Ramsey Numbers, Electronic Journal of Combinatorics, Dynamic survey DS1. updated 2009.

[9] D. Speyer, B. Sturmfels, The tropical Grassmannian, Adv. Geom., 4(3):389–411, 2004.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 95–106

Models and refined models for involutory reflection groups and classical Weyl groups

Fabrizio Caselli[1] and Roberta Fulci[1]

[1] *Dipartimento di Matematica, Università di Bologna, P.zza di Porta S. Donato 5, 40126 Bologna, Italy.*
E-mail address: `caselli@dm.unibo.it, fulci@dm.unibo.it`

Abstract. A finite subgroup G of $GL(n,\mathbb{C})$ is involutory if the sum of the dimensions of its irreducible complex representations is given by the number of absolute involutions in the group, i.e. elements $g \in G$ such that $g\bar{g} = 1$, where the bar denotes complex conjugation. A uniform combinatorial model is constructed for all non-exceptional irreducible complex reflection groups which are involutory including, in particular, all infinite families of finite irreducible Coxeter groups. If G is a classical Weyl group this result is much refined in a way which is compatible with the Robinson-Schensted correspondence on involutions.

Résumé. Un sousgroupe fini G de $GL(n,\mathbb{C})$ est dit involutoire si la somme des dimensions de ses representations irréductibles complexes est donné par le nombre de involutions absolues dans le groupe, c'est-a-dire le nombre de éléments $g \in G$ tels que $g\bar{g} = 1$, où le bar dénotes la conjugaison complexe. Un model combinatoire uniform est construit pour tous les groupes de réflexions complexes irréductibles qui sont involutoires, en comprenant, toutes les familles de groupes de Coxeter finis irreductibles. Si G est un groupe de Weyl ce resultat peut se raffiner dans une manière compatible avec la correspondence de Robinson-Schensted sur les involutions.

Keywords: Complex reflection groups, Gelfand models, Classical Weyl groups

1 Introduction

In their paper [7] Bernstein, Gelfand and Gelfand introduced the problem of the construction of a model of a group G, i.e. a representation which is the direct sum of all irreducible complex representations of G with multiplicity one. We can find several constructions of models in the literature for the symmetric group [2, 3, 13, 15, 16, 17] and for some other special classes of complex reflection groups [1, 4, 5, 6].
A complex reflection group, or simply a reflection group, is a subgroup of $GL(V)$, where V is a finite dimensional complex vector space, generated by reflections, i.e. by elements of finite order which fix a hyperplane pointwise. There is a well-known classification of irreducible reflection groups due to Shephard-Todd [20] including an infinite family $G(r,p,n)$ depending on 3 parameters together with 34 exceptional cases. As mentioned above one can find in the literature models for some reflection groups such as the wreath product groups $G(r,1,n)$ as well as the groups $G(2,2,n)$, which are better known as the Weyl groups of type D.
If G is a finite subgroup of $GL(n,\mathbb{C})$, a specialization of a theorem of Bump and Ginzburg [8] gives a

combinatorial description of the character of a model of the group G if its dimension is given by the number of absolute involutions of G (i.e. elements $g \in G$ such that $g\bar{g} = 1$). We say that a group satisfying this condition is involutory. It turns out that a complex reflection group $G(r, p, n)$ is involutory if and only if $\mathrm{GCD}(p, n) = 1, 2$ and that one can construct an explicit model for all these groups in a uniform way. This construction involves in a crucial way the theory of projective reflection groups developed in [9]. Indeed a byproduct of this construction is also a model for some related projective reflection groups.

The model of the group G considered in this paper has a basis indexed by the absolute involutions of the dual group G^* (see §2) and it is clear from the definition that the subspace spanned by the basis elements indexed by the absolute involutions in a symmetric conjugacy class is a submodule. If G is a classical Weyl group we show that any such submodule is given by the sum of all irreducible representations indexed by the shapes corresponding to the indexing involutions by means of the projective Robinson-Schensted correspondence. This decomposition becomes particularly interesting for Weyl groups of type D with respect to the so-called split representations.

The paper is organized as follows. In §2 we collect the notation and the preliminary results which are needed. In §3 we classify all projective reflection groups of the form $G(r, p, q, n)$ (see §2 for the definition) which are involutory. In §4 we show an explicit model for all involutory reflection groups. In §5 a first decomposition is given for the model of the generic involutory reflection group $G(r, p, n)$, which reflects the existence of the split representations. In §6 and §7 a finer decomposition is given for the groups of type B_n and D_n.

2 Notation and preliminaries

In this section we collect the notations that are used in this paper as well as the preliminary results that are needed.

We let $\mathbb{Z}$ be the set of integer numbers and $\mathbb{N}$ be the set of nonnegative integer numbers. For $a, b \in \mathbb{Z}$, with $a \leq b$ we let $[a, b] \stackrel{\text{def}}{=} \{a, a + 1, \ldots, b\}$ and, for $n \in \mathbb{N}$ we let $[n] \stackrel{\text{def}}{=} [1, n]$. For $r \in \mathbb{N}$ we let $\mathbb{Z}_r \stackrel{\text{def}}{=} \mathbb{Z}/r\mathbb{Z}$. If $r \in \mathbb{N}$, $r > 0$, we denote by ζ_r the primitive r-th root of unity $\zeta_r \stackrel{\text{def}}{=} e^{\frac{2\pi i}{r}}$.

The main subject of this work are the complex reflection groups, or simply reflection groups, with particular attention to their combinatorial representation theory. The most important example of a complex reflection group is the group of permutations of $[n]$, known as the symmetric group, that we denote by S_n. We know by the work of Shephard-Todd [20] that all but a finite number of irreducible reflection groups are the groups $G(r, p, n)$ that we are going to describe. If A is a matrix with complex entries we denote by $|A|$ the real matrix whose entries are the absolute values of the entries of A. The *wreath product* groups $G(r, n) = G(r, 1, n)$ are given by all $n \times n$ matrices satysfying the following conditions:

- the non-zero entries are r-th roots of unity;
- there is exactly one non-zero entry in every row and every column (i.e. $|A|$ is a permutation matrix).

If p divides r then the reflection group $G(r, p, n)$ is the subgroup of $G(r, n)$ given by all matrices $A \in G(r, n)$ such that $\frac{\det A}{\det |A|}$ is a $\frac{r}{p}$-th root of unity.
Following [9], a *projective reflection group* is a quotient of a reflection group by a scalar subgroup. Observe that a scalar subgroup of $G(r, n)$ is necessarily a cyclic group of the form $C_q =< \zeta_q I >$ of order q, for some $q|r$.

It is also easy to characterize all possible scalar subgroups of the groups $G(r,p,n)$: in fact the scalar matrix $\zeta_q I$ belongs to $G(r,p,n)$ if and only if $q|r$ and $pq|rn$. In this case we let $G(r,p,q,n) \stackrel{\text{def}}{=} G(r,p,n)/C_q$. If $G = G(r,p,q,n)$ then the projective reflection group $G^* \stackrel{\text{def}}{=} G(r,q,p,n)$, where the roles of the parameters p and q are interchanged, is always well-defined. We say that G^* is the *dual* of G and we refer the reader to [9] for the main properties of this duality. In this work we will see another important occurrence of the relationship between a group G and its dual G^*.

If the non-zero entry in the i-th row of $g \in G(r,n)$ is $\zeta_r^{z_i}$ we let $z_i(g) \stackrel{\text{def}}{=} z_i \in \mathbb{Z}_r$ and say that $z_1(g), \ldots, z_n(g)$ are the *colors* of g. We can also note that g belongs to $G(r,p,n)$ if and only if $z(g) \stackrel{\text{def}}{=} \sum z_i(g) \equiv 0 \mod p$.

For $g \in G(r,n)$ we let $|g| \in S_n$ be the permutation defined by $|g|(i) = j$ if $g_{i,j} \neq 0$. We may observe that an element $g \in G(r,n)$ is uniquely determined by the permutation $|g|$ and by its colors $z_i(g)$ for all $i \in [n]$.

If $g \in G(r,n)$ we let $\bar{g} \in G(r,n)$ be the complex conjugate of g. We can also observe that $\bar{g}$ is determined by the conditions $|\bar{g}| = |g|$ and $z_i(\bar{g}) = -z_i(g)$ for all $i \in [n]$. Since the bar operator stabilizes the cyclic subgroup $C_q =< \zeta_q I >$ it is well-defined also on the projective reflection groups $G(r,p,q,n)$.

In [9] we can find a parametrization of the irreducible representations of the groups $G(r,p,q,n)$, that we briefly recall for the reader's convenience. Given a partition $\lambda = (\lambda_1, \ldots, \lambda_l)$ of n, the *Ferrers diagram of shape* λ is a collection of boxes, arranged in left-justified rows, with λ_i boxes in row i. We denote by $\mathrm{Fer}(r,n)$ the set of r-tuples $(\lambda^{(0)}, \ldots, \lambda^{(r-1)})$ of Ferrers diagrams such that $\sum |\lambda^{(i)}| = n$. If $\mu \in \mathrm{Fer}(r,n)$ we define the *color* of μ by $z(\mu) = \sum_i i|\lambda^{(i)}|$ and, if $p|r$ we let $\mathrm{Fer}(r,p,n) \stackrel{\text{def}}{=} \{\mu \in \mathrm{Fer}(r,n) : z(\mu) \equiv 0 \mod p\}$. If $q \in \mathbb{N}$ is such that $q|r$ and $pq|nr$ then the cyclic group C_q acts on $\mathrm{Fer}(r,p,n)$ by a shift of r/q positions of its elements (see [9, Lemma 6.1]). Paralleling the definition for the projective reflection groups we denote the corresponding quotient set by $\mathrm{Fer}(r,p,q,n)$. We denote by $(C_p)_\mu$ the stabilizer of μ in C_p. For example, if

$$\mu = \left[\ydiagram{2,1}, \ydiagram{1}, \ydiagram{2,1}, \ydiagram{1}\right] \text{ and } \mu' = \left[\ydiagram{1}, \ydiagram{2,1}, \ydiagram{1}, \ydiagram{2,1}\right],$$

then μ and μ' are elements in $\mathrm{Fer}(4,2,8)$ which represent the same class in $\mathrm{Fer}(4,2,4,8)$. We also observe that in this case the stabilizer in $C_4(\mu) = C_4(\mu')$ is the cyclic group C_2 of order 2.

Proposition 2.1 *The irreducible complex representations of $G(r,p,q,n)$ can be parametrized by pairs (μ, ρ), where $\mu \in \mathrm{Fer}(r,q,p,n)$ and $\rho \in (C_p)_\mu$, where $(C_p)_\mu$ is the stabilizer of any element in the class μ by the action of C_p.*

If $\mu \in \mathrm{Fer}(r,n)$ we denote by $\mathcal{ST}_\mu$ the set of all possible fillings of the boxes in μ with all the numbers from 1 to n appearing once, in such way that rows are increasing from left to right and columns are incresing from top to bottom in every single Ferrers diagram of μ. We let $\mathcal{ST}(r,n) \stackrel{\text{def}}{=} \cup_{\mu \in \mathrm{Fer}(r,n)} \mathcal{ST}_\mu$ and we define $\mathcal{ST}(r,p,n)$ and $\mathcal{ST}(r,p,q,n)$ as already done for Ferrers diagrams. For example, the two elements

$$T = \left[\begin{ytableau} 2 & 8 \\ 4 \end{ytableau}, \begin{ytableau} 1 \end{ytableau}, \begin{ytableau} 3 & 5 \\ 7 \end{ytableau}, \begin{ytableau} 6 \end{ytableau}\right] \text{ and } T' = \left[\begin{ytableau} 6 \end{ytableau}, \begin{ytableau} 2 & 8 \\ 4 \end{ytableau}, \begin{ytableau} 1 \end{ytableau}, \begin{ytableau} 3 & 5 \\ 7 \end{ytableau}\right]$$

belong to $\mathcal{ST}(4,2,8)$ and represent the same class in $\mathcal{ST}(4,2,4,8)$.

The classical Robinson-Schensted correspondence [22, §7.11] for the symmetric groups was generalized to the Stanton-White correspondence [23] for the wreath products $G(r,n)$. A further generalization of the correspondence, which is valid for all projective reflection groups $G(r,p,q,n)$, is explicitely shown in [9, §10]. We refer to this correspondence as the *projective Robinson-Schensted correspondence*. We do not describe this correspondence explicitly, but we briefly state it for future reference.

Theorem 2.2 *There exists an explicit map*

$$\begin{array}{rcl} G(r,p,q,n) & \longrightarrow & \mathcal{ST}(r,p,q,n) \times \mathcal{ST}(r,p,q,n) \\ g & \longmapsto & [P(g), Q(g)], \end{array}$$

satisfying the following properties:

1. *$P(g)$ and $Q(g)$ have the same shape in $\mathrm{Fer}(r,p,q,n)$ for all $g \in G(r,p,q,n)$;*
2. *if $P, Q \in \mathcal{ST}(r,p,q,n)$ have the same shape $\mu \in \mathrm{Fer}(r,p,q,n)$ then*

$$|\{g \in G(r,p,q,n) : P(g) = P \text{ and } Q(g) = Q\}| = |(C_q)_\mu|,$$

$(C_q)_\mu$ being, as above, the stabilizer in C_q of any element in the class μ.

If G is a finite group we let $\mathrm{Irr}(G)$ be the set of irreducible complex representations of G. If M is a complex vector space and $\rho : G \to GL(M)$ is a representation of G we say that the pair (M, ρ) is a G-*model* if the character χ_ρ is the sum of the characters of all irreducible representations of G over $\mathbb{C}$, i.e. M is isomorphic as a G-module to the direct sum of all irreducible modules of G with multiplicity one. Sometimes we simply say that M is a G-model if we do not need to know the map ρ explicitly or if it is clear from the context. It is clear that two G-models are always isomorphic as G-modules, and so we can also speak about "the" G-model. The last result in this section is a beautiful theorem of Bump and Ginzburg, which generalizes a classical theorem of Frobenius and Schur [11], and allows us in some cases to determine the character of the model of a finite group if we know its dimension.

Theorem 2.3 ([8], Theorem 7) *Let G be a finite group, $\tau \in Aut(G)$ with $\tau^2 = 1$ and M be a G-model. Assume that*

$$\dim(M) = \#\{g \in G : g\tau(g) = z\},$$

where z is a central element in G such that $z^2 = 1$. Then

$$\chi_M(g) = \#\{u \in G : u\tau(u) = gz\}.$$

3 Involutory projective reflection groups

In this section we start the investigation of a model for the projective reflection groups $G(r,p,q,n)$. The main result here is the characterization of the groups $G(r,p,q,n)$ such that the dimension of a $G(r,p,q,n)$-model is equal to the number of absolute involutions in $G(r,p,q,n)$. In these groups we can directly apply Theorem 2.3 to obtain a combinatorial description of the character of the model. The next result relates the dimension of a model with the projective Robinson-Schensted correspondence.

Proposition 3.1 *Let $G = G(r,p,q,n)$. The dimension of a G-model is equal to the number of elements g in the dual group G^* which correspond by means of the projective Robinson-Schensted correspondence to pairs of the form $[P,P]$, for some $P \in \mathcal{ST}(r,q,p,n)$.*

The next target is to show that absolute involutions in G^* correspond to pairs of the form $[P, P]$ under the projective Robinson-Schensted correspondence, and then to characterize those groups for which the converse holds, i.e. the groups where the fact that $v \mapsto [P, P]$ implies that v is an absolute involution.

If $g \in G(r, p, q, n)$, we say that g is a *symmetric* element if any (equivalently every) lift of g in $G(r, n)$ is a symmetric matrix. We similarly define *antisymmetric* elements in $G(r, p, q, n)$. Observe that we can have antisymmetric elements only if r is even. The following result is a characterization of absolute involutions in $G(r, p, q, n)$.

Lemma 3.2 *Let $g \in G(r, p, q, n)$. Then g is an absolute involution, i.e. $g\bar{g} = 1$, if and only if either g is symmetric or q is even and g is antisymmetric.*

We denote by $I(r, p, q, n)$ the set of absolute involutions in $G(r, p, q, n)$.

Theorem 3.3 *Let $G = G(r, p, q, n)$. Then*

$$\sum_{\phi \in \mathrm{Irr}(G)} \dim \phi \geq |I(r, q, p, n)|$$

and equality holds if and only if either $\mathrm{GCD}(p, n) = 1, 2$*, or* $\mathrm{GCD}(p, n) = 4$ *and* $r \equiv p \equiv q \equiv n \equiv 4 \mod 8$.

We conclude this section by observing that a projective reflection group $G = G(r, p, q, n)$ and its dual group G^* always have the same number of absolute involutions. This fact will be the keypoint in the description of the character of the model for the groups satisfying the conditions of Theorem 3.3.

Proposition 3.4 *We always have $|I(r, p, q, n)| = |I(r, q, p, n)|$.*

The proof of this proposition is by direct computation. A “nice” bijective proof is desirable.

We say that a projective reflection group $G = G(r, p, q, n)$ is *involutory* if the dimension of a model of G is equal to the number of absolute involutions in G. By Proposition 3.4 we have that $G(r, p, q, n)$ is involutory if and only if it satisfies the conditions in Theorem 3.3.

If we restrict our attention to standard reflection groups we may note that a group $G(r, p, n)$ is involutory if and only if $\mathrm{GCD}(p, n) = 1, 2$. In particular all infinite families of finite irreducible Coxeter groups (these are $A_n = G(1, 1, n)$, $B_n = G(2, 1, n)$, $D_n = G(2, 2, n)$, $I_2(r) = G(r, r, 2)$) are involutory. In the next section we establish a unified construction of a model for all involutory reflection groups (and the corresponding quotients).

The fact that $G(r, p, n)$ is involutory if $\mathrm{GCD}(p, n) = 1, 2$ can also be deduced from known results in the following alternative way. From the characterization of automorphism of complex reflection groups appearing in [18, §1] one can deduce that, under these hypothesis, any irreducible representation ϕ of $G(r, p, n)$ can be realized by a matrix representation $\phi : G(r, p, n) \to GL_n(\mathbb{C})$ satisfying $\phi(\bar{g}) = \overline{\phi(g)}$. Then a straightforward application of the twisted Schur-Frobenius theory developed in [14] implies that $G(r, p, n)$ is involutory.

4 Models

From the results of the previous section we have that the dimension of the model of an involutory reflection group G, is equal to the number of absolute involutions of G and also to the number of absolute involutions of G^*. In this section we show how we can give the structure of a G-model to the formal vector space having a basis indexed by the absolute involutions in G^*.

Unless otherwise stated, we let $G = G(r,p,n)$ be an involutory reflection group, i.e. such that $\mathrm{GCD}(p,n) = 1,2$. By Theorem 2.3 we have that the character χ of a G-model is given by

$$\chi(g) = |\{u \in G : u\bar{u} = g\}|.$$

Once we have an algebraic-combinatorial description of the dimension and of the character of a model for $G(r,p,n)$ we have two of the main ingredients of the proof of our main result. Before stating it, we need some more definitions. If $\sigma, \tau \in S_n$ with $\tau^2 = 1$ we let

$$\mathrm{Inv}(\sigma) = \{\{i,j\} : (j-i)(\sigma(j) - \sigma(i)) < 0\} \quad \text{and} \quad \mathrm{Pair}(\tau) = \{\{i,j\} : \tau(i) = j \neq i\}.$$

If $g \in G(r,p,n)$ and $v \in I(r,p,n)^*$ we let

$$\begin{aligned} s(g,v) &= \# \big(\mathrm{Inv}(|g|) \cap \mathrm{Pair}(|v|)\big) \\ a(g,v) &= z_1(\tilde{v}) - z_{|g|^{-1}(1)}(\tilde{v}) \in \mathbb{Z}_r \end{aligned}$$

where $\tilde{v}$ is any lift of v in $G(r,n)$. Note that since $a(g,v)$ is the difference of two colors of $\tilde{v}$ it is well-defined. Furthermore, given $g, g' \in G(r,n)$, we let

$$< g, g' >= \sum_i z_i(g) z_i(g') \in \mathbb{Z}_r.$$

Also, it is easy to see that, given $g \in G = G(r,p,n)$, the function of the dual group $G^* = G(r,1,p,n)$

$$\begin{aligned} T_g : G(r,1,p,n) &\to \mathbb{Z}_r \\ g' &\mapsto < g, g' > \end{aligned}$$

is well defined, i.e., taken any two lifts $\bar{g}$ and $\hat{g}$ of g' in $G(r,n)$, we have $< g, \bar{g} > \equiv < g, \hat{g} > \mod r$. We denote by $I(r,p,n)^* = I(r,1,p,n)$ the set of absolute involutions in G^* and we recall (Lemma 3.2) that these elements can be either symmetric or antisymmetric.

Theorem 4.1 *Let* $\mathrm{GCD}(p,n) = 1,2$ *and let*

$$M(r,p,n)^* \stackrel{\text{def}}{=} \bigoplus_{v \in I(r,p,n)^*} \mathbb{C} C_v$$

and $\varrho : G(r,p,n) \to GL(M(r,p,n)^*)$ *be defined by*

$$\varrho(g)(C_v) \stackrel{\text{def}}{=} \begin{cases} \zeta_r^{<g,v>} (-1)^{s(g,v)} C_{|g|v|g|^{-1}} & \textit{if } v \textit{ is symmetric} \\ \zeta_r^{<g,v>} \ \zeta_r^{a(g,v)} \ C_{|g|v|g|^{-1}} & \textit{if } v \textit{ is antisymmetric}. \end{cases} \tag{1}$$

Then $(M(r,p,n)^*, \varrho)$ *is a* $G(r,p,n)$*-model.*

The proof of this theorem consists in the explicit and rather involved computation of the character of this representation, and in verifying that this character agrees with the character described in Theorem 2.3.

If $q|r$ and $pq|rn$ (i.e. the group $G(r,p,q,n)$ is defined) we can consider the submodule $M(r,q,p,n) \subseteq M(r,p,n)^*$ spanned by all elements C_v such that $v \in I(r,q,p,n)$. The next result shows that $M(r,q,p,n)$ is the sum of all irreducible representations of $G(r,p,n)$ indexed by elements $\mu \in \mathrm{Fer}(r,q,p,n)$.

Corollary 4.2 *Let* $\mathrm{GCD}(p,n) = 1,2$. *Then the pair* $(M(r,q,p,n), \varrho)$, *where*

$$\varrho : G(r,p,q,n) \to GL(M(r,q,p,n))$$

is defined as in Theorem 4.1, is a $G(r,p,q,n)$*-model.*

We will see in the following sections several important generalizations of these results if the group G is a classical Weyl group.

5 Splitting split representations

If $\mathrm{GCD}(p,n) = 2$, there is another natural decomposition of $M(r,p,n)^*$ into two $G(r,p,n)$-submodules. The submodule $\mathrm{Sym}(r,p,n)^*$ spanned by symmetric elements and the submodule $\mathrm{Asym}(r,p,n)^*$ spanned by antisymmetric elements. Recall from Proposition 2.1 that an irreducible representation μ of $G(r,n)$ when restricted to $G(r,p,n)$ either remains irreducible if the stabilizer $(C_p)_\mu$ is trivial, or splits into two irreducible representations of $G(r,p,n)$ if $(C_p)_\mu$ has two elements (note that there are no other possibilities since $\mathrm{GCD}(p,n) = 2$), and that all irreducible representations of $G(r,p,n)$ are obtained in this way.

Theorem 5.1 *Let* χ *be the character of* $\mathrm{Sym}(r,p,n)^*$ *and* ϕ *be an irreducible representation of* $G(r,n)$. *If* ϕ *does not split in* $G(r,p,n)$ *then* $<\chi,\chi_\phi>=1$. *If* ϕ *splits into two irreducible representations* ϕ^+,ϕ^- *of* $G(r,p,n)$ *then*

$$<\chi,\chi_{\phi^+}>=1 \iff <\chi,\chi_{\phi^-}>=0.$$

If we restrict our attention to the case of Weyl groups $D_n = G(2,2,n)$, the proof of this result is based on the following observation which is a direct consequence of the explicit formulas for the split characters of the groups D_n (see [21, 19]).

Proposition 5.2 *Let* $g \in S_n$ *be of cycle-type* 2α. *Then one can label the split representations of* D_n *by* $(\lambda,\lambda)^+$ *and* $(\lambda,\lambda)^-$ *so that*

$$\sum_{\lambda\vdash n/2} (\chi_{(\lambda,\lambda)^+} - \chi_{(\lambda,\lambda)^-})(g) = 2^{\ell(\alpha)}\chi_M(\alpha),$$

where χ_M *is the character of the model for* $S_{n/2}$.

Consider now the two representations of D_n $(\mathrm{Asym}(2,2,n)^*,\rho^+)$ and $(\mathrm{Asym}(2,2,n)^*,\rho^-)$, given by

$$\rho^+(g)(C_v) \stackrel{\text{def}}{=} (-1)^{<g,v>}C_{|g|v|g|^{-1}}, \qquad \rho^-(g)(C_v) \stackrel{\text{def}}{=} (-1)^{<g,v>}(-1)^{a(g,v)}C_{|g|v|g|^{-1}}$$

(notice that $\rho^-(g) = \varrho(g)|_{\mathrm{Asym}(2,2,n)^*}$). An explicit computation of the characters of the representations ρ^+ and ρ^- and Proposition 5.2 show that

$$\sum_{\lambda\vdash n/2}\chi_{(\lambda,\lambda)^+}(g) - \sum_{\lambda\vdash n/2}\chi_{(\lambda,\lambda)^-}(g) = \chi_{\rho^+}(g) - \chi_{\rho^-}(g) \qquad \forall\, g \in D_n.$$

Comparing the dimensions of the representations involved, and recalling the linear independence of characters, we can conclude that

$$\chi_{\rho^+}(g) = \sum_{\lambda\vdash n/2}\chi_{(\lambda,\lambda)^+}(g) \quad \text{and} \quad \chi_{\rho^-}(g) = \sum_{\lambda\vdash n/2}\chi_{(\lambda,\lambda)^-}(g):$$

this means that $(\mathrm{Asym}(2,2,n)^*,\varrho) \cong \bigoplus_{\lambda\vdash n/2}(\lambda,\lambda)^-$, as claimed.

6 Refinement for B_n

Let us have a closer look at the model (M, ϱ) for $G = G(r,n)$. There is an immediate decomposition of M into submodules that we are going to describe.

Let $g, h \in G(r,n)$. We say that g and h are S_n-conjugate if there exists $\sigma \in S_n$ such that $g = \sigma h \sigma^{-1}$. If c is an S_n-conjugacy class of absolute involutions in G we denote by $M(c)$ the subspace of M spanned by the elements in c, and it is clear that

$$M = \bigoplus_c M(c) \quad \text{as } G\text{-modules},$$

where the sum runs through all S_n-conjugacy classes of absolute involutions. It is natural to ask if we can describe the irreducible decomposition of the submodules $M(c)$. This decomposition is known if G is the symmetric group S_n (see [1, 13]). We will focus on the case of B_n and we show that the irreducible decompositions of these submodules are well behaved with respect to the RS correspondences, a problem which was raised in [2]. The meaning of 'well behaved with respect to the RS correspondence' will be clarified in Theorem 6.1.

Let v be an involution of B_n. We denote by $R(v)$ the element of $\mathrm{Fer}(2,n)$ which is the shape of the tableaux of the image of v via the Robinson-Schensted correspondence. Namely $R(v) \stackrel{\text{def}}{=} (\lambda, \mu)$, where

$$v \xrightarrow{RS} [P,P], \quad P \in \mathcal{ST}(2,n), \quad P \text{ of shape } (\lambda,\mu).$$

For notational convenience we let $R(c) = \cup_{v \in c} R(v)$. The main goal of this section is the following result.

Theorem 6.1 *Let c be an S_n-conjugacy class of involutions in B_n. Then the following decomposition holds:*

$$M(c) \cong \bigoplus_{(\lambda,\mu) \in R(c)} \rho_{\lambda,\mu}.$$

In order to prove Theorem 6.1, first of all we need to parametrize the S_n-conjugacy classes of involutions explicitly. With this purpose we let

- $\mathrm{fix}(v) \stackrel{\text{def}}{=} \#\{i : i > 0 \text{ and } v(i) = i\}$
- $\mathrm{fix}^-(v) \stackrel{\text{def}}{=} \#\{i : i > 0 \text{ and } v(i) = -i\}$
- $\mathrm{pair}(v) \stackrel{\text{def}}{=} \#\{(i,j) : 0 < i < j,\ v(i) = j \text{ and } v(j) = i\}$
- $\mathrm{pair}^-(v) \stackrel{\text{def}}{=} \#\{(i,j) : 0 < i < j,\ v(i) = -j \text{ and } v(j) = -i\}$.

Proposition 6.2 *Two involutions v, w in B_n are S_n-conjugate if and only if*

$$\mathrm{fix}(v) = \mathrm{fix}(w), \quad \mathrm{pair}(v) = \mathrm{pair}(w),$$

$$\mathrm{fix}^-(v) = \mathrm{fix}^-(w), \quad \mathrm{pair}^-(v) = pair^-(w).$$

Furthermore, given an involution v in B_n, let $R(v) = (\lambda, \mu)$. Then λ has $\mathrm{fix}(v)$ *odd columns and* $\mathrm{fix}(v) + 2\,\mathrm{pair}(v)$ *boxes, while μ has* $\mathrm{fix}^-(v)$ *odd columns and* $\mathrm{fix}^-(v) + 2\,\mathrm{pair}^-(v)$ *boxes.*

We can thus name the S_n-conjugacy classes of the involutions of B_n in this way:

$$c_{f_0,f_1,p_0,p_1} \stackrel{\text{def}}{=} \{v : \text{fix}(v) = f_0;\ \text{fix}^-(v) = f_1;\ \text{pair}(v) = p_0;\ \text{pair}^-(v) = p_1\}.$$

The description given for the S_n-conjugacy classes ensures that the subspace M_0 of M generated by the involutions $v \in B_n$ with $\text{fix}(v) = \text{fix}^-(v) = 0$, is a B_n-submodule. The crucial step in the proof of Theorem 6.1 is the following partial result regarding this submodule (we observe that in this case n is necessarily even, $n = 2m$): M_0 is the direct sum of all the irreducible representations of B_{2m} indexed by pairs of diagrams whose columns have an even number of boxes, each of such representations occurring once. To show this we need the following argument which generalizes an idea appearing in [13].

Lemma 6.3 *Let Π_m be representations of B_{2m}, m ranging in $\mathbb{N}$. Then the following are equivalent:*

a) for every m, Π_m is the direct sum of all the irreducible representations of B_{2m} indexed by pairs of diagrams whose columns have an even number of boxes, each of such representations occurring once;

b) for every m,

(b0) Π_0 is unidimensional;

(b1) the following isomorphism holds:

$$\Pi_m \downarrow_{B_{2m-1}} \cong \Pi_{m-1} \uparrow^{B_{2m-1}}; \tag{2}$$

(b2) the module Π_m contains all the irreducible representations of B_{2m} indexed by the pairs of diagrams $(1^{2j}, 1^{2(m-j)})$, $j \in [0, m]$, where 1^k is the single Ferrers diagram with one column of length k.

This lemma can be proved constructively by means of a generalization to B_n of the branching rule (see [12]). The implication b)⇒ a) of the preceding lemma can be applied to the case $\Pi_m = M_0$.

The group B_0 is the identity group so property b0) is trivially verified.

Let us denote by N_0 the B_{2m-2}-module constructed in the same way. To check property b1), we have to show that

$$M_0 \downarrow_{B_{2m-1}} \cong N_0 \uparrow^{B_{2m-1}}\ . \tag{3}$$

The following argument is used. Let M_0^h be the submodule of M_0 generated by the involutions v satisfying $\text{fix}(v) = \text{fix}^-(v) = 0$, $\text{pair}(v) = h$ and $\text{pair}^-(v) = m - h$. Each M_0^h, once restricted to B_{2m-1}, splits into two submodules according to the color of $2m$. We denote by $M_0^{h,+}$ the submodule of M_0^h containing involutions v such that $z_{2m}(v) = 0$, and similarly for $M_0^{h,-}$. So we have

$$M_0 \downarrow_{B_{2m-1}} = \bigoplus_{h=0}^{m} \left(M_0^{h,+} \oplus M_0^{h,-} \right).$$

One checks that $N_0^h \uparrow^{B_{2m-1}} \cong M_0^{h+1,+} \oplus M_0^{h,-}$ and property (b1) follows.

As for property (b2) one can proceed as follows. For any $S \subseteq [2m]$ let $C_S = \sum C_v$, where the sum is over all involutions $v \in B_{2m}$ with $\text{fix}(v) = \text{fix}^-(v) = 0$ and such that $z_i(v) = 0$ if and only if

$i \in S$. Then one can check that the subspace spanned by all C_S with $|S| = 2h$ affords the representation parametrized by the single-rowed diagrams $(2h, 2(n-h))$. From this it is possible to derive the representation $(1^{2h}, 1^{2(m-h)})$.

Let us now turn to the case of the general submodule $M(c)$. For every $k \in [0,n]$, let f_0, f_1, p_0, p_1 be nonnegative integers such that $f_0 + f_1 = k$, $2(p_0 + p_1) = n - k$. By means of Proposition 6.2 we have to show that

$$M(c_{f_0,f_1,p_0,p_1}) \cong \bigoplus_{(\lambda,\mu)\in R(c_{f_0,f_1,p_0,p_1})} \varrho_{\lambda,\mu},$$

where

$$R(c_{f_0,f_1,p_0,p_1}) = \{(\lambda,\mu) \text{ such that } \lambda \vdash f_0 + 2p_0,\ \mu \vdash f_1 + 2p_1, \\ \lambda \text{ has } f_0 \text{ odd columns},\ \mu \text{ has } f_1 \text{ odd columns}\}.$$

Generalizing the ideas developed for M_0, one shows that

$$M(c_{f_0,f_1,p_0,p_1}) \cong \mathrm{Ind}_{B_{n-(f_0+f_1)}\times B_{f_0+f_1}}^{B^n} (M_0 \otimes \varrho_{\iota_{f_0},\iota_{f_1}}),$$

where M_0 is the $B_{n-(f_0+f_1)}$-module constructed as above, and ι_k is the single-rowed Ferrers diagram of length k. This isomorphism can be achieved by standard representation theory, while the rest of the proof can be carried out by applying the partial result obtained on M_0 and a generalization of the Littlewood-Richardson rule to the case of B_n.

Example 6.4 *Let $v \in B_6$ given by $|v| = [6,4,3,2,5,1]$ and $z(v) = [1,0,0,0,1,1]$. Then $f_0 = f_1 = p_0 = p_1 = 1$ and the S_n-conjugacy class c of v has 180 elements. Then the B_6-module $M(c)$ is given by the sum of the irreducible representations indexed by $(\lambda,\mu) \in \mathrm{Fer}(2,6)$ such that both λ and μ are partitions of 3 and have exactly one column of odd length. In particular*

$$M(c) \cong \rho_{\left(\ydiagram{2,1},\,\ydiagram{1,1,1}\right)} \oplus \rho_{\left(\ydiagram{1,1,1},\,\ydiagram{2,1}\right)} \oplus \rho_{\left(\ydiagram{2,1},\,\ydiagram{2,1}\right)} \oplus \rho_{\left(\ydiagram{1,1,1},\,\ydiagram{1,1,1}\right)}.$$

7 Refinement for D_n

We have already seen that in an involutory reflection group $G(r,p,n)$ the submodule generated by the antisymmetric absolute involutions $\mathrm{Asym}(r,p,n)^*$ is isomorphic to the multiplicity-free sum of all the irreducible representations $\rho_{(\lambda,\lambda)^-}$, while all the other irreducible representations of $G(r,p,n)$ are afforded by $\mathrm{Sym}(r,p,n)^*$. We will make use of what was proved for B_n to give a finer decomposition of $\mathrm{Sym}(2,2,n)^*$ for the groups D_n.

Let $\bar{v}$ be a symmetric involution of $D_n^* = B_n/\pm I$ and v and $-v$ be its lifts in B_n. We also denote by $\bar{c}$ the S_n-conjugacy class of $\bar{v}$ in D_n^* and by c and c' the S_n-conjugacy classes of v and $-v$ in B_n. Generalizing the notation used in §6 we let $R(\bar{v})$ be the element of $\mathrm{Fer}(2,1,2,n)$ which is the shape of the tableaux of the image of $\bar{v}$ via the projective Robinson-Schensted correspondence. Namely $R(\bar{v}) \overset{\mathrm{def}}{=} (\lambda,\mu)$, where

$$v \xrightarrow{RS} [P,P], \quad P \in \mathcal{ST}(2,1,2,n), \quad P \text{ of shape } (\lambda,\mu).$$

We also let

$$R(\bar{c}) = \bigcup_{\bar{w}\in\bar{c}} R(\bar{w}).$$

One can verify that the restrictions of the B_n-modules $M(c)$ and $M(c')$ to D_n are isomorphic. If v and $-v$ are not S_n-conjugate then a direct application of Theorem 6.1 provides

$$M(\bar{c}) \cong \bigoplus_{(\lambda,\mu)\in R(\bar{c})} \rho_{\lambda,\mu}.$$

Note that in this case we obtain unsplit representations only since $R(v) = (\lambda, \mu)$ implies $R(-v) = (\mu, \lambda)$. If v and $-v$ are S_n-conjugate, using Theorems 6.1 and 5.1 we can conclude that

$$M(\bar{c}) \cong \bigoplus_{\substack{(\lambda,\mu)\in R(\bar{c}):\\ \lambda\neq\mu}} \rho_{\lambda,\mu} \oplus \bigoplus_{(\lambda,\lambda)\in R(\bar{c})} \rho_{(\lambda,\lambda)^+}.$$

Example 7.1 *Let $v \in B_6$ given by $|v| = [6,4,3,2,5,1]$ and $z(v) = [1,0,0,0,1,1]$. Then $\bar{c}$, the S_n-conjugacy class of $\bar{v}$, has 90 elements and the decomposition of the D_n-module $M(\bar{c})$ is given by all representations indexed by $(\lambda,\mu) \in \mathrm{Fer}(2,1,2,6)$ where both λ and μ are partitions of 3 and have exactly one column of odd length, with the additional condition that if $\lambda = \mu$ the split representation to be considered is $(\lambda,\lambda)^+$. Therefore*

$$M(\bar{c}) \cong \rho_{\left(\ydiagram{2,1},\, \ydiagram{1,1,1}\right)} \oplus \rho_{\left(\ydiagram{2,1},\, \ydiagram{2,1}\right)^+} \oplus \rho_{\left(\ydiagram{1,1,1},\, \ydiagram{1,1,1}\right)^+}.$$

Acknowledgements. The first author wishes to thank Christian Stump and Itaru Terada for useful conversations during the preparation of this work.

References

[1] R. Adin, A. Postnikov and Y. Roichman, Combinatorial Gelfand models, J. Algebra 320 (2008), 1311–1325.

[2] R. Adin, A. Postnikov and Y. Roichman, A Gelfand model for wreath products, Israel J. Math., in press.

[3] J.L. Aguado and J.O. Araujo, A Gelfand model for the symmetric group, Communications in Algebra 29 (2001), 1841–1851.

[4] J.O. Araujo, A Gelfand model for a Weyl group of type B_n, Beiträge Algebra Geom. 44 (2003), 359–373.

[5] J.O. Araujo and J.J Bigeon, A Gelfand model for a Weyl group of type D_n and the branching rules $D_n \hookrightarrow B_n$, J. Algebra 294 (2005), 97–116.

[6] R.W. Baddeley, Models and involution models for wreath products and certain Weyl groups. J. London Math. Soc. 44 (1991), 55–74.

[7] I.N.Bernstein, I.M. Gelfand and S.I. Gelfand, Models of representations of compact Lie groups (Russian), Funkcional. Anal. i Prilozen. 9 (1975), 61–62.

[8] D. Bump and D. Ginzburg, Generalized Frobenius-Schur numbers, J.Algebra 278 (2004), 294–313.

[9] F. Caselli, Projective reflection groups, preprint, `arXiv:0902.0684`.

[10] F. Caselli, Involutory reflection groups and their models, preprint, `arXiv:0905.3649`.

[11] G. Frobenius and I. Schur, Über die reellen Darstellungen de rendlichen Gruppen, S'ber. Akad. Wiss. Berlin (1906), 186–208.

[12] G. Hiss, R. Kessar, *Scopes reduction and Morita equivalence classes of blocks in finite classical groups*, J. Algebra 230 (2000), no. 2, 378-423.

[13] N.F.J. Inglis, R.W. Richardson and J. Saxl, An explicit model for the complex representations of S_n, Arch. Math. (Basel) 54 (1990), 258–259.

[14] N. Kawanaka and H. Matsuyama, A twisted version of the Frobenius-Schur indicator and multiplicity-free permutation representations, Hokkaido Math. J. 19 (1990), 495–506.

[15] A.A. Klyachko, Models for complex representations of groups $GL(n.q)$ and Weyl groups (Russian), Dokl. Akad. Nauk SSSR 261 (1981), 275–278.

[16] A.A. Klyachko, Models for complex representations of groups $GL(n.q)$ (Russian), Mat. Sb. 120 (1983), 371–386.

[17] V. Kodiyalam and D.-N. Verma, A natural representation model for symmetric groups, `arXiv:math/0402216`.

[18] I. Marin, J. Michel, Automorphisms of complex reflection groups, preprint `arXiv:math/0701266`.

[19] G. Pfeiffer, Character Tables of Weyl Groups in GAP, Bayreuth. Math. Schr. 47 (1994), 165–222.

[20] G. C. Shephard and J. A. Todd, Finite unitary reflection groups, Canadian J. Math. 6 (1954), 274–304.

[21] J.R. Stembridge, *On the eigenvalues of representations of reflection groups and wreath products*, Pacific J. Math. 140 (1989) n.2, 353-396

[22] R. P. Stanley, Enumerative combinatorics, vol. 2, Cambridge Studies in Advanced Mathematics 62, Cambridge University Press, Cambridge, 1999.

[23] D.W. Stanton and D.E. White, A Schensted algorithm for rim hook tableaux, J. Combin. Theory Ser. A 40 (1985), 211–247.

Chamber Structure For Double Hurwitz Numbers

Renzo Cavalieri[1] and Paul Johnson[2†] and Hannah Markwig[3‡]

[1]*Colorado State University, Department of Mathematics, Weber Building, Fort Collins, CO 80523-1874, USA*
[2]*Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2AZ, UK*
[3]*CRC "Higher Order Structures in Mathematics", Georg August Universität Göttingen, Bunsenstr. 3-5, 37073 Göttingen, Germany*

Abstract. Double Hurwitz numbers count covers of the sphere by genus g curves with assigned ramification profiles over 0 and ∞, and simple ramification over a fixed branch divisor. Goulden, Jackson and Vakil (2005) have shown double Hurwitz numbers are piecewise polynomial in the orders of ramification, and Shadrin, Shapiro and Vainshtein (2008) have determined the chamber structure and wall crossing formulas for $g = 0$. We provide new proofs of these results, and extend them in several directions. Most importantly we prove wall crossing formulas for all genera.

The main tool is the authors' previous work expressing double Hurwitz number as a sum over labeled graphs. We identify the labels of the graphs with lattice points in the chambers of certain hyperplane arrangements, which give rise to piecewise polynomial functions. Our understanding of the wall crossing for these functions builds on the work of Varchenko (1987). This approach to wall crossing appears novel, and may be of broader interest.

This extended abstract is based on a new preprint by the authors.

Résumé. Les nombres de Hurwitz doubles dénombrent les revêtements de la sphère par une surface de genre g avec ramifications prescrites en 0 et ∞, et dont les autres valeurs critiques sont non dégénérées et fixées. Goulden, Jackson et Vakil (2005) ont prouvé que les nombres de Hurwitz doubles sont polynomiaux par morceaux en l'ordre des ramifications prescrites, et Shadrin, Shapiro et Vainshtein (2008) ont déterminé la structure des chambres et ont établis des formules pour traverser les murs en genre 0. Nous proposons des nouvelles preuves de ces résultats, et les généralisons dans plusieurs directions. En particulier, nous prouvons des formules pour traverser les murs en tout genre.

L'outil principal est le précédent travail des auteurs exprimant les nombres de Hurwitz doubles comme somme de graphes étiquetés. Nous identifions les étiquetages avec les points entiers à l'intérieur d'une chambre d'un arrangement d'hyperplans, qui sont connu pour donner une fonction polynomiale par morceaux. Notre étude des formules pour traverser les murs de cettes fonctions se base sur un travail antérieur de Varchenko (1987). Cette approche paraît nouvelle, et peut être d'un large intérêt.

Ce résumé élargi se base sur un papier nouveaux des auteurs.

Keywords: Hurwitz Numbers, Lattice Points, Hyperplane arrangements, Graphs

†NSF grants DMS-0602191 and DMS-0902754
‡German Research Foundation (Deutsche Forschungsgemeinschaft (DFG)) through the Institutional Strategy of the University of Göttingen

1 Introduction

Hurwitz theory studies holomorphic maps between Riemann surfaces with specified ramification. Double Hurwitz numbers count covers of $\mathbb{P}^1$ with assigned ramification profiles over 0 and ∞, and simple ramification over a fixed branch divisor.

A systematic study of double Hurwitz numbers in Goulden et al. (2005) shows double Hurwitz numbers are piecewise polynomial in the entries of the partitions defining the special ramification. In Shadrin et al. (2008), this result was investigated further in genus 0; the regions of polynomiality are determined, and a recursive wall crossing formula for how the polynomials change is obtained. This paper gives a unified approach to these results that strengthens them in several ways - the most important being the extension of the results of Shadrin et al. (2008) to positive genus.

This extended abstract is based on Cavalieri et al. (2009).

2 Statement of Results

The double Hurwitz number $H_g(\mathbf{x})$ (where $\mathbf{x} = (x_1, \ldots, x_n)$) counts the number of maps $\pi : C \to \mathbb{P}^1$, where C is a connected, genus g curve and π has profiles $\mathbf{x_0} := \{x_i | x_i > 0\}$ (resp. $\mathbf{x}_\infty := \{x_i | x_i < 0\}$) over 0 (resp. ∞), and simple ramification over $r = 2g - 2 + n$ fixed other points. The preimages of 0 and ∞ are marked. Each cover is counted with weight $1/|\mathrm{Aut}(\pi)|$. Since r and g are related by the Riemann-Hurwitz formula, we sometimes use $H^r(\mathbf{x})$ to denote $H_g(\mathbf{x})$ when it makes formulas more attractive.

A ramified cover is essentially equivalent information to a monodromy representation; an equivalent definition of Hurwitz number counts the number of homomorphisms φ from the fundamental group Π_1 of $\mathbb{P}^1 \setminus \{0, \infty, p_1, \ldots, p_r\}$ to the symmetric group S_d such that:

- the image of a loop around 0 has cycle type $\mathbf{x_0}$;
- the image of a loop around ∞ has cycle type $\mathbf{x}_\infty$;
- the image of a loop around p_i is a transposition;
- the subgroup $\varphi(\Pi_1)$ acts transitively on the set $\{1, \ldots, d\}$.

This number is divided by $|S_d|$, to account both for automorphisms and for different monodromy representations corresponding to the same cover. One organizes this count in terms of graphs as in (Cavalieri et al., Lemma 4.1), a fact which is the starting point of our investigation (see Section 3).

Let $\mathcal{H}$ be the hyperplane $\mathcal{H} = \{\sum_i x_i = 0\} \subset \mathbb{R}^n$. We think of H_g (equiv. H^r) as a map

$$H_g : \mathcal{H} \cap \mathbb{Z}^n \to \mathbb{Q} : \mathbf{x} \mapsto H_g(\mathbf{x}).$$

Our first result is a new proof of the following theorem in Goulden et al. (2005):

Theorem 2.1 (GJV) *The function $H_g(\mathbf{x})$ is a piecewise polynomial function of degree $4g - 3 + n$.*

Our techniques allow us to extend this result and answer a question implicit in the work of Goulden, Jackson and Vakil:

Theorem 2.2 *$H_g(\mathbf{x})$ is either even or odd, depending on the parity of the leading degree $4g - 3 + n$.*

We then extend the results of Shadrin et al. (2008) to all genera. We determine the regions on which $H_g(\mathbf{x})$ is polynomial:

Theorem 2.3 *The chambers of polynomiality of $H_g(\mathbf{x})$ are bounded by **walls** corresponding to the **resonance** hyperplanes W_I, given by the equation $W_I = \left\{\mathbf{x}_I = \sum_{i\in I} x_i = 0\right\}$, for any $I \subset \{1,\ldots,n\}$.*

We then describe wall crossing formulas for general genus. Denote the chambers of the resonance arrangement as H-chambers;

Definition 2.4 *Let C_1 and C_2 be two H-chambers adjacent along the wall W_I, with C_1 being the chamber with $x_I < 0$. The Hurwitz number $H^r(\mathbf{x})$ is given by polynomials, say $P_1(\mathbf{x})$ and $P_2(\mathbf{x})$, on these two regions. By a wall crossing formula, we mean a formula for the polynomial*

$$WC_I^r(\mathbf{x}) = P_2(\mathbf{x}) - P_1(\mathbf{x}).$$

With the notation $WC_I^r(\mathbf{x})$ there is no ambiguity about which direction we cross the wall. Since $\mathbf{x}$ lies on the hyperplane $\sum_{i=1}^n x_i = 0$, each wall has two possible labels: W_I and W_{I^c}. We choose the name so that $\mathbf{x}_I$ is increasing.

We use $H^{r\bullet}(\mathbf{x})$ to denote Hurwitz numbers with potentially disconnected covers. Our main theorem is:

Theorem 2.5 (Wall crossing formula)

$$WC_I^r(\mathbf{x}) = \sum_{s+t+u=r} \sum_{|\mathbf{y}|=|\mathbf{z}|=|\mathbf{x}_I|} (-1)^t \binom{r}{s,t,u} \frac{\prod \mathbf{y}_i}{\ell(\mathbf{y})!} \frac{\prod \mathbf{z}_j}{\ell(\mathbf{z})!} H^s(\mathbf{x}_I, \mathbf{y}) H^{t\bullet}(-\mathbf{y}, \mathbf{z}) H^u(\mathbf{x}_{I^c}, -\mathbf{z}) \tag{1}$$

Here $\mathbf{y}$ is an ordered tuple of $\ell(\mathbf{y})$ positive integers with sum $|\mathbf{y}|$, and similarly with $\mathbf{z}$.

The walls W_I correspond to values of $\mathbf{x}$ where the cover could potentially be disconnected, or where $x_i = 0$. Crossing this second type of wall corresponds to moving a ramification between 0 and ∞. In the traditional view of double Hurwitz numbers, the number of ramification points over 0 and ∞ were fixed separately, rather than just the total number of ramification points. Theorem 2.5 suggests that it is natural to treat them as part of the same problem: the wall crossing formula for $x_i = 0$ is identical to the other wall crossing formulas.

3 Overview of Methods

This paper is an exploration of the consequences of a formula in the author's previous work, Cavalieri et al., which expresses double Hurwitz numbers $H_g(\mathbf{x})$ as a sum over certain directed trivalent graphs Γ with several labelings, which we call monodromy graphs:

Definition 3.1 *For fixed g and $\mathbf{x} = (x_1,\ldots,x_n)$, a graph Γ is a* monodromy graph *if:*

- *Γ is a connected, genus g, directed graph.*
- *Γ has n 1-valent vertices called* leaves*; the edges leading to them are* ends*. All ends are directed inward, and are labeled by the weights $x_1,\ldots,x_n$. If $x_i > 0$, we say it is an* in-end*, otherwise it is an* out-end.

- *All other vertices of* Γ *are* 3*-valent, and are called* internal vertices. *Edges that are not ends are called* internal edges.
- *After reversing the orientation of the out-ends,* Γ *does not have directed loops, sinks or sources.* (i).
- *The internal vertices are ordered compatibly with the partial ordering induced by the directions of the edges.*
- *Every internal edge* e *of the graph is equipped with a* weight $w(e) \in \mathbb{N}$. *The weights satisfy the* balancing condition *at each internal vertex: the sum of all weights of incoming edges equals the sum of the weights of all outgoing edges.*

It follows from (Cavalieri et al., Lemma 4.1) that the Hurwitz number is computed as:

$$H_g(\mathbf{x}) = \sum_{\Gamma} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{e} w(e), \tag{2}$$

the sum over all monodromy graphs Γ for g and $\mathbf{x}$, and the product over the interior edges of Γ.

In genus zero, the edge labelings $w(e)$ are determined uniquely by $\mathbf{x}$. This makes the genus zero case much easier to treat, and the results for this case were already presented in Cavalieri et al.. In higher genus, if we fix a directed graph and the labels $\mathbf{x}$ for the ends (such data will be called a *directed* $\mathbf{x}$*-graph*), there are many ways to assign edge labels $w(e)$ that satisfy the balancing condition.

The crux of this paper is to understand the space of edge labelings (which we call *flows*) for each directed $\mathbf{x}$-graph. The space of flows consists of the lattice points in a certain bounded polytope which we call an F-chamber. The contribution $s_{\Gamma(d)}(\mathbf{x})$ of a fixed directed $\mathbf{x}$-graph $\Gamma(d)$ to $H_g(\mathbf{x})$ equals

$$s_{\Gamma(d)}(\mathbf{x}) = \frac{1}{|\mathrm{Aut}(\Gamma(d))|} \cdot m(\mathcal{C}) \cdot \sum_{b \in \mathcal{C} \cap \Lambda} \prod_{e} L_e(\mathbf{x}, b) \tag{3}$$

where $\mathcal{C}$ is the F-chamber associated to $\Gamma(d)$, Λ denotes the lattice and $m(\mathcal{C})$ equals the number of ways to order the vertices of $\Gamma(d)$ as required for a monodromy graph. Here we have written $L_e(\mathbf{x}, b)$ for $w(e)$, as the weight of each edge will be a linear function in $\mathbf{x}$ and the coordinates of Λ.

We illustrate this in an example that we continue to develop throughout. Consider the directed $\mathbf{x}$-graph $\Gamma(\mathbf{x}, d, v)$ on the left hand side in Figure 1. In this example, we use the notation $\Gamma(\mathbf{x}, d, v)$ to indicate that the graph comes with directed edges (d) and with a vertex ordering (v). In the figure, the vertices are labelled to indicate the vertex ordering. We want to understand all monodromy graphs that equal $\Gamma(\mathbf{x}, d, v)$ after forgetting the weights of the internal edges. There are no monodromy graphs that equal $\Gamma(\mathbf{x}, d, v)$ after forgetting the weights if $x_1 + x_3 \leq 0$, so we assume that $x_1 + x_3 > 0$.

We have two degrees of freedom to choose weights for the interior edges such that the balancing condition is satisfied, one for each independent cycle of Γ. Once we label one of the interior edges with the weight i, and another with j, all other weights are determined by the balancing condition, as shown in the right hand side of Figure 1. All possible collections of edge labels are indexed by the lattice points in the polytope defined requiring these labels to be nonnegative:

$$i \geq 0, j \geq 0, j + i - x_2 \geq 0, -x_4 - i - j \geq 0, -x_4 - j \geq 0, j - x_2 \geq 0.$$

(i) We do not consider leaves to be sinks or sources.

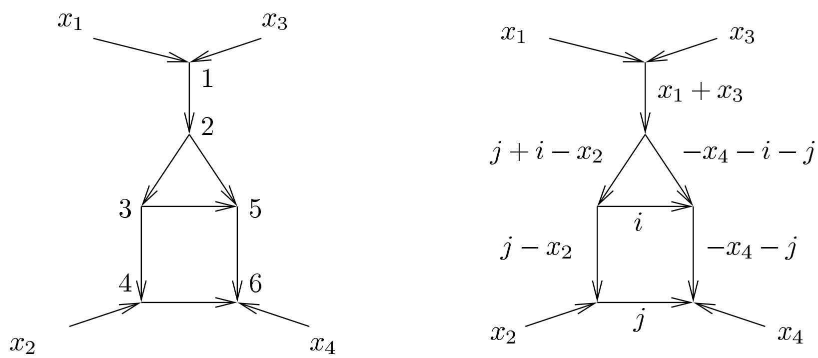


Fig. 1: A directed $\mathbf{x}$-graph and the weights of internal edges determined by the balancing condition

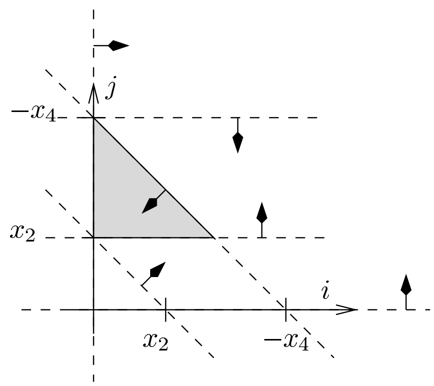


Fig. 2: The F-chamber corresponding to $\Gamma(\mathbf{x}, d, v)$

Figure 2 shows all hyperplanes $w(e) = 0$ with a normal vector indicating on which side of the hyperplane the inequality $w(e) > 0$ is satisfied; this defines the F-chamber corresponding to $\Gamma(\mathbf{x}, d, v)$.

The contribution of $\Gamma(\mathbf{x}, d, v)$ to $H_g(\mathbf{x})$ is given by

$$(x_1 + x_3) \cdot \sum_{i=0}^{-x_4-x_2} \sum_{j=x_2}^{-i-x_4} i \cdot j \cdot (j + i - x_2) \cdot (-x_4 - i - j) \cdot (-x_4 - j) \cdot (j - x_2)$$

where the sum goes over all lattice points (i, j) in the polygon above ($\Gamma(\mathbf{x}, d, v)$ has no automorphisms).

Theorem 2.1 follows from Equation 3 and the general theory of lattice points in polytopes. As we change $\mathbf{x}$, the facets of F-chamber $\mathcal{C}$ translate (their normal directions remain constant). Since for all integral $\mathbf{x}$, the vertices of the F-chamber $\mathcal{C}$ are integers, our sums are piecewise polynomial, and the walls occur when the topology of $\mathcal{C}$ changes.

In the general setup of the theory the resulting polynomials need not be odd or even, so Theorem 2.2 is more subtle: it is related to Ehrhart reciprocity, and depends essentially on the fact that the polynomial

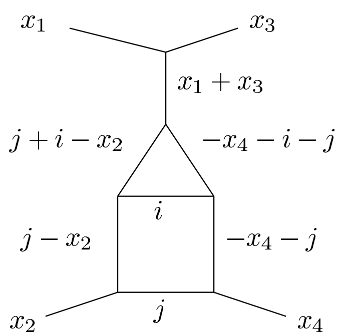


Fig. 3: Labels $w'(e)$ for the undirected graph

we are summing over the polytope vanishes on the boundary of the polytope.

To prove Theorems 2.3 and 2.5 requires understanding for what values of $\mathbf{x}$ the F-chambers change topology, and how they change topology, respectively. To answer these questions, it is helpful to notice that the F-chambers for distinct $\mathbf{x}$-graphs with the same underlying undirected graph Γ fit together as the set $\mathcal{BC}_\Gamma(x)$ of bounded chambers of a natural hyperplane arrangement $\mathcal{A}_\Gamma(\mathbf{x})$ associated to Γ and $\mathbf{x}$.

Returning to our example, we can retain the orientation of the edges in Figure 1 as a reference orientation, and the labels $w'(e)$ for the internal edges obtained from the balancing condition as in Figure 3.

We switch to $w'(e)$ instead of $w(e)$ because we do no longer restrict the edge labels to be positive; instead, any possible value of i and j are allowed. For each edge of Γ, the set of i and j where $w'(e) = 0$ will give a hyperplane, and together these form the hyperplane arrangement $\mathcal{A}_\Gamma(\mathbf{x})$. Inside a chamber of $\mathcal{A}_\Gamma(\mathbf{x})$, a sign for $w'(e)$ is picked for every edge, and thus an orientation for every edge; the chamber will be the F-chamber for that directed $\mathbf{x}$-graph. Figure 4 shows the hyperplane arrangement $\mathcal{A}_\Gamma(\mathbf{x})$, with each F-chamber labeled by the corresponding directed graphs with the induced orientations. Since the orientation of the ends and the edge with label $x_1 + x_3$ does not depend on i and j, we do not include these edges in the pictures.

Only the bounded F-chambers (shaded) correspond to directed $\mathbf{x}$-graphs that contribute to the Hurwitz number. The unbounded F-chambers correspond to graphs with a directed loop, and so the vertices have no compatible total orderings and the multiplicity of these chambers are zero.

For different chambers, the product $\prod w(e)$ differs at most by the sign, since the edge weights $w(e)$ equal plus or minus the edge label $w'(e)$, depending on the side of the hyperplane $w'(e) = 0$ the F-chamber is situated. Thus we can define a sign $\operatorname{sign}(\mathcal{C})$ for each F-chamber $\mathcal{C}$ that is determined by the number of edges that are reversed when compared to the reference orientation.

Summing all the contributions from directed $\mathbf{x}$-graphs $\Gamma(d)$ with the same underlying undirected $\mathbf{x}$-graph Γ, we get the contribution S_Γ of the undirected $\mathbf{x}$-graph Γ to the Hurwitz number as

$$S_\Gamma(\mathbf{x}) = \frac{1}{|\mathrm{Aut}(\Gamma)|} \sum_{\mathcal{C}} \operatorname{sign}(\mathcal{C}) m(\mathcal{C}) \sum_{b \in \mathcal{C} \cap \Lambda} \prod L_e(\mathbf{x}, b),$$

where the sum goes over all bounded chambers $\mathcal{C}$ of $\mathcal{A}_\Gamma(\mathbf{x})$.

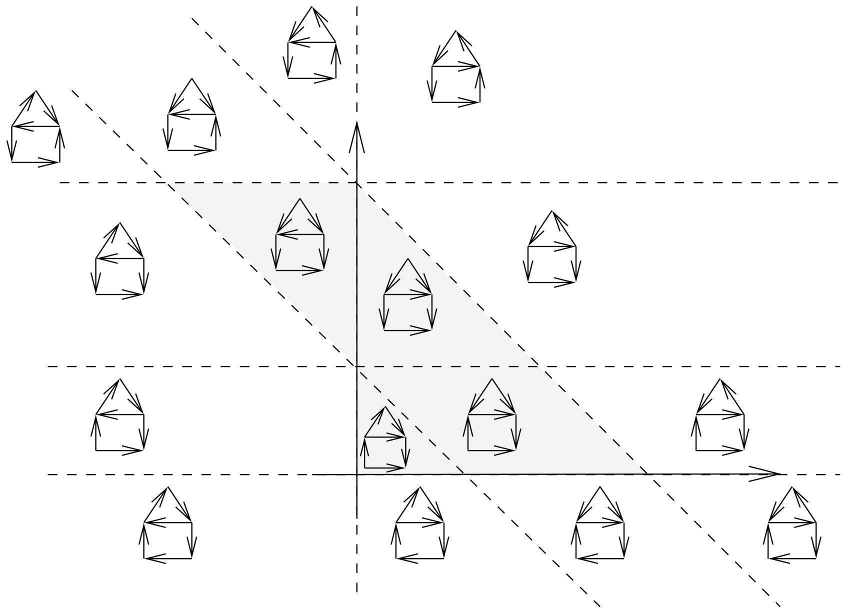


Fig. 4: The parameter space for monodromy graphs corresponding to a given $\mathbf{x}$-graph

Remark 3.2 *In Equation (2)* Γ *is a monodromy graph, while here* Γ *is an* $\mathbf{x}$*-graph, and so the meaning of* *Aut*(Γ) *is different. An automorphism of a monodromy graph must fix all vertices, while an automorphism of an* $\mathbf{x}$*-graph only needs to fix the ends. These extra automorphisms account for the fact that the same monodromy graph can occur in multiple ways from a single* $\mathbf{x}$*-graph.*

Even for a generic choice of $\mathbf{x}$ the arrangement $\mathcal{A}_\Gamma(\mathbf{x})$ is not simple - that is, there are hyperplanes that do not intersect transversally. This follows from the balancing condition: if two edge labels incident to a vertex are both zero, then the third edge label must be as well. As a consequence, for each vertex we have three hyperplanes intersecting in codimension two. But for generic $\mathbf{x}$, these are the only nontransverse intersections. When we pass through a value of $\mathbf{x}$ with more nontransverse intersections than expected, the topology of the arrangement $\mathcal{A}_\Gamma(\mathbf{x})$ changes, and so do the Hurwitz polynomials. We prove Theorem 2.3 by showing that if $e_1, \ldots, e_k$ are k edges whose hyperplanes intersect in codimension $k-1$ at $\mathbf{x}$, but generically intersect transversally, then these edges disconnect Γ, and each component will contain at least one end. Flows in the intersection of the hyperplanes correspond to flows on the graph where the edges are cut, and so if I is the set of ends on one component, we see that $\mathbf{x}$ must have been a point on the wall W_I.

Our main result is the wall crossing formula (Theorem 2.5). The idea of the proof is simple: matching the contributions to both sides by every directed $\mathbf{x}$-graph. In genus 0, realizing this strategy is straightforward because there is a natural geometric bijection (Cut) between graphs contributing to the wall crossing (LHS) and pairs of graphs contributing to the product of Hurwitz numbers on the RHS (the middle term can easily be seen to equal 1 in genus 0). In higher genus Cut is no longer a function, and a delicate process of inclusion/exclusion is required, leading us to foray into algebraic combinatorics. While to determine the walls it is enough to know where topology of $\mathcal{A}_\Gamma(\mathbf{x})$ changes, to derive the wall crossing formula we must understand how the topology changes; i.e. if $\mathbf{x}_1$ and $\mathbf{x}_2$ lie in two H-chambers C_1 and C_2, adjacent along the wall W_I, how does $\mathcal{A}_\Gamma(\mathbf{x}_1)$ differ from $\mathcal{A}_\Gamma(\mathbf{x}_2)$? This understanding is essential for relating the difference of the contributions $S_\Gamma(\mathbf{x}_1) - S_\Gamma(\mathbf{x}_2)$ to the resp. Hurwitz numbers.

The information how $\mathcal{A}_\Gamma(\mathbf{x}_1)$ differs from $\mathcal{A}_\Gamma(\mathbf{x}_2)$ is conveniently encoded in a linear map

$$\nabla_{12} : \mathbb{R}[\mathcal{BC}_\Gamma(\mathbf{x}_1)] \to \mathbb{R}[\mathcal{BC}_\Gamma(\mathbf{x}_2)]$$

called the Gauss-Manin connection. The basic picture is that as $\mathbf{x}$ passes through a wall, certain F-chambers vanish, and others appear. For any F-chamber, the change in shape as $\mathbf{x}$ crosses a wall can be described in terms of adding or subtracting these appearing F-chambers; ∇_{12} is the map that sends a given F-chamber to this signed sum of F-chambers. It turns out to be easier to declare F-chambers to form an orthonormal basis of $\mathbb{R}[\mathcal{BC}_\Gamma(\mathbf{x}_1)]$, and study the adjoint ∇_{12}^* which records which F-chambers of $\mathcal{A}_\Gamma(\mathbf{x}_1)$ map to a given one in $\mathcal{A}_\Gamma(\mathbf{x}_2)$.

The key point is that integrating a polynomial f over an F-chamber $\mathcal{C}(\mathbf{x})$ gives only a piecewise polynomial function; but if we replace $\mathcal{C}(\mathbf{x})$ by $\nabla_{12}\mathcal{C}(\mathbf{x})$ when we cross a wall, then we get a globally defined polynomial. Results of Varchenko (1987) show that if we replace integration by summing over lattice points, the same result is true if we deal properly with lattice points in the boundary of the polytope. Since our polynomials vanish there, we don't have to worry about this, and so ∇_{12} encodes essentially all the information for Hurwitz wall crossing.

Returning once more to our running example, in Figure 4 showing $\mathcal{A}_\Gamma(\mathbf{x})$, we implicitly assumed that $0 > x_2 + x_4$. The topology of the hyperplane arrangement changes if $0 = x_2 + x_4$. Fix the wall $W_{\{2,4\}}$ and let C_1 and C_2 be two adjacent H-chambers. Assume that in C_1, we have $0 < x_2 + x_4$, and in C_2, we have $x_2 + x_4 < 0$. Figure 5 shows the hyperplane arrangements $\mathcal{A}_\Gamma(\mathbf{x}_1)$ and $\mathcal{A}_\Gamma(\mathbf{x}_2)$ for two points $\mathbf{x}_1 \in C_1$ and $\mathbf{x}_2 \in C_2$. The hyperplanes appear with their defining equations. They are drawn with different line styles in order to emphasize how they move. The bounded F-chambers are labelled with letters. Since the edge with weight $x_1 + x_3$ gives the inequality $x_1 + x_3 > 0$ on the right which is not satisfied on the left, every F-chamber on the right is an appearing chamber, and every F-chamber on the left is vanishing. This can also be seen from the corresponding graphs: since the top most interior edge with weight $x_1 + x_3$ always points down on the right, there is a flow from top to bottom. Figure 6 shows the directed $\mathbf{x}$-graphs corresponding to some of the F-chambers.

To understand the Gauss-Manin connection for this example, we pick an appearing F-chamber on the right, e.g. A, and ask ourselves what F-chambers on the left contain it in their support when carried over the wall, i.e. we determine $\nabla_{\Gamma,12}^*(A)$. To do this, we take chambers on the left, e.g. E, and carry them over, i.e. we first determine $\nabla_{\Gamma,12}(E)$. When we carry E over, we get B and keep the orientation (we switch the summation index twice). In the same way, we get $\nabla_{\Gamma,12}(F) = A$. If we interpret the inequalities of G on the right, we have to switch one summation index, and then we end up with $A + B + C$. Thus $\nabla_{\Gamma,12}(G) = -A - B - C$. Finally, H becomes $D + B + C$. Thus, $\nabla_{\Gamma,12}^*(A) = F - G$, $\nabla_{\Gamma,12}^*(B) = E - G + H$, $\nabla_{\Gamma,12}^*(C) = -G + H$ and $\nabla_{\Gamma,12}^*(D) = H$.

Our key result is that we can express the linear map ∇_{12}^* combinatorially. We define a simple I-cut of a directed $\mathbf{x}$-graph Γ to be a minimal set of edges E so that E disconnects Γ into exactly two components, one containing the ends in I, and the other containing all the ends in I^c. An I-cut in general is a union of simple I-cuts. This could be the empty union, which we call the empty cut. Note that a general I-cut might be expressible as a union of simple I-cuts in many different ways.

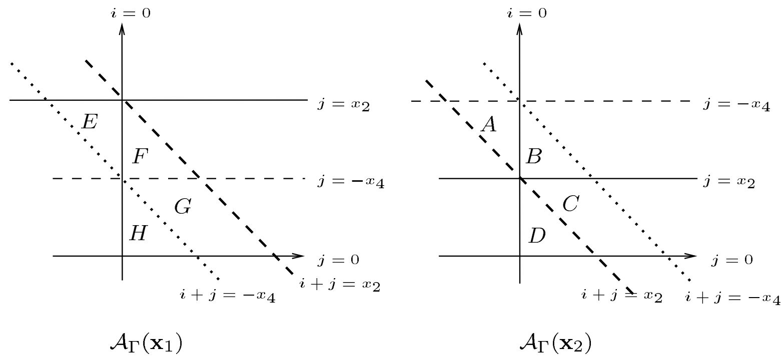


Fig. 5: The hyperplane arrangements $\mathcal{A}_\Gamma(\mathbf{x}_1)$ and $\mathcal{A}_\Gamma(\mathbf{x}_2)$ for two points $\mathbf{x}_1$ and $\mathbf{x}_2$ on opposite sides of a wall.

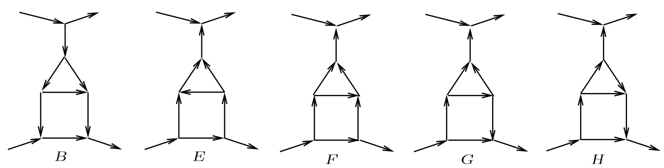


Fig. 6: The directed $\mathbf{x}$-graphs corresponding to the F-chambers B, E, F, G and H of figure 5.

The set of simple I-cuts forms a poset under inclusion, which we denote $C_\Gamma(I)$. As an example, we show the poset of $\{1,3\}$-cuts of the graph Γ_B from above as an example (see Figure 7).

In this case, the $C_\Gamma(I)$ is simply the boolean lattice generated by the simple cuts, although this is not true in general. A key lemma is in our paper is the following:

Lemma 3.3 *$C_\Gamma(I)$ is isomorphic to the face lattice of a certain cone, defined in terms of a different hyperplane arrangement associated to Γ.*

The main importance of Lemma 3.3 is that it shows $C_\Gamma(I)$ is Eulerian, and all Möbius inversion type questions can be translated into questions about Euler characteristics of subsets of the cone. A more immediate consequence is that $C_\Gamma(I)$ is ranked; the empty cut has rank zero.

The key step in our proof of the Main Theorem 2.5 is the following theorem, which expresses ∇^*_{12} in terms of the poset of cuts $C_\Gamma(I)$.

Theorem 3.4 *Let A and B be F-chambers in $\mathcal{BC}_\Gamma(\mathbf{x}_2)$ and $\mathcal{BC}_\Gamma(\mathbf{x}_1)$, respectively. Let Γ_A and Γ_B denote the corresponding orientations of the edges of Γ, and let S be the subset of edges of Γ where these orientations differ. Then*

$$\langle \nabla^*_{12} A, B \rangle = (-1)^{|S|} \sum_{S \subset C \in C_{\Gamma_A}(I)} (-1)^{rk(C)}.$$

Here the notation means we sum only over the I-cuts of Γ_A that contain S.

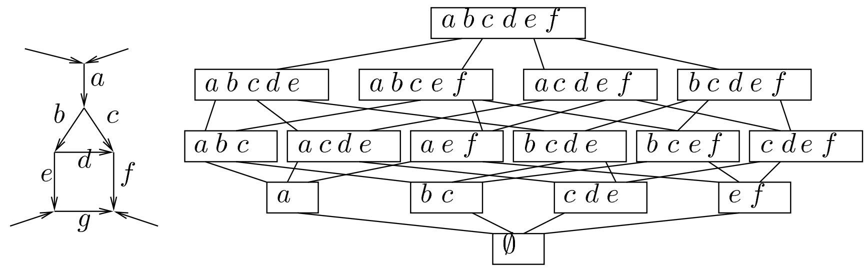


Fig. 7: The directed $\mathbf{x}$-graph Γ_B and its poset of $\{1,3\}$-cuts

Theorem 3.4 does not depend on the graph being trivalent, and could be of independent interest.

The proof of Theorem 3.4 is the technical heart of the paper, and rests upon the following observations from Varchenko (1987): cones are preserved by the Gauss-Manin connection, and every chamber can be written as a signed sum of cones. Thus, it suffices to show ∇_{12}^* preserves cones. We are able to do this by using the understanding of $C_\Gamma(I)$ afforded by Lemma 3.3.

We now illustrate the statement of Theorem 3.4 in the case of our example. Consider the appearing chamber B on the right of the wall. We have seen that $\nabla_{\Gamma,12}^*(B) = E - G + H$, and so we understand the left hand side of Theorem 3.4. We have also determined the poset of I-cuts of the directed graph Γ_B (Figure 7), and so we are able to compute the right hand side as well.

First, let us verify that Theorem 3.4 gives $\langle \nabla_{12}^* B, E\rangle = 1$. We see that to get Γ_E from Γ_B, we must change the orientation of the edges a, b, c, d, e and f, and so $S = \{a,b,c,d,e,f\}$. . There is only one cut that contains S, the maximal cut. Its rank is four (see Figure 7). Since $|S| = 6$, Theorem 3.4 gives $\langle \nabla_{12}^* B, E\rangle = (-1)^6 \cdot (-1)^4 = 1$.

Similarly, we will verify that Theorem 3.4 gives $\langle \nabla_{12}^* B, G\rangle = -1$. In this case, the set S of edges where the orientations of Γ_G and Γ_B differ is $\{a,b,c,e\}$. There are three cuts that cut these edges, namely $abcde$ and $abcef$, both of rank three, and $abcdef$ of rank four. Since $|S| = 4$, Theorem 3.4 gives $\langle \nabla_{12}^* B, G\rangle = (-1)^4 \cdot ((-1)^3 + (-1)^3 + (-1)^4) = -1$.

Additionally, chambers that do not appear in $\nabla_{12}^* B$ should appear with coefficient zero in the right hand side of Theorem 3.4. Let us check that we get $\langle \nabla_{12}^* B, F\rangle = 0$. To get Γ_F from Γ_B, the set S of edges we must reverse is $\{a,b,c,e,f\}$. The cuts that contain S are $abcef$ of rank three and $abcdef$ of rank four, and so we get $(-1)^5 \cdot ((-1)^3 + (-1)^4) = 0$.

A more complicated wall crossing formula than Theorem 2.5 follows rather quickly from Theorem 3.4.

For a cut $C \in C_\Gamma(I)$, removing the edges in C from Γ will cut Γ into multiple components graphs, each of which can can be interpreted as a graph appearing for a simpler Hurwitz number.

As we sum over all Γ, we will sometimes see essentially the same cut C appearing for different Γ - that is, the components of $\Gamma \setminus C$ will be different graphs, but will contribute to the same Hurwitz problem (have the same number of vertices and in and out going ends), and glue together in the same manner. This is the situation illustrated in Figure 8. As a result, we obtain:

Theorem 3.5 (Heavy Formula)

$$WC_I^r(\mathbf{x}) = \sum_{N=0}^{\infty} \sum_{s+(\sum_{j=1}^N t_j)+u=r} \sum_{|\lambda|=|\eta|=d} \sum_{data\ in\ \star} (-1)^N \binom{r}{s, t_1, \ldots, t_N, u} \frac{\prod (\mu^{(i,j)})_k}{\prod \ell(\mu^{(i,j)} j)!}$$

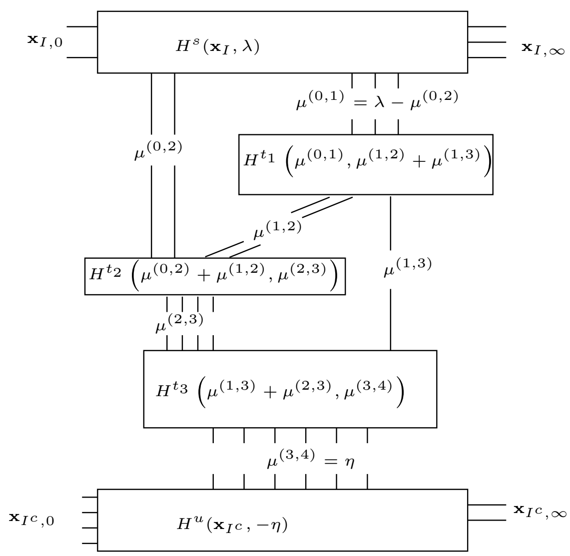


Fig. 8: The data denoted by $\star$ in the heavy formula, Theorem 3.5

$$H^s(\mathbf{x}_I, \lambda)\left(\prod_{j=1}^{N} H^{t_j}(\star)\right) H^u(\mathbf{x}_{I^c}, -\eta)$$

The data denoted by $\star$ is illustrated in Figure 8: it consists in disconnecting a graph with an I-cut in all possible ways with the right numerical invariants. The μ_i^j denote the partitions of weights of the edges connecting the i-th to the j-th connected component, we use $(\mu_i^j)_k$ to denote its parts.

The derivation of Theorem 3.5 from Theorem 2.5 is essentially inclusion-exclusion, and an application of Lemma 3.3.

4 Motivation and Connections to other work

Although our methods are essentially combinatorial, much of the motivation of Goulden et al. (2005), and hence our work, comes from algebraic geometry, in particular the ELSV formula Ekedahl et al. (2001). There, it is shown that similar polynomiality occurs for single Hurwitz numbers (where there is no ramification over ∞), and that the coefficients of these polynomials are the intersection of certain classes in the moduli space of curves $\overline{\mathcal{M}}_{g,n}$. This connection has been vital in understanding these intersections. In Goulden et al. (2005), it is suggested that a similar relationship should hold for one part double Hurwitz, where the map is totally ramified over zero - i.e., where $\mathbf{x}$ has only one positive part. They conjecture that the moduli space of curves should be replaced by some yet to be determined universal Picard space, which would give us a similar understanding of the intersection theory there. One part double Hurwitz numbers are simply one chamber of the Hurwitz problem, however, and it would be wonderful to extend the conjecture of Goulden et al. (2005) to give a formula for double Hurwitz numbers on all chambers, with the wall crossing phenomenon explained in terms of changes in the moduli space. Our work could perhaps be of use in investigating such a conjecture.

Acknowledgements

We would like to thank Federico Ardila and Michael Shapiro for helpful discussions.

References

R. Cavalieri, P. Johnson, and H. Markwig. Tropical Hurwitz numbers. JACO (to appear), arXiv:0804.0579.

R. Cavalieri, P. Johnson, and H. Markwig. Wall crossings for double Hurwitz numbers. Preprint, arXiv:1003.1805, 2009.

T. Ekedahl, S. Lando, M. Shapiro, and A. Vainshtein. Hurwitz numbers and intersections on moduli spaces of curves. *Invent. Math.*, 146(2):297–327, 2001.

I. Goulden, D. Jackson, and R. Vakil. Towards the geometry of double Hurwitz numbers. *Adv. Math.*, 198 (1):43–92, 2005.

S. Shadrin, M. Shapiro, and A. Vainshtein. Chamber behavior of double Hurwitz numbers in genus 0. *Adv. Math.*, 217(1):79–96, 2008.

A. Varchenko. Combinatorics and topology of the arrangement of affine hyperplanes in the real space. *Functional Anal. Appl.*, 21(1):9–19, 1987.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 119–126

Generalized Ehrhart polynomials

Sheng Chen[1] and Nan Li[2] and Steven V Sam[2†]

[1] *Department of Mathematics, Harbin Institute of Technology, Harbin, China 150001*
[2] *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

Abstract. Let P be a polytope with rational vertices. A classical theorem of Ehrhart states that the number of lattice points in the dilations $P(n) = nP$ is a quasi-polynomial in n. We generalize this theorem by allowing the vertices of $P(n)$ to be arbitrary rational functions in n. In this case we prove that the number of lattice points in $P(n)$ is a quasi-polynomial for n sufficiently large. Our work was motivated by a conjecture of Ehrhart on the number of solutions to parametrized linear Diophantine equations whose coefficients are polynomials in n, and we explain how these two problems are related.

Résumé. Soit P un polytope avec sommets rationelles. Un théorème classique des Ehrhart déclare que le nombre de points du réseau dans les dilatations $P(n) = nP$ est un quasi-polynôme en n. Nous généralisons ce théorème en permettant à des sommets de $P(n)$ comme arbitraire fonctions rationnelles en n. Dans ce cas, nous prouvons que le nombre de points du réseau en $P(n)$ est une quasi-polynôme pour n assez grand. Notre travail a été motivée par une conjecture d'Ehrhart sur le nombre de solutions à linéaire paramétrée Diophantine équations dont les coefficients sont des polyômes en n, et nous expliquer comment ces deux problèmes sont liés.

Keywords: Diophantine equations, Ehrhart polynomials, lattice points, quasi-polynomials

1 Introduction.

In this article, we relate two problems, one from classical number theory, and one from lattice point enumeration in convex bodies. Motivated by a conjecture of Ehrhart and a result of Xu, we study linear systems of Diophantine equations with a single parameter. To be more precise, we suppose that the coefficients of our system are given by polynomial functions in a variable n, and also that the number of solutions $f(n)$ in positive integers for any given value of n is finite. We are interested in the behavior of the function $f(n)$, and in particular, we prove that $f(n)$ is **eventually a quasi polynomial**, i.e., there exists some period s and polynomials $f_i(t)$ for $i = 0, \dots, s-1$ such that for $t \gg 0$, the number of solutions for $n \equiv i \pmod{s}$ is given by $f_i(n)$. The other side of our problem can be stated in a similar fashion: suppose that $P(n)$ is a convex polytope whose vertices are given by rational functions in n. Then the number of integer points inside of $P(n)$, as a function of n, enjoys the same properties as that of f as above. We now describe in more detail some examples and the statements of our results.

† Supported by an NSF graduate fellowship and an NDSEG fellowship.

1.1 Diophantine equations.

As a warmup to our result, we begin with two examples. The first is a result of Popoviciu. Let a and b be relatively prime positive integers. We wish to find a formula for the number of positive integer solutions (x, y) to the equation $ax + by = n$. For a real number x, let $\lfloor x \rfloor$ denote the greatest integer less than or equal to x, and define $\{x\} = x - \lfloor x \rfloor$ to be the fractional part of x. Then the number of such solutions is given by the formula

$$\frac{n}{ab} - \left\{ \frac{na^{-1}}{b} \right\} - \left\{ \frac{nb^{-1}}{a} \right\} + 1,$$

where a^{-1} and b^{-1} satisfy $aa^{-1} \equiv 1 \pmod b$ and $bb^{-1} \equiv 1 \pmod a$. See [BR, Chapter 1] for a proof. In particular, this function is a quasi-polynomial in n.

For the second example which is a generalization of the first example, consider the number of solutions $(x, y, z) \in \mathbf{Z}^3_{\geq 0}$ to the matrix equation

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} \tag{1}$$

where the x_i and y_i are fixed positive integers and $x_{i+1}y_i < x_i y_{i+1}$ for $i = 1, 2$. Write $Y_{ij} = x_i y_j - x_j y_i$. We assume that $\gcd(Y_{12}, Y_{13}, Y_{23}) = 1$, so that there exist integers (not unique) f_{ij}, g_{ij} such that

$$\gcd(f_{12}Y_{13} + g_{12}Y_{23}, Y_{12}) = 1, \quad \gcd(f_{13}Y_{12} + g_{13}Y_{23}, Y_{13}) = 1, \quad \gcd(f_{23}Y_{13} + g_{23}Y_{12}, Y_{23}) = 1.$$

Now define two regions $\Omega_i = \{(x, y) \mid \frac{y_i}{x_i} < \frac{y}{x} < \frac{y_{i+1}}{x_{i+1}}\}$ for $i = 1, 2$. Then if $m = (m_1, m_2) \in \mathbf{Z}^2$ is in the positive span of the columns of the matrix in (1), there exist Popoviciu-like formulas for the number of solutions of (1) which depend only on whether $m \in \Omega_1$ or $m \in \Omega_2$, and the numbers $Y_{ij}, f_{ij}, g_{ij}, x_i, y_i$. See [Xu, Theorem 4.3] for the precise statement.

In particular, one can replace the x_i, y_i, and m_i by polynomials in n in such a way that for all values of n, the condition $\gcd(Y_{12}, Y_{13}, Y_{23}) = 1$ holds. For a concrete example, consider the system

$$\begin{pmatrix} 2n+1 & 3n+1 & n^2 \\ 2 & 3 & n+1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3n^3+1 \\ 3n^2+n-1 \end{pmatrix}.$$

Then for $n \gg 0$, we have that

$$\frac{3}{3n+1} < \frac{3n^2+n-1}{3n^3+1} < \frac{n+1}{n^2},$$

so that for these values of n, there exists a quasi-polynomial that counts the number of solutions (x, y, z).

Given these examples, we are ready to state our general theorem. We denote by $\mathbf{QP}_{\gg 0}$ the set of functions $f \colon \mathbf{Z} \to \mathbf{Z}$ which are eventually quasi-polynomial.

Theorem 1.1 *Let $A(n)$ be an $m \times k$ matrix, and $b(n)$ be a column vector of length m, both with entries in $\mathbf{Z}[n]$. If $f(n)$ denotes the number of nonnegative integer vectors x satisfying $A(n)x = b(n)$ (assuming that these values are finite), then $f \in \mathbf{QP}_{\gg 0}$.*

This theorem generalizes the conjecture [Sta, Exercise 4.12]. See [Ehr, p. 139] for a conjectural multivariable analogue.

1.2 Lattice point enumeration.

We first recall a classical theorem due to Pick. Let $P \subset \mathbf{R}^2$ be a convex polygon with integral vertices. If $A(P)$, $I(P)$, and $B(P)$ denote the area of P, the number of integer points in the interior of P, and the number of integer points on the boundary of P, respectively, then one has the equation

$$A(P) = I(P) + \frac{1}{2}B(P) - 1.$$

Now let us examine what happens with dilates of P: define $nP = \{nx \mid x \in P\}$. Then of course $A(nP) = A(P)n^2$ and $B(nP) = nB(P)$ whenever n is a positive integer, so we can write

$$A(P)n^2 = I(nP) + \frac{1}{2}B(P)n - 1,$$

or equivalently,

$$\#(nP \cap \mathbf{Z}^2) = I(nP) + B(nP) = A(P)n^2 + \frac{1}{2}B(P)n + 1,$$

which is a polynomial in n. The following theorem of Ehrhart says that this is always the case independent of the dimension, and we can even relax the integral vertex condition to rational vertices:

Theorem 1.2 (Ehrhart) *Let $P \subset \mathbf{R}^d$ be a polytope with rational vertices. Then the function $L_P(n) = \#(nP \cap \mathbf{Z}^d)$ is a quasi-polynomial of degree $\dim P$. Furthermore, if D is an integer such that DP has integral vertices, then D is a period of $L_P(n)$. In particular, if P has integral vertices, then $L_P(n)$ is a polynomial.*

Proof: See [Sta, Theorem 4.6.25] or [BR, Theorem 3.23]. □

The function $L_P(t)$ is called the **Ehrhart quasi-polynomial** of P. One can see this as saying that if the vertices of P are $v_i = (v_{i1}, \dots, v_{id})$, then the vertices of nP are given by the linear functions $v_i(n) = (v_{i1}n, \dots, v_{id}n)$. We generalize this as

Theorem 1.3 *Given polynomials $v_{ij}(x), w_{ij}(x) \in \mathbf{Z}[x]$ for $0 \le i \le s$ and $1 \le j \le d$, let n be a positive integer such that $w_{ij}(n) \neq 0$ for all i, j. This is satisfied by n sufficiently large, so we can define a rational polytope $P(n) = \mathrm{conv}(p^0(n), p^1(n), \dots, p^s(n)) \in \mathbf{R}^d$, where $p^i(n) = (\frac{v_{i1}(n)}{w_{i1}(n)}, \dots, \frac{v_{id}(n)}{w_{id}(n)})$. Then $\#(P(n) \cap \mathbf{Z}^d) \in \mathbf{QP}_{\gg 0}$.*

We call the function $\#(P(n) \cap \mathbf{Z}^d)$ a **generalized Ehrhart polynomial**.

2 Equivalence of the two problems

As we shall see, the two problems of the Diophantine equations and lattice point enumeration are closely intertwined. In this section, we want to show that Theorem 1.1 is equivalent to Theorem 1.3. Before this, let us see the equivalence of Theorem 1.3 with the following result. For notation, if x and y are vectors, then $x \ge y$ if $x_i \ge y_i$ for all i.

Theorem 2.1 *For $n \gg 0$, define a rational polytope $P(n) = \{x \in \mathbf{R}^d \mid V(n)x \ge c(n)\}$, where $V(x)$ is an $r \times d$ matrix, and $c(x)$ is an $r \times 1$ column vector, both with entries in $\mathbf{Z}[x]$. Then $\#(P(n) \cap \mathbf{Z}^d) \in \mathbf{QP}_{\gg 0}$.*

Notice that the difference of Theorem 1.3 and Theorem 2.1 is that one defines a polytope by its vertices and the other by hyperplanes. So we will show their equivalence by presenting a generalized version of the algorithm connecting "vertex description" and "hyperplane description" of a polytope.

The connection is based on the fact that we can compare two rational functions $f(n)$ and $g(n)$ when n is sufficiently large. For example, if $f(n) = n^2 - 4n + 1$ and $g(n) = 5n$, then $f(n) > g(n)$ for all $n > 9$, we denote this by $f(n) >_{\text{even}} g(n)$ ("even" being shorthand for "eventually"). Therefore, given a point and a hyperplane, we can test their relative position. To be precise, let $p(n) = (r_1(n), \dots, r_k(n))$ be a point where the $r_i(n)$ are rational functions and let $F(x, n) = a_1(n)x_1 + a_2(n)x_2 + \cdots + a_k(n)x_k = 0$ be a hyperplane where all the $a_i(n)$ are polynomials of n. Then exactly one of the following will be true:

$$F(p, n) =_{\text{even}} 0; \quad F(p, n) >_{\text{even}} 0; \quad F(p, n) <_{\text{even}} 0.$$

Given this, we can make the following definition. We say that two points $p(n)$ and $q(n)$ **lie** (resp., **weakly lie**) **on the same side of** $F(p, n)$ if $F(p, n)F(q, n) >_{\text{even}} 0$ (resp., $F(p, n)F(q, n) \geq_{\text{even}} 0$).

2.1 Equivalence of Theorem 1.3 and Theorem 2.1.

Going from the "vertex description" to the "hyperplane description":

Given all vertices of a polytope $P(n)$, whose coordinates are all rational functions of n, we want to get its "hyperplane description" for $n \gg 0$. Let $F(x, n)$ be a hyperplane defined by a subset of vertices. If all vertices lie weakly on one side of $F(x, n)$, we will keep it together with ≥ 0, or ≤ 0 or $= 0$ indicating the relative position of this hyperplane and the polytope. We can get all the hyperplanes defining the polytope by this procedure.

Going from the "hyperplane description" to the "vertex description":

Let $P(n) = \{x \in \mathbf{R}^d \mid V(n)x \geq c(n)\}$ be a polytope, where $V(x)$ is an $r \times d$ matrix, and $c(x)$ is an $r \times 1$ column vector, both with entries in $\mathbf{Z}[x]$. Without loss of generality, we may assume that $P(n)$ is full-dimensional. We want to find its vertex description. Let $f_1(n), \dots, f_r(n)$ be the linear functionals defined by the rows of $V(n)$. So we can rewrite $P(n)$ as

$$P(n) = \{x \in \mathbf{R}^d \mid \langle f_i(n), x\rangle \geq c_i(n) \text{ for all } i\}.$$

The vertices of $P(n)$ can be obtained as follows. For every d-subset $I \subseteq \{1, \dots, m\}$, if the equations $\{\langle f_i(n), x\rangle = c_i(n) \mid i \in I\}$ are linearly independent for $n \gg 0$, and their intersection is nonempty, then it consists of a single point, which we denote by $v_I(n)$. If $\langle f_j(n), v_I(n)\rangle \geq c_j(n)$ for all j, then $v_I(n) \in \mathbf{Q}(n)^d$ is a vertex of $P(n)$, and all vertices are obtained in this way. We claim that the subsets I for which $v_I(n)$ is a vertex remains constant if we take n sufficiently large. First, the notion of being linearly independent equations can be tested by showing that at least one of the $d \times d$ minors of the rows of $V(n)$ indexed by I does not vanish. Since these minors are all polynomial functions, they can only have finitely many roots unless they are identically zero. Hence taking $n \gg 0$, we can assume that $\{f_i(n) \mid i \in I\}$ is either always linearly dependent or always linearly independent. Similarly, the sign of $\langle f_j(n), v_I(n)\rangle$ is determined by the sign of a polynomial, and hence is constant for $n \gg 0$.

2.2 Equivalence of Theorem 1.1 and Theorem 2.1.

We can easily transform an inequality to an equality by introducing some slack variables and we can also represent an equality $f(n, x) = 0$ by two inequalities $f(n, x) \geq 0$ and $-f(n, x) \geq 0$. So the main difference between the two theorems is that Theorem 1.1 is counting nonnegative solutions while Theorem 2.1 is counting all integral solutions. But we can deal with this by adding constraints on each variable.

A more interesting connection between Theorem 1.1 and Theorem 2.1 is worth mentioning here. First consider any fixed integer n. Then the entries of $A(n)$ and $b(n)$ in the linear Diophantine equations $A(n)x = b(n)$ of Theorem 1.1 all become integers. For an integer matrix, we can calculate its Smith normal form. Similarly, we can use a generalized Smith normal form for matrices over $\mathbf{QP}_{\gg 0}$ to get a transformation from Theorem 1.1 to Theorem 2.1.

Theorem 2.2 *For any matrix $M \in (\mathbf{QP}_{\gg 0})^{k\times s}$, define a matrix function $D\colon \mathbf{Z} \to \mathbf{Z}^{k\times s}$ such that $D(n)$ is the Smith normal form of $M(n)$. Then $D \in (\mathbf{QP}_{\gg 0})^{k\times s}$ and there exists $U \in (\mathbf{QP}_{\gg 0})^{k\times k}$, $V \in (\mathbf{QP}_{\gg 0})^{s\times s}$ such that $U(n), V(n)$ are unimodular (determinant is $+1$ or -1) for $n \gg 0$ and $UMV = D$. We call this matrix function D the* ***generalized Smith normal form*** *of M.*

Then given $A(n)$ and $b(n)$, by Theorem 2.2, we can put $A(n)$ into generalized Smith normal form: $D(n) = U(n)A(n)V(n)$ for some matrix

$$D(n) = (\mathrm{diag}(d_1(n), \ldots, d_r(n), 0, \ldots, 0)|\mathbf{0})$$

with nonzero entries only on its main diagonal, and unimodular matrices $U(n)$ and $V(n)$. Then the equation $A(n)x = b(n)$ can be rewritten as $D(n)V(n)^{-1}x = U(n)b(n)$. Set $y = V(n)^{-1}x$ and $b'(n) = U(n)b(n)$. By the form of $D(n)$, we have a solution y if and only if $d_i(n)$ divides $b'_i(n)$ for $i = 1, \ldots, r$, and for any given solution, the values $y_{r+1}, \ldots, y_k$ can be arbitrary. However, since $V(n)y = x$, we require that $V(n)y \geq 0$, and any such y gives a nonnegative solution x to the original problem. Simplifying $V(n)y \geq 0$, where $V(n) = (v_1(n), \ldots, v_k(n))$, we get $V'(n)X \geq c(n)$, where $V'(n) = (v_{r+1}(n), \ldots, v_k(n))$, $X = (y_{r+1}, \ldots, y_k)$ and $c(n) = -(v_1(n)y_1 + \cdots + v_r(n)y_r)$. Although $V'(n)$ and $c(n)$ has entries in $\mathbf{QP}_{\gg 0}$, we can assume that they are polynomials by dealing with each constituent of the quasi-polynomials separately. So we reduce Theorem 1.1 to Theorem 2.1.

The proof of Theorem 2.2 is based on a theory of generalized division and GCD over the ring $\mathbf{Z}[x]$, which mainly says that for $f(x), g(x) \in \mathbf{Z}[x]$, the functions $\left\lfloor \frac{f(n)}{g(n)} \right\rfloor$, $\left\{ \frac{f(n)}{g(n)} \right\}$, and $\gcd(f(n), g(n))$ lie in the ring $\mathbf{QP}_{\gg 0}$. Once interesting consequence of these results is that every finitely generated ideal in $\mathbf{QP}_{\gg 0}$ is principal, despite the fact that $\mathbf{QP}_{\gg 0}$ is not Noetherian. We developed this theory in order to appoach Theorem 1.1 at first, but subsequently have found a proof that circumvents its use. Further development of these results will appear elsewhere.

3 Lemmas and examples

By the equivalence discussed in Section 2, we only need to prove Theorem 1.1. We give an outline of the proof. The key idea is an elementary "writing in base n" trick, whose use allows us to reduce equations with polynomial coefficients to linear functions. The idea of the following "writing in base n" trick is the following: given a linear Diophantine equation

$$a_1(n)x_1 + a_2(n)x_2 + \cdots + a_k(n)x_k = m(n)$$

with polynomial coefficients $a_i(n)$ and $m(n)$, fix an integer n so that the coefficients all become integers. Now consider a solution $(x_1, x_2, \ldots, x_k)$ with $x_i \in \mathbf{Z}_{\geq 0}$. Substituting the values of $(x_1, x_2, \ldots, x_k)$ into the equation, both sides become integers. Then we use the fact that any integer has a unique representation in base n (n is a fixed number), and compare the coefficient of each power of n in both sides of the equation.

One can show (Lemma 3.1) that the form of this representation in base n is uniform for both sides when n is sufficiently large. Moreover, the coefficient of each power of n in both sides of the equation are all linear functions of n. Using Lemma 3.2, this uniform expression can be reduced to a system of inequalities of the form $f(x) \geq An + B$ where A, B are integers and $f(x)$ is a linear form with constant coefficients. Then by Lemma 3.4, we can reduce these equations with linear coefficients to a case where we can apply Ehrhart's theorem (Theorem 1.2) to show that the number of solutions are quasi-polynomials of n. This completes the proof of Theorem 1.1.

We finish this section with the statements of the above mentioned lemmas and include examples.

Lemma 3.1 *Given $p(x) \in \mathbf{Z}[x]$ with $p(n) > 0$ for $n \gg 0$ (i.e., $p(x)$ has positive leading coefficient), there is a unique representation of $p(n)$ in base n:*

$$p(n) = c_d(n)n^d + \cdots + c_1(n)n + c_0(n),$$

where $c_i(n)$ is a linear function of n such that for $n \gg 0$, $0 \leq c_i(n) \leq n-1$ for $i = 0, 1, \ldots, d$ and $0 < c_d(n) \leq n-1$. We denote $d = \deg_n(p(n))$.

Note that $\deg_n(p(n))$ may not be equal to $\deg(p(n))$. For example, $n^2 - n + 3$ is represented as $c_1(n)n + c_0(n)$ with $d = 1$, $c_1(n) = n-1$, and $c_0(n) = 3$.

Now fix a positive integer n. We have a unique expression of any integer x written in base n, if we know an upper bound d of the highest power, as $x = x_d n^d + x_{d-1}n^{d-1} + \cdots + x_1 n + x_0$ with $0 \leq x_i < n$. This gives us a bijection between the set $\{(x_1, \ldots, x_k) \in \mathbf{Z}^k_{\geq 0}\}$ and the set

$$\{0 \leq (x_{ij})_{\substack{1 \leq i \leq k \\ 0 \leq j \leq d-d_i}} < n, x_{ij} \in \mathbf{Z}\}.$$

Then by a direct "base n" comparison starting from the lowest power to the highest power, we have the following lemma.

Lemma 3.2 *For $n \gg 0$, there is a one to one correspondence between the following two sets:*

$$S_1 = \{(x_1, \ldots, x_k) \in (\mathbf{Z}_{\geq 0})^k \mid a_1(n)x_1 + a_2(n)x_2 + \cdots + a_k(n)x_k = m(n)\}$$

where $a_i(n) = \sum_{\ell=0}^{d_i} a_{i\ell}n^\ell$ (as a usual polynomial) with $a_{id_i} > 0$, $i = 1, \ldots, k$, and $m(n) = \sum_{\ell=0}^{d} b_\ell n^\ell$ (represented in base n as in Lemma 3.1), with $b_d > 0$ and $d \geq \max_{1 \leq i \leq k}\{d_i\}$.

$$S_2 = \{0 \leq (x_{ij})_{\substack{1 \leq i \leq k \\ 0 \leq j \leq d-d_i}} < n, x_{ij} \in \mathbf{Z} \mid \text{ all constraints on } x = (x_{ij}) \text{ are of the form } An + B \leq f(x)\}$$

where $A, B \in \mathbf{Z}$ and $f(x)$ is a linear form of x with constant coefficients.

For a lower bound on n in the above lemma, the sum of all absolute value of coefficients $1 + \sum_{i=1}^{k} \sum_{\ell=0}^{d_i} |a_{i\ell}| + \sum_{\ell=0}^{d} |b_\ell|$ is sufficient.

Example 3.3 We give an example of Lemma 3.2. Consider nonnegative integer solutions for

$$2x_1 + (n+1)x_2 + n^2x_3 = 4n^2 + 3n - 5.$$

For any $n > 5$, RHS $= 4n^2 + 2n + (n-5)$ is the expression in base n. Now consider the left hand side. Writing x_1, x_2, x_3 in base n, let $x_1 = x_{12}n^2 + x_{11}n + x_{10}$, $x_2 = x_{21}n + x_{20}$ and $x_3 = x_{30}$ with $0 \le x_{ij} < n$. Then we have

$$\text{LHS} = (2x_{12} + x_{21} + x_{30})n^2 + (2x_{11} + x_{21} + x_{20})n + (2x_{10} + x_{20}).$$

Now we can write the left hand side in base n with extra constraints on (x_{ij})'s.

We start with comparing the coefficient of n^0 in both sides. We have the following three cases:

$$\begin{aligned}
A_0^0 &= \{0 \le 2x_{10} + x_{20} < n, 2x_{10} + x_{20} = n - 5\},\\
A_1^0 &= \{n \le 2x_{10} + x_{20} < 2n, 2x_{10} + x_{20} = (n-5) + n\},\\
A_2^0 &= \{2n \le 2x_{10} + x_{20} < 3n, 2x_{10} + x_{20} = (n-5) + 2n\}.
\end{aligned}$$

We next consider the n^1 term. If x satisfies A_i^0 for n^0, $i \in I_0 = \{0, 1, 2\}$, then the equation is reduced to

$$(2x_{12} + x_{21} + x_{30})n^2 + (2x_{11} + x_{21} + x_{20} + i)n = 4n^2 + 2n.$$

Now compare the n^1 terms. We have five cases for each $i \in I_0 = \{0, 1, 2\}$.

$$A_{ij}^1 = \{jn \le 2x_{11} + x_{21} + x_{20} + i < (j+1)n,\ 2x_{11} + x_{21} + x_{20} + i = jn + 2\},$$

where $j \in I_1 = \{0, 1, 2, 3, 4\}$.

Last, we compare the n^2 terms. Note that since we assume $n \gg 0$, the n^0 term won't affect the n^2 term, so the computation of n^2 term only depends on the term n^1. If x satisfies the jth condition for n^1, the equation then becomes

$$(2x_{12} + x_{21} + x_{30} + j)n^2 = 4n^2.$$

So for each $j \in I_1$, we have

$$A_j^2 = \{2x_{12} + x_{21} + x_{30} + j = 4\}.$$

Overall, we have the set

$$\{(x_1, x_2, x_3) \in \mathbf{Z}_{\ge 0}^3 \mid 2x_1 + (n+1)x_2 + n^2x_3 = 4n^2 + 3n - 5\}$$

is in bijection with the set

$$\{x = (x_{12}, x_{11}, x_{10}, x_{21}, x_{20}, x_{30}) \in \mathbf{Z}_{\ge 0}^6, 0 \le x_{ij} < n\},$$

such that x satisfies the conditions

$$\begin{pmatrix} A_0^0 & A_1^0 & A_2^0 \end{pmatrix} \begin{pmatrix} A_{00}^1 & A_{01}^1 & \cdots & A_{04}^1 \\ A_{10}^1 & A_{11}^1 & \cdots & A_{14}^1 \\ A_{20}^1 & A_{21}^1 & \cdots & A_{24}^1 \end{pmatrix} \begin{pmatrix} A_0^2 \\ A_1^2 \\ \vdots \\ A_4^2 \end{pmatrix}.$$

Here we borrow the notation of matrix multiplication AB to represent intersection of sets $A \cap B$ and matrix summation $A + B$ to represent set union $A \cup B$. Note that here all constrains A_i^j on $x = (x_{12}, x_{11}, x_{10}, x_{21}, x_{20}, x_{30})$ are in the form of $An + B \leq f(x)$, where $A, B \in \mathbf{Z}$ and $f(x)$ is a linear form of x with constant coefficients.

The following lemma allows us to reduce these equations (or inequalities) with linear function coefficients to the case when we can apply Ehrhart's theorem (Theorem 1.2) and show that the number of solutions are quasi-polynomials of n.

Lemma 3.4 *If $P(n) \subset \mathbf{R}^d$ is a polytope defined by inequalities of the form $An + B \leq f(x)$, where $A, B \in \mathbf{Z}$ and $f(x)$ is a linear form of x with constant coefficients, then $\#(P(n) \cap \mathbf{Z}^d) \in \mathbf{QP}_{\gg 0}$.*

Example 3.5 For n a positive integer, let $P(n)$ be the polygon defined by the inequalities $x \geq 0$, $y \geq 0$ and $-2x - y \geq -n - 1$. Then $P'(n)$ is defined by the inequalities $x \geq 0$, $y \geq 0$, and $2x + y \leq n$, and $P_1(n)$ is defined by the inequalities $x \geq 0$, $y \geq 0$, and $n + 1 = 2x + y$. We can rewrite the equality as $y = n + 1 - 2x$, and then the other inequalities become $x \geq 0$ and $n + 1 \geq 2x$.

We see that $P'(n)$ is the convex hull of the points $\{(0,0), (0,n), (n/2, 0)\}$, while $P_1(n)$ is the interval $[0, (n+1)/2]$. The total number of integer points in $P'(n)$ and $P_1(n)$ is given by the quasipolynomial

$$\#(P(n) \cap \mathbf{Z}^2) = \begin{cases} k^2 + 3k + 2 & \text{if } n = 2k \\ k^2 + 4k + 4 & \text{if } n = 2k+1 \end{cases}.$$

Its rational generating function is

$$\sum_{n \geq 0} \#(P(n) \cap \mathbf{Z}^2) t^n = \frac{t^5 - 3t^3 + 4t + 2}{(1 - t^2)^3} = \frac{t^3 - 2t^2 + 2}{(1-t)^3(1+t)}.$$

References

[BR] Matthias Beck and Sinai Robins, *Computing the Continuous Discretely: Integer-point enumeration in polyhedra*, Undergraduate Texts in Mathematics, Springer, New York, 2007, available for download from `http://math.sfsu.edu/beck/ccd.html`.

[Ehr] E. Ehrhart, *Polynômes arithmétiques et Méthode des Polyèdres en Combinatoire*, International Series of Numerical Mathematics, vol. 35, Birkhäuser Verlag, Basel/Stuttgart, 1977.

[Sta] Richard P. Stanley, *Enumerative Combinatorics, Vol. I*, Cambridge Studies in Advanced Mathematics **49**, Cambridge University Press, 1997.

[Xu] Zhiqiang Xu, An explicit formulation for two dimensional vector partition functions, *Integer Points in Polyhedra—Geometry, Number Theory, Representation Theory, Algebra, Optimization, Statistics*, Contemporary Mathematics **452** (2008), 163–178.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 127–137

Descent polynomials for permutations with bounded drop size

Fan Chung[1], Anders Claesson[2], Mark Dukes[3†]and Ronald Graham[1]

[1] *University of California at San Diego, La Jolla CA 92093, USA.*
[2] *The Mathematics Institute, School of Computer Science, Reykjavík University, 103 Reykjavík, Iceland.*
[3] *Science Institute, University of Iceland, 107 Reykjavík, Iceland.*

Abstract. Motivated by juggling sequences and bubble sort, we examine permutations on the set $\{1, 2, \ldots, n\}$ with d descents and maximum drop size k. We give explicit formulas for enumerating such permutations for given integers k and d. We also derive the related generating functions and prove unimodality and symmetry of the coefficients.

Résumé. Motivés par les "suites de jonglerie" et le tri à bulles, nous étudions les permutations de l'ensemble $\{1, 2, \ldots, n\}$ ayant d descentes et une taille de déficience maximale k. Nous donnons des formules explicites pour l'énumération de telles permutations pour des entiers k et d fixés, ainsi que les fonctions génératrices connexes. Nous montrons aussi que les coefficients possèdent des propriétés d'unimodalité et de symétrie.

Keywords: Permutations, descent polynomial, drop size, Eulerian distribution.

1 Introduction

There have been extensive studies of various statistics on $\mathcal{S}_n$, the set of all permutations of $\{1, 2, \ldots, n\}$. For a permutation π in $\mathcal{S}_n$, we say that π has a *drop* at i if $\pi_i < i$ and that the *drop size* is $i - \pi_i$. We say that π has a *descent* at i if $\pi_i > \pi_{i+1}$. One of the earliest results [8] in permutation statistics states that the number of permutations in $\mathcal{S}_n$ with k drops equals the number of permutations with k descents. A concept closely related to drops is that of *excedances*, which is just a drop of the inverse permutation. In this paper we focus on drops instead of excedances because of their connection with our motivating applications concerning bubble sort and juggling sequences.

Other statistics on a permutation π include such things as the number of *inversions* (pairs (i, j) such that $i < j$ and $\pi_i > \pi_j$) and the *major index* of π (the sum of i for which a descent occurs). The enumeration of and generating functions for these statistics can be traced back to the work of Rodrigues in 1839 [9] but was mainly influenced by McMahon's treatise in 1915 [8]. There is an extensive literature studying the distribution of the above statistics and their q-analogs, see for example Foata and Han [4], or the papers of Shareshian and Wachs [10, 11] for more recent developments.

This joint work originated from its connection with a paper [2] on sequences that can be translated into juggling patterns. The set of juggling sequences of period n containing a specific state, called the ground

† AC and MD were supported by grant no. 090038011 from the Icelandic Research Fund.

state, corresponds to the set $\mathcal{B}_{n,k}$ of permutations in $\mathcal{S}_n$ with drops of size at most k. As it turns out, $\mathcal{B}_{n,k}$ can also be associated with the set of permutations that can be sorted by k operations of bubble sort. These connections will be further described in the next section. We note that the *maxdrop* statistic has not been treated in the literature as extensively as many other statistics in permutations. As far as we know, this is the first time that the distribution of descents with respect to maxdrop has been determined.

First we give some definitions concerning the statistics and polynomials that we examine. Given a permutation π in $\mathcal{S}_n$, let $\mathrm{Des}(\pi) = \{1 \le i < n : \pi_i > \pi_{i+1}\}$ be the descent set of π and let $\mathrm{des}(\pi) = |\,\mathrm{Des}(\pi)|$ be the number of descents. We use $\mathrm{maxdrop}(\pi)$ to denote the value of the maximum drop (or maxdrop) of π,

$$\mathrm{maxdrop}(\pi) = \max\{\, i - \pi(i) : 1 \le i \le n \,\}.$$

Let $\mathcal{B}_{n,k} = \{\pi \in \mathcal{S}_n : \mathrm{maxdrop}(\pi) \le k\}$. It is known, and also easy to show, that $|\mathcal{B}_{n,k}| = k!(k+1)^{n-k}$; e.g., see [2, Thm. 1] or [7, p. 108]. Let

$$b_{n,k}(r) = |\{\pi \in \mathcal{B}_{n,k} : \mathrm{des}(\pi) = r\}|,$$

and define the (k-maxdrop-restricted) descent polynomial

$$B_{n,k}(x) = \sum_{r \ge 0} b_{n,k}(r) x^r = \sum_{\pi \in \mathcal{B}_{n,k}} x^{\mathrm{des}(\pi)}.$$

Examining the case of $k = 2$, we observed the coefficients $b_{n,2}(r)$ of $B_{n,2}(x)$ appear to be given by every *third* coefficient of the simple polynomial

$$(1 + x^2)(1 + x + x^2)^{n-1}.$$

Looking at the next two cases, $k = 3$ and $k = 4$, yielded more mysterious polynomials: $b_{n,3}(r)$ appeared to be every fourth coefficient of

$$(1 + x^2 + 2x^3 + x^4 + x^6)(1 + x + x^2 + x^3)^{n-2}$$

and $b_{n,4}(r)$ every fifth coefficient of

$$(1 + x^2 + 2x^3 + 4x^4 + 4x^5 + 4x^7 + 4x^8 + 2x^9 + x^{10} + x^{12})(1 + x + x^2 + x^3 + x^4)^{n-3}.$$

After a fierce battle with these polynomials, we were able to show that $b_{n,k}(r)$ is the coefficient of $u^{r(k+1)}$ in the polynomial

$$P_k(u)\left(1 + u + \cdots + u^k\right)^{n-k} \tag{1}$$

where

$$P_k(u) = \sum_{j=0}^{k} A_{k-j}(u^{k+1})(u^{k+1} - 1)^j \sum_{i=j}^{k} \binom{i}{j} u^{-i}, \tag{2}$$

and A_k denotes the kth Eulerian polynomial (defined in the next section). Further to this, we give an expression for the generating function $\mathbf{B}_k(z, y) = \sum_{n \ge 0} B_{n,k}(y) z^n$, namely

$$\mathbf{B}_k(z, y) = \frac{1 + \displaystyle\sum_{t=1}^{k} \left(A_t(y) - \sum_{i=1}^{t} \binom{k+1}{i} (y-1)^{i-1} A_{t-i}(y) \right) z^t}{1 - \displaystyle\sum_{i=1}^{k+1} \binom{k+1}{i} z^i (y-1)^{i-1}}. \tag{3}$$

We also give some alternative formulations for P_k which lead to some identities involving Eulerian numbers as well as proving the symmetry and unimodality of the polynomials $B_{n,k}(x)$.

Many questions remain. For example, is there a more natural bijective proof for the formulas that we have derived for $B_{n,k}$ and $\mathbf{B}_k$? Why do permutations that are k-bubble sortable define the aforementioned juggling sequences?

2 Descent polynomials, bubble sort and juggling sequences

We first state some standard notation. The polynomial

$$A_n(x) = \sum_{\pi \in \mathcal{S}_n} x^{\mathrm{des}(\pi)}$$

is called the nth *Eulerian polynomial*. For instance, $A_0(x) = A_1(x) = 1$ and $A_2(x) = 1 + x$. Note that $B_{n,k}(x) = A_n(x)$ for $k \geq n-1$, since $\mathrm{maxdrop}(\pi) \leq n-1$ for all $\pi \in \mathcal{S}_n$. The coefficient of x^k in $A_n(x)$ is denoted $\left\langle {n \atop k} \right\rangle$ and is called an *Eulerian number*. It is well known that [5]

$$\frac{1-w}{e^{(w-1)z} - w} = \sum_{k,n \geq 0} \left\langle {n \atop k} \right\rangle w^k \frac{z^n}{n!}. \tag{4}$$

The Eulerian numbers are also known to be given explicitly as [3, 5] $\left\langle {n \atop k} \right\rangle = \sum_{i=0}^{n} \binom{n+1}{i} (k+1-i)^n (-1)^i$.

We define the operator bubble which acts recursively on permutations via

$$\mathsf{bubble}(LnR) = \mathsf{bubble}(L)Rn.$$

In other words, to apply bubble to a permutation π in $\mathcal{S}_n$, we split π into (possibly empty) blocks L and R to the left and right, respectively, of the largest element of π (which initially is n), interchange n and R, and then recursively apply this procedure to L. We will use the convention that $\mathsf{bubble}(\emptyset) = \emptyset$; here $\emptyset$ denotes the empty permutation. This operator corresponds to one *pass* of the classical bubble sort operation. Several interesting results on the analysis of bubble sort can be found in Knuth [7, pp. 106–110]. We define the *bubble sort complexity* of π as

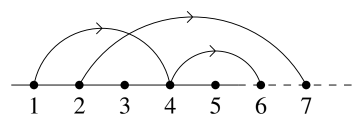
$$\mathrm{bsc}(\pi) = \min\{k : \mathsf{bubble}^k(\pi) = \mathrm{id}\},$$

the number of times bubble must be applied to π to give the identity permutation. The following lemma is easy to prove using induction.

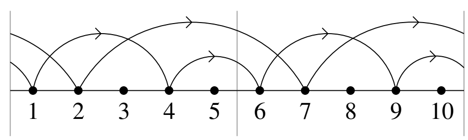
Lemma 1 (i) *For all permutations* π *we have* $\mathrm{maxdrop}(\pi) = \mathrm{bsc}(\pi)$.
(ii) *The bubble sort operator maps* $\mathcal{B}_{n,k}$ *to* $\mathcal{B}_{n,k-1}$.

The class of permutations $\mathcal{B}_{n,k}$ appears in a recent paper [2] on enumerating juggling patterns that are usually called *siteswaps* by (mathematically inclined) jugglers. Suppose a juggler throws a ball at time i so that the ball will be in the air for a time t_i before landing at time $t_i + i$. Instead of an infinite sequence, we will consider periodic patterns, denoted by $T = (t_1, t_2, \ldots, t_n)$. A *juggling sequence* is just one in which two balls never land at the same time. It is not hard to show [1] that a necessary and sufficient condition for a sequence to be a juggling sequence is that all the values $t_i + i \pmod{n}$ are distinct. In particular, it follows that that the average of t_i is just the numbers of balls being juggled. Here is an example:

If $T = (3, 5, 0, 2, 0)$ then at time 1 a ball is thrown that will land at time $1 + 3 = 4$. At time 2 a ball is thrown that will land at time $2 + 5 = 7$. At time 3 a ball is thrown that will land at time $3 + 0 = 3$. Alternatively one can say that no ball is thrown at time 3. This is represented in the following diagram.



Repeating this for all intervals of length 5 gives



For a given juggling sequence, it is often possible to further decompose into shorter juggling sequences, called *primitive juggling sequences*, which themselves cannot be further decomposed. These primitive juggling sequences act as basic building blocks for juggling sequences [2]. However, in the other direction, it is not always possible to combine primitive juggling sequences into a longer juggling sequence. Nevertheless, if primitive juggling sequences share a common *state* (which one can think of as a *landing schedule*), then we can combine them to form a longer and more complicated juggling sequences. In [2] primitive juggling sequences associated with a specified state are enumerated. Here we mention the related fact concerning $\mathcal{B}_{n,k}$:

There is a bijection mapping permutations in $\mathcal{B}_{n,k}$ to primitive juggling sequences of period n with k balls that all share a certain state, called the ground state.

The bijection maps π to $\phi(\pi) = (t_1, \ldots, t_n)$ with $t_i = k - i + \pi_i$. As a consequence of the above fact and Lemma 1, we can use bubble sort to transform a juggling sequence using k balls to a juggling sequence using $k - 1$ balls.

To make this more precise, let $T = (t_1, \ldots, t_n)$ be a juggling sequence that corresponds to $\pi \in \mathcal{B}_{n,k}$, and suppose that $T' = (s_1, \ldots, s_n)$ is the juggling sequence that corresponds to $\mathsf{bubble}(\pi)$. Assume that the ball B thrown at time j is the one that lands latest out of all the n throws. In other words, $t_j + j$ is the largest element in $\{t_i + i\}_{i=1}^{n}$. Now, write $T = Lt_jR$ where $L = (t_1, \ldots, t_{j-1})$ and $R = (t_{j+1}, \ldots, t_n)$. Then we have

$$T' = f_k(T) = f_k(L)Rs,$$

where $s = t_j + j - (n + 1)$. In other words, we have removed the ball B thrown at time j and thus throw all balls after time j one time unit sooner. Then at time n we throw the ball B so that it lands one time unit sooner than it would have originally landed. Then we repeat this procedure to all the balls thrown before time j.

3 The polynomials $B_{n,k}(y)$

In this section we will characterise the polynomials $B_{n,k}(y)$. This is done by first finding a recurrence for the polynomials and then solving the recurrence by exploiting some aspects of their associated characteristic polynomials. The latter step is quite involved and so we present the special case dealing with $B_{n,4}(y)$ first.

3.1 Deriving the recurrence for $B_{n,k}(y)$

We will derive the following recurrence for $B_{n,k}(y)$.

Theorem 1 *For* $n \geq 0$*,*

$$B_{n+k+1,k}(y) = \sum_{i=1}^{k+1} \binom{k+1}{i} (y-1)^{i-1} B_{n+k+1-i,k}(y) \tag{5}$$

with the initial conditions

$$B_{i,k}(y) = A_i(y), \quad 0 \leq i \leq k.$$

We use the notation $[a, b] = \{i \in \mathbb{Z} : a \leq i \leq b\}$ and $[b] = [1, b]$. Let $A = \{a_1, \ldots, a_n\}$ with $a_1 < \cdots < a_n$ be any finite subset of $\mathbb{N}$. The *standardization* of a permutation π on A is the permutation $\mathrm{st}(\pi)$ on $[n]$ obtained from π by replacing the integer a_i with the integer i. Thus π and $\mathrm{st}(\pi)$ are order isomorphic. For example, $\mathrm{st}(19452) = 15342$. If the set A is fixed, the inverse of the standardization map is well defined, and we denote it by $\mathrm{st}_A^{-1}(\sigma)$; for instance, with $A = \{1, 2, 4, 5, 9\}$, we have $\mathrm{st}_A^{-1}(15342) = 19452$. Note that st and st_A^{-1} each preserve the descent set.

For any set $S \subseteq [n-1]$ we define $\mathcal{A}_{n,k}(S) = \{\pi \in \mathcal{B}_{n,k} : \mathrm{Des}(\pi) \supseteq S\}$ and

$$t_n(S) = \max\{i \in \mathbb{N} : [n-i, n-1] \subseteq S\}.$$

Note that $t_n(S) = 0$ in the case that $n-1$ is not a member of S. Now, for any permutation $\pi = \pi_1 \ldots \pi_n$ in $\mathcal{A}_{n,k}(S)$ define

$$f(\pi) = (\sigma, X), \text{ where } \sigma = \mathrm{st}(\pi_1 \ldots \pi_{n-i-1}), X = \{\pi_{n-i}, \ldots, \pi_n\} \text{ and } i = t_n(S).$$

Example 1 *Let* $S = \{3, 7, 8\}$*, and choose the permutation* $\pi = 138425976$ *in* $\mathcal{A}_{9,3}(S)$*. Notice that* $\mathrm{Des}(\pi) = \{3, 4, 7, 8\} \supset S$*. Now* $t_9(S) = 2$*. This gives* $f(\pi) = (\sigma, X)$ *where* $\sigma = \mathrm{st}(138425) = 136425$ *and* $X = \{\pi_7, \pi_8, \pi_9\} = \{6, 7, 9\}$*. Hence* $f(138425976) = (136425, \{6, 7, 9\})$*.*

Lemma 2 *For any* π *in* $\mathcal{A}_{n,k}(S)$*, the image* $f(\pi)$ *is in the Cartesian product*

$$\mathcal{A}_{n-i-1,k}(S \cap [n - t_n(S) - 2]) \times \binom{[n-k, n]}{t_n(S) + 1},$$

where $\binom{X}{m}$ *denotes that set of all* m*-element subsets of the set* X*.*

Proof: Given $\pi \in \mathcal{A}_{n,k}(S)$, let $f(\pi) = (\sigma, X)$. Suppose $i = t_n(S)$. Then there are descents at positions $n-i, \ldots, n-1$ (this is an empty sequence in case $i = 0$). Thus

$$n \geq \pi_{n-i} > \pi_{n-i+1} > \cdots > \pi_{n-1} > \pi_n \geq n-k,$$

where the last inequality follows from the assumption that $\mathrm{maxdrop}(\pi) \leq k$. Hence X is an $(i+1)$-element subset of $[n-k, n]$, as claimed. Clearly $\sigma \in \mathcal{S}_{n-i-1}$.

Next we shall show that σ is in $\mathcal{A}_{n-i-1,k}$. Notice that the entries of $(\pi_1,\ldots,\pi_{n-i-1})$ that do not change under standardization are those π_ℓ which are less than π_n. Since these values remain unchanged, the values $\ell-\pi_\ell$ are also unchanged and are thus at most k.

Let $(\pi_{a(1)},\ldots,\pi_{a(m)})$ be the subsequence of values which are greater than π_n. The smallest value that any of these may take after standardization is $\pi_n \geq n-k$. So $\sigma_{a(j)} \geq \pi_n \geq n-k$ for all $j \in [1,m]$. Thus $a(j)-\sigma_{a(j)} \leq a(j)-(n-k) = k-(n-a(j)) \leq k$ for all $j \in [1,m]$. Therefore $\ell-\sigma_\ell \leq k$ for all $\ell \in [1,n-i-1]$ and so $\sigma \in \mathcal{A}_{n-i-1,k}$.

The descent set is preserved under standardization, and consequently σ is in $\mathcal{A}_{n-i-1,k}(S\cap[n-i-2])$, as claimed. $\square$

We now define a function g which will be shown to be the inverse of f. Let π be a permutation in $\mathcal{A}_{m,k}(T)$, where T is a subset of $[m-1]$. We will add $i+1$ elements to π to yield a new permutation σ in $\mathcal{A}_{m+i,k}(T\cup[m+1,m+i])$. Choose any $(i+1)$-element subset X of the interval $[m+i+1-k,m+i+1]$, and let us write $X=\{x_1,\ldots,x_{i+1}\}$, where $x_1 \leq \cdots \leq x_{i+1}$. Define

$$g(\pi,X) = \mathrm{st}_V^{-1}(\pi_1\ldots\pi_m)\,x_{i+1}x_i\ldots x_1, \text{ where } V=[m+i+1]\setminus X.$$

Example 2 *Let* $T=\{1\}$*, and choose the permutation* $\pi=3142$ *in* $\mathcal{A}_{4,3}(T)$*. Notice that* $\mathrm{Des}(\pi)=\{1,3\}\supseteq T$*. Choose* $i=2$ *and select a subset* X *from* $[4+2+1-3,4+2+1]=\{4,5,6,7\}$ *of size* $i+1=3$*. Let us select* $X=\{4,6,7\}$*. Now we have* $g(\pi,X)=\mathrm{st}_V^{-1}(3142)\,764=3152764$*, where* V *is the set* $[4+2+1]\setminus\{4,6,7\}=\{1,2,3,5\}$.

Lemma 3 *If* (π,X) *is in the Cartesian product*

$$\mathcal{A}_{m,k}(T)\times\binom{[m+i+1-k,m+i+1]}{i+1}$$

for some $i>0$ *then* $g(\pi,X)$ *is in*

$$\mathcal{A}_{m+i+1,k}(T\cup[m+1,m+i]).$$

Proof: Let $\sigma=g(\pi,X)$. For the first m elements of σ, since $\sigma_j \geq \pi_j$ for all $1\leq j\leq m$, we have $j-\sigma_j \leq j-\pi_j$ which gives

$$\max\{j-\sigma_j : j\in[m]\} \leq \max\{j-\pi_j : j\in[m]\} \leq k.$$

The final $i+1$ elements of σ are decreasing so the $\mathrm{maxdrop}$ of these elements will be the $\mathrm{maxdrop}$ of the final element,

$$m+i+1-\sigma_{m+i+1} = m+i+1-x_1 \leq m+i+1-(m+i+1-k) = k.$$

Thus $\mathrm{maxdrop}(\sigma)\leq k$ and so $\sigma\in\mathcal{B}_{m+i+1,k}$. The descents of σ will be in the set $T\cup[m+1,m+i]$ since descents are preserved under standardization and the final $i+1$ elements of σ are listed in decreasing order. Hence $\sigma\in A_{m+i+1,k}(T\cup[m+1,m+i])$, as claimed. $\square$

We omit the straightforward, but a bit tedious, proof of the following important Lemma.

Lemma 4 *The function* f *is a bijection, and* g *is its inverse.*

Corollary 1 *Let $a_{n,k}(S) = |\mathcal{A}_{n,k}(S)|$ and $i = t_n(S)$. Then*

$$a_{n,k}(S) = \binom{k+1}{i+1} a_{n-(i+1),k}(S \cap [1, n-(i+1)]).$$

Proposition 1 *For all $n \geq 0$,*

$$\mathcal{B}_{n,k}(y+1) = \sum_{i=1}^{k+1} \binom{k+1}{i} y^{i-1} \mathcal{B}_{n-i,k}(y+1).$$

Proof: Notice that

$$\begin{aligned}
\mathcal{B}_{n,k}(y+1) &= \sum_{\pi \in \mathcal{B}_{n,k}} (y+1)^{\mathrm{des}(\pi)} \\
&= \sum_{\pi \in \mathcal{B}_{n,k}} \sum_{i=0}^{\mathrm{des}(\pi)} \binom{\mathrm{des}(\pi)}{i} y^i \\
&= \sum_{\pi \in \mathcal{B}_{n,k}} \sum_{S \subseteq \mathrm{Des}(\pi)} y^{|S|} \\
&= \sum_{S \subseteq [n-1]} y^{|S|} \sum_{\pi \in \mathcal{A}_{n,k}(S)} 1 = \sum_{S \subseteq [n-1]} y^{|S|} a_{n,k}(S).
\end{aligned}$$

From Corollary 1, multiply both sides by $y^{|S|}$ and sum over all $S \subseteq [n-1]$. We have

$$\begin{aligned}
\mathcal{B}_{n,k}(y+1) &= \sum_{S \subseteq [n-1]} y^{|S|} \binom{k+1}{t_n(S)+1} a_{n-(t_n(S)+1),k}(S \cap [n-(t_n(S)+2)]) \\
&= \sum_{i \geq 0} \sum_{\substack{S \subseteq [n-1] \\ t_n(S)=i}} y^i y^{|S|-i} \binom{k+1}{i+1} a_{n-(i+1),k}(S \cap [n-(i+2)]) \\
&= \sum_{i \geq 0} \binom{k+1}{i+1} y^i \sum_{\substack{S \subseteq [n-1] \\ t_n(S)=i}} a_{n-(i+1),k}(S \cap [n-(i+2)]) y^{|S|-i} \\
&= \sum_{i \geq 0} \binom{k+1}{i+1} y^i \sum_{S \subseteq [n-(i+1)]} a_{n-(i+1),k}(S) y^{|S|} \\
&= \sum_{i \geq 0} \binom{k+1}{i+1} y^i \mathcal{B}_{n-(i+1),k}(y+1) \\
&= \sum_{i \geq 1} \binom{k+1}{i} y^{i-1} \mathcal{B}_{n-i,k}(y+1).
\end{aligned}$$

□

Proof of Theorem 1: Replacing n and y by $n+k+1$ and $y-1$, respectively, in Proposition 1 yields the recurrence (5):

$$B_{n+k+1,k}(y) = \sum_{i=1}^{k+1} \binom{k+1}{i} (y-1)^{i-1} B_{n+k+1-i,k}(y)$$

for $n \geq 0$, with the initial conditions $B_{i,k}(y) = A_i(y)$ for $0 \leq i \leq k$. □

By multiplying the above recurrence by z^n and summing over all $n \geq 0$, we have the generating function $\mathbf{B}_k(z,y)$ given in equation (3).

3.2 Solving the recurrence for $B_{n,4}(y)$.

Before we proceed to solve the recurrence for $B_{n,k}$, we first examine the special case of $k = 4$ which is quite illuminating. We note that the characteristic polynomial for the recurrence for $B_{n,4}$ is

$$h(z) = z^5 - 5z^4 + 10(1-y)z^3 - 10(1-y)^2 z^2 + 5(1-y)^3 z - (1-y)^4 = \frac{(z-1+y)^5 - yz^5}{1-y}.$$

Substituting $y = t^5$ in the expression above, we see that the roots of $h(z)$ are just

$$\rho_j(t) = \frac{1-t^5}{1-\omega^j t}, \quad 0 \leq j \leq 4,$$

where $\omega = \exp(\frac{2\pi \mathrm{i}}{5})$ is a primitive 5th root of unity. Hence, the general term for $B_{n,4}(t)$ can written as

$$B_{n,4}(t) = \sum_{i=0}^{4} \alpha_i(t) \rho_i^n(t)$$

where the $\alpha_i(t)$ are appropriately chosen coefficients (polynomials in t). To determine the $\alpha_i(t)$ we need to solve the following system of linear equations:

$$\sum_{i=0}^{4} \alpha_i(t) \rho_i^j(t) = B_{j,4}(t) = A_j(t^5), \quad 0 \leq j \leq 4.$$

Thus, $\alpha_i(t)$ can be expressed as the ratio $N_{4,i+1}(t)/D_4(t)$ of two determinants. The denominator $D_4(t)$ is just a standard Vandermonde determinant whose $(i+1, j+1)$ entry is $\rho_i^j(t)$. The numerator $N_{4,i+1}(t)$ is formed from $D_4(t)$ by replacing the elements $\rho_i^j(t)$ in the $(i+1)$st row by $A_j(t^5)$. A quick computation (using the symbolic computation package Maple) gives:

$$\begin{aligned} D_4(t) &= 25\sqrt{5}\,(1-t^5)^6 t^{10}; \\ N_{4,1}(t) &= 5\sqrt{5}\,(t^{12} + t^{10} + 2t^9 + 4t^8 + 4t^7 + 4t^5 + 4t^4 + 2t^3 + t^2 + 1)(1-t^5)^3(1-t)^3 t^{10} \end{aligned}$$

and, in general, $N_{4,i+1}(t) = N_{4,1}(\omega^i t)$. Substituting the value $\alpha_0(t) = N_{4,1}(t)/D_4(t)$ into the first term in the expansion of $B_{n,4}$, we get

$$\begin{aligned} &\alpha_0(t)(1+t+t^2+t^3+t^4)^n \\ &\qquad = \tfrac{1}{5}(t^{12} + t^{10} + 2t^9 + 4t^8 + 4t^7 + 4t^5 + 4t^4 + 2t^3 + t^2 + 1)(1+t+t^2+t^3+t^4)^{n-3}. \end{aligned}$$

Now, since the other four terms $\alpha_i(t)(1+t+t^2+t^3+t^4)^n$ arise by replacing t by $\omega^i t$ then in the sum of all five terms, the only powers of t that survive are those which have powers which are multiples of 5. Thus, we can conclude that if we write

$$(t^{12}+t^{10}+2t^9+4t^8+4t^7+4t^5+4t^4+2t^3+t^2+1)(1+t+t^2+t^3+t^4)^{n-3}=\sum_r \beta(r)t^r$$

then $b_{n,4}(d)=\beta(5d)$. In other words, the number of permutations $\pi\in\mathcal{B}_{n,4}$ with d descents is given by the coefficient of t^{5d} in the expansion of the above polynomial. Incidentally, the corresponding results for the earlier $\mathcal{B}_{n,i}$ are as follows: $b_{n,1}(d)=\beta(2d)$ in the expansion of

$$(1+t)^n=\sum_r \beta(r)t^r,$$

so $b_{n,1}(d)=\binom{n}{2d}$; $b_{n,2}(d)=\beta(3d)$ in the expansion of

$$(1+t^2)(1+t+t^2)^{n-1}=\sum_r \beta(r)t^r;$$

and $b_{n,3}(d)=\beta(4d)$ in the expansion of

$$(1+t^2+2t^3+t^4+t^6)(1+t+t^2+t^3)^{n-2}=\sum_r \beta(r)t^r.$$

The preceding arguments have now set the stage for dealing with the general case of $B_{n,k}$. Of course, the arguments will be somewhat more involved but it is hoped that treating the above special case will be a useful guide for the reader.

3.3 Solving the recurrence for $B_{n,k}(y)$

Theorem 2 *We have $B_{n,k}(y)=\sum_d \beta_k\big((k+1)d\big)y^d$, where*

$$\sum_j \beta_k(j)u^j = P_k(u)\left(1+u+\cdots+u^k\right)^{n-k}$$

and

$$P_k(u)=\sum_{j=0}^{k} A_{k-j}(u^{k+1})(u^{k+1}-1)^j \sum_{i=j}^{k}\binom{i}{j}u^{-i}. \tag{6}$$

The first few values of $P_k(u)$ are shown below.

k	$P_k(u)$
0	1
1	$1+u$
2	$1+u+2u^2+u^3+u^4$
3	$1+u+2u^2+4u^3+4u^4+4u^5+4u^6+2u^7+u^8+u^9$
4	$1+u+2u^2+4u^3+8u^4+11u^5+11u^6+14u^7+16u^8+$
	$+14u^9+11u^{10}+11u^{11}+8u^{12}+4u^{13}+2u^{14}+u^{15}+u^{16}$

There is clearly a lot of structure in the polynomials $P_k(u)$ which will be discussed in the next section.

4 The structure of $P_k(u)$

Consider Equation 6 of Theorem 2. We write $P_k(u) = \sum_{i=0}^{k^2} \alpha_i u^i$ and define the *stretch* of $P_k(u)$ to be

$$PP_k(u) = \alpha_0 + \alpha_{k^2} u^{k^2+k} + \sum_{i=0}^{k} \sum_{j=0}^{k-2} \alpha_{1+i+(k+1)j} u^{2+i+(k+1)j+j}.$$

What this does to $P_k(u)$ is to insert 0 coefficients at every $(k+1)$st term, starting after α_0. Thus, the stretched polynomials corresponding to the values of $P_k(u)$ given in the array above are:

k	$PP_k(u)$
0	1
1	$1+u^2$
2	$1+u^2+2u^3+u^4+u^6$
3	$1+u^2+2u^3+4u^4+4u^5+4u^7+4u^8+2u^9+u^{10}+u^{12}$
4	$1+u^2+2u^3+4u^4+8u^5+11u^6+11u^8+14u^9+16u^{10}+$
	$+14u^{11}+11u^{12}+11u^{14}+8u^{15}+4u^{16}+2u^{17}+u^{18}+u^{20}$

Note that if $P_k(u)$ has degree k^2 then $PP_k(u)$ has degree k^2+k.

Theorem 3 *For all $k \geq 1$,*

$$P_{k+1}(u) = PP_k(u) \cdot (1+u+u^2+\cdots+u^{k+1}).$$

Theorem 4 *The coefficients of $P_k(u)$ are symmetric and unimodal.*

Proof: It follows from Theorem 3 that we can construct the coefficient sequence for $P_{k+1}(u)$ from that of $P_k(u)$ by the following rule (where we assume that all coefficients of u^t in $P_k(u)$ are 0 if $t < 0$ or $t > k^2$). Namely, suppose we write $P_k(u) = \sum_{i=0}^{k^2} \alpha_i u^i$ so that we have the coefficient sequence $A_k = (\alpha_0, \alpha_1, \ldots, \alpha_{k^2})$. Now form the new sequence $B_k = (\beta_0, \beta_1, \ldots \beta_{k^2+k})$ by the rule

$$\beta_i = \sum_{j=i-k}^{i} \alpha_j, \quad 0 \leq i \leq k^2+k.$$

Finally, starting with β_0, insert *duplicate* values for the coefficients

$$\beta_0, \beta_{k+1}, \beta_{2(k+1)}, \ldots, \beta_{t(k+1)}, \ldots, \beta_{(k-1)(k+1)} \text{ and } \beta_{k(k+1)}.$$

Thus, this will generate the sequence

$$(\beta_0, \beta_0, \beta_1, \beta_2, \ldots, \beta_k, \beta_{k+1}, \beta_{k+1}, \beta_{k+2}, \ldots, \beta_{k^2+k-1}, \beta_{k^2+k}, \beta_{k^2+k}).$$

This new sequence will in fact just be the coefficient sequence A_{k+1} for $P_{k+1}(u)$. For example, starting with $P_1(u) = 1+u$, we have $A_1 = (1,1)$ and so $B_1 = (1,2,1)$. Now, inserting the duplicate values for $\beta_0 = 1$ and $\beta_2 = 1$, we get the coefficient sequence $A_2 = (1, \mathbf{1}, 2, 1, \mathbf{1})$ for $P_2(u) = 1+u+2u^2+u^3+u^4$. Repeating this process for A_2, we sum blocks of length 3 to get $B_2 = (1,2,4,4,4,2,1)$. Inserting duplicates for entries at positions $0, 3$ and 6 gives us the new coefficient sequence $A_3 = (1, \mathbf{1}, 2, 4, 4, \mathbf{4}, 4, 2, 1, \mathbf{1})$ of $P_3 = 1+u+2u^2+4u^3+4u^4+4u^5+4u^6+2u^7+u^8+u^9$, etc. It is also clear from this procedure that if A_k is symmetric and unimodal, then so is B_k, and consequently, so is A_{k+1}. This is what we claimed. □

4.1 An Eulerian identity

Note that since $P_k(u)$ is symmetric and has degree u^{k^2}, we have $P_k(u) = u^{k^2} P_k(\frac{1}{u})$. Replacing $P_k(u)$ by its expression in (6), we obtain (with some calculation) the interesting identity

$$\sum_{j=0}^{a+b}(-1)^j \binom{a}{j}(1-x)^j A_{a+b-j}(x) = x\sum_{j=0}^{a+b}\binom{b}{j}(1-x)^j A_{a+b-j}(x) + \binom{b}{a+b}(1-x)^{a+b+1}$$

for *all* integers a and b provided that $a+b \geq 0$.

References

[1] J. Buhler, D. Eisenberg, R. Graham and C. Wright, Juggling drops and descents, *Amer. Math. Monthly* **101** (1994), 507–519.

[2] F. Chung, and R. L. Graham, Primitive juggling sequences, *Amer. Math. Monthly* **115** (2008), 185–194.

[3] L. Euler, Methodus universalis series summandi ulterius promota, *Commentarii academiae scientiarum imperialis Petropolitanae* **8** (1736), 147–158. Reprinted in his *Pera Omnia*, series 1, volume 14, 124–137.

[4] D. Foata and G. Han, q-series in Combinatorics; permutation statistics (Lecture Notes), preliminary edition, 2004.

[5] R. L. Graham, D. E. Knuth and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 1994.

[6] D. E. Knuth: *The Art of Computer Programming*, Vol. 1, *Fundamental algorithms*. Addison-Wesley, Reading, 1969.

[7] D. E. Knuth, *The Art of Computer Programming*, Vol. 3, *Sorting and Searching*, Addison-Wesley, Reading, 2nd ed., 1998.

[8] P. A. MacMahon, *Combinatory Analysis*, 2 volumes, Cambridge University Press, London, 1915-1916. Reprinted by Chelsea, New York, 1960.

[9] O. Rodrigues, Note sur les inversions, ou dérangements produits dans les permutations, *J. de Math.* **4** (1839), 236–240.

[10] J. Shareshian and M. L. Wachs, q-Eulerian polynomials: excedance number and major index, *Electron. Res. Announc. Amer. Math. Soc.* **13** (2007), 33–45.

[11] J. Shareshian and M. L. Wachs, Eulerian quasisymmetric functions, preprint 2009.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 138–149

Weighted branching formulas for the hook lengths

Ionuţ Ciocan-Fontanine[1], Matjaž Konvalinka[2] and Igor Pak[3]

[1] *School of Mathematics, University of Minnesota, Minneapolis MN, USA, and School of Mathematics, Korea Institute for Advanced Study, Seoul, Korea*
[2] *Department of Mathematics, Vanderbilt University, Nashville TN, USA*
[3] *Department of Mathematics, UCLA, Los Angeles CA, USA*

Abstract. The famous hook-length formula is a simple consequence of the branching rule for the hook lengths. While the Greene-Nijenhuis-Wilf probabilistic proof is the most famous proof of the rule, it is not completely combinatorial, and a simple bijection was an open problem for a long time. In this extended abstract, we show an elegant bijective argument that proves a stronger, weighted analogue of the branching rule. Variants of the bijection prove seven other interesting formulas. Another important approach to the formulas is via weighted hook walks; we discuss some results in this area. We present another motivation for our work: J-functions of the Hilbert scheme of points.

Résumé. La formule bien connue de la longueur des crochets est une conséquence simple de la règle de branchement des longueurs des crochets. La preuve la plus répandue de cette règle est de nature probabiliste et est due à Greene-Njenhuis-Wilf. Elle n'est toutefois pas complètement combinatoire et une simple bijection a été pendant longtemps un problème ouvert. Dans ce résumé étendu, nous proposons un argument bijectif élégant qui démontre une version à poids plus forte de cette règle. Des variantes de cette bijection permettent d'obtenir sept autres formules intéressantes. Une autre approche importante de ces formules est via les marches des crochets à poids. Nous discutons certains résultats dans cette direction. Enfin, nous présentons aussi une autre motivation à l'origine de ce travail: les J-fonctions du schéma d'Hilbert des points.

Resumen. La famosa fórmula de la longitud de codos es una consecuencia simple de la ley de ramificación de las longitudes de los codos. Mientras que la prueba probabilística de la fórmula de Greene-Nijenhuis-Wilf es la más famosa, ésta no es del todo combinatoria. Por mucho tiempo el problema de encontrar una prueba biyectiva de la formula estuvo abierto. En este resumen extendido, mostramos un argumento biyectivo elegante que prueba una variante ponderada más robusta de la ley de ramificación. Variantes de la biyección prueban otras siete fórmulas interesantes. Otro enfoque importante a las fórmulas es a traves de caminos ponderados de codos: discutimos unos resultados en esta área. Presentamos otra motivación: las J-funciones del esquema de Hilbert de puntos.

Keywords: Hilbert scheme of points, hook-length formula, bijective proofs

1 Introduction and main results

The classical hook-length formula gives an elegant product formula for the number of standard Young tableaux. Since its discovery by Frame, Robinson and Thrall in [9], it has been reproved, generalized and extended in several different ways, and applications have been found in a number of fields of mathematics.

Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$, $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_\ell > 0$, be a partition of n, $\lambda \vdash n$, and let $[\lambda] = \{(i,j) \in \mathbb{Z}^2 \colon 1 \leq i \leq \ell, 1 \leq j \leq \lambda_i\}$ be the corresponding *Young diagram*. The *conjugate partition* λ' is defined by $\lambda'_j = \max\{i : \lambda_i \geq j\}$. The *hook* $H_{\mathbf{z}} \subseteq [\lambda]$ is the set of squares weakly to the right and below of $\mathbf{z} = (i,j) \in [\lambda]$, and the *hook length* $h_{\mathbf{z}} = h_{ij} = |H_{\mathbf{z}}| = \lambda_i + \lambda'_j - i - j + 1$ is the size of the hook.

A *standard Young tableau* of shape λ is a bijective map $f : [\lambda] \to \{1, \ldots, n\}$, such that $f(i_1, j_1) < f(i_2, j_2)$ whenever $i_1 \leq i_2$, $j_1 \leq j_2$, and $(i_1, j_1) \neq (i_2, j_2)$. We denote the number of standard Young tableaux of shape λ by f^λ. The hook-length formula states that if λ is a partition of n, then

$$f^\lambda = \frac{n!}{\prod_{\mathbf{z} \in [\lambda]} h_{\mathbf{z}}}.$$

For example, for $\lambda = (3,2,2) \vdash 7$, the hook-length formula gives $f^{322} = \frac{7!}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 2 \cdot 1 \cdot 1} = 21$.

One way to prove the hook-length formula is by induction on n. Namely, it is obvious that in a standard Young tableau, n must be in one of the *corners*, squares (i,j) of $[\lambda]$ satisfying $(i+1,j), (i,j+1) \notin [\lambda]$. Therefore $f^\lambda = \sum_{\mathbf{c} \in \mathcal{C}[\lambda]} f^{\lambda - \mathbf{c}}$, where $\mathcal{C}[\lambda]$ is the set of all corners of λ, and $\lambda - \mathbf{c}$ is the partition whose diagram is $[\lambda] \setminus \{\mathbf{c}\}$. That means that in order to prove the hook-length formula, we have to prove that $F^\lambda = n!/\prod h_{\mathbf{z}}$ satisfy the same recursion. It is easy to see that this is equivalent to the following *branching rule for the hook lengths*:

$$\sum_{(r,s) \in \mathcal{C}[\lambda]} \frac{1}{n} \prod_{i=1}^{r-1} \frac{h_{is}}{h_{is} - 1} \prod_{j=1}^{s-1} \frac{h_{rj}}{h_{rj} - 1} = 1. \tag{1}$$

In an important development, Green, Nijenhuis and Wilf introduced the *hook walk* which proves (1) by a combination of a probabilistic and a short but delicate induction argument [13]. Zeilberger converted the hook walk proof into a bijective proof [26], but laments on the "enormous size of the input and output" and "the recursive nature of the algorithm" (ibid, §3). With time, several variations of the hook walk have been discovered, most notably the q-version of Kerov [16], and its further generalization, the (q,t)-version of Garsia and Haiman [10]. In the recent paper [7], a direct bijective proof of (1) is presented. In fact, a bijective proof is presented of the following more general identity, called the *weighted branching formula*.

$$\begin{aligned}
&\left[\sum_{(p,q) \in [\lambda]} x_p y_q\right] \cdot \left[\prod_{(i,j) \in [\lambda] \setminus \mathcal{C}[\lambda]} \Big(x_{i+1} + \ldots + x_{\lambda'_j} + y_{j+1} + \ldots + y_{\lambda_i}\Big)\right] \\
&= \sum_{(r,s) \in \mathcal{C}[\lambda]} \left[\prod_{\substack{(i,j) \in [\lambda] \setminus \mathcal{C}[\lambda] \\ i \neq r, j \neq s}} \Big(x_{i+1} + \ldots + x_{\lambda'_j} + y_{j+1} + \ldots + y_{\lambda_i}\Big)\right] \\
&\cdot \left[\prod_{i=1}^{r} (x_i + \ldots + x_r + y_{s+1} + \ldots + y_{\lambda_i})\right] \cdot \left[\prod_{j=1}^{s} \Big(x_{r+1} + \ldots + x_{\lambda'_j} + y_j + \ldots + y_s\Big)\right]
\end{aligned}$$

We refer to this formula as WBR. Here $x_1, \ldots, x_{\ell(\lambda)}, y_1, \ldots, y_{\lambda_1}$ are some commutative variables. If we substitute all x_i and y_j by 1, we get

$$n \cdot \prod_{\mathbf{z}\in[\lambda]\setminus\mathcal{C}[\lambda]} (h_{\mathbf{z}} - 1) \;=\; \sum_{(r,s)\in\mathcal{C}[\lambda]} \left[\prod_{\substack{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]\\ i\neq r, j\neq s}} (h_{\mathbf{z}} - 1) \right] \prod_{i=1}^{r} h_{is} \prod_{j=1}^{s} h_{rj},$$

which is equivalent to (1).

In [18], a weighted analogue of the formula

$$\prod_{\mathbf{z}\in[\lambda]} (h_{\mathbf{z}} + 1) \;=\; \sum_{(r,s)\in\mathcal{C}'[\lambda]} \left[\prod_{\substack{(i,j)\in[\lambda]\\ i\neq r, j\neq s}} (h_{\mathbf{z}} + 1) \right] \prod_{i=1}^{r-1} h_{is} \prod_{j=1}^{s-1} h_{rj} \tag{2}$$

is proved. Here $\mathcal{C}'[\lambda]$ is the set of *outer corners* of λ, squares $(i,j) \notin [\lambda]$ satisfying $i = 1$ or $(i-1,j) \in [\lambda]$, and $j = 1$ or $(i, j-1) \in [\lambda]$. The motivation for this formula is as follows, see [22]. Division by $\prod_{\mathbf{z}\in[\lambda]}(h_{\mathbf{z}} + 1)$ and $\prod_{\mathbf{z}\in[\lambda]} h_{\mathbf{z}}$ yields

$$\frac{1}{\prod_{\mathbf{z}\in[\lambda]} h_{\mathbf{z}}} \;=\; \sum_{(r,s)\in\mathcal{C}'[\lambda]} \prod_{i=1}^{r-1} \frac{1}{h_{is}+1} \prod_{j=1}^{s-1} \frac{1}{h_{rj}+1} \prod_{\substack{(i,j)\in[\lambda]\\ i\neq r, j\neq s}} \frac{1}{h_{\mathbf{z}}}$$

We multiply by $(n+1)!$ and use the hook-length formula. We get $(n+1)f^{\lambda} = \sum_{\mathbf{c}\in\mathcal{C}'[\lambda]} f^{\lambda+\mathbf{c}}$, where $\lambda + \mathbf{c}$ is the partition whose diagram is $[\lambda] \cup \{\mathbf{c}\}$. Let us introduce the notation $\mu \to \lambda$ or $\lambda \leftarrow \mu$ if $\lambda = \mu - \mathbf{c}$ for a corner $\mathbf{c}$ of μ, or, equivalently, if $\mu = \lambda + \mathbf{c}$ for an outer corner $\mathbf{c}$ of λ. We then have

$$\sum_{\mu\vdash n+1} (f^{\mu})^2 = \sum_{\mu\vdash n+1} f^{\mu} \left(\sum_{\lambda\leftarrow\mu} f^{\lambda} \right) = \sum_{\lambda\vdash n} f^{\lambda} \left(\sum_{\mu\to\lambda} f^{\mu} \right) = (n+1) \sum_{\lambda\vdash n} (f^{\lambda})^2.$$

Induction proves the famous formula $\sum_{\lambda\vdash n} (f^{\lambda})^2 = n!$.

It turns out that the correct weighted analogue is

$$\prod_{(i,j)\in[\lambda]} \Big(x_i + \ldots + x_{\lambda'_j} + y_j + \ldots + y_{\lambda_i} \Big) = \sum_{(r,s)\in\mathcal{C}'[\lambda]} \prod_{\substack{(i,j)\in[\lambda]\\ i\neq r, j\neq s}} \Big(x_i + \ldots + x_{\lambda'_j} + y_j + \ldots + y_{\lambda_i} \Big)$$
$$\cdot \left[\prod_{i=1}^{r-1} (x_{i+1} + \ldots + x_{r-1} + y_s + \ldots + y_{\lambda_i}) \right] \cdot \left[\prod_{j=1}^{s-1} \Big(x_r + \ldots + x_{\lambda'_j} + y_{j+1} + \ldots + y_{s-1} \Big) \right].$$

We refer to this result as complementary weighted branching rule, or CWBR.

This extended abstract is organized as follows. In Section 2, we describe the work that led us to WBR. In Section 3, we give bijective proofs of WBR and CWBR. Simple variants of the proofs lead to six other interesting identities. In Section 4, we present new theorems on weighted hook walks, and some recursions for f^{λ} which arise as corollaries. We finish with some final remarks in Section 5.

This extended abstract is based on papers [6], [7] and [18].

2 Motivation: J-functions of the Hilbert scheme of points

In the last fifteen years, deep relations have been uncovered between representation theory and the geometry of the Hilbert scheme of points in the complex affine plane $\mathrm{Hilb}_n(\mathbb{C}^2)$. See, say, Nakajima, [19] and Haiman, [15]. The equivariant quantum cohomology $QH^*_{(\mathbb{C}^*)^2}(\mathrm{Hilb}_n)(\mathbb{C}^2)$ of the Hilbert scheme has been recently determined by Okounkov and Pandharipande, and the authors have also shown that it agrees with the (equivariant) *relative Donaldson-Thomas theory* of $\mathbb{P}^1 \times \mathbb{C}^2$, see [20], [21].

A different perspective on the study of the relationship between $QH^*_{(\mathbb{C}^*)^2}(\mathrm{Hilb}_n(\mathbb{C}^2))$ and DT-theory is undertaken in [3]. The main point there is to exploit the fact that the Hilbert scheme is a Geometric Invariant Theory (GIT) quotient via the celebrated "ADHM construction" of Atiyah-Drinfeld-Hitchin-Manin [1].

On the one hand, this allows one to employ the machinery of the abelian/nonabelian correspondence in Gromov-Witten theory of [5], [2] to analyze the quantum cohomology of $\mathrm{Hilb}_n(\mathbb{C}^2)$. In particular, one can give a formula (a priori conjectural) for the *J-function* of the Hilbert scheme – a certain generating function for Gromov-Witten invariants of a nonsingular algebraic variety, essentially encoding the same information as the quantum cohomology ring. On the other hand, the ADHM construction of $\mathrm{Hilb}_n(\mathbb{C}^2)$ is also highly relevant to the Donaldson-Thomas side of the story, due to work of Diaconescu. Namely, in [8] he used it to obtain a gauge-theoretic partial compactification of the space of maps $\mathbb{P}^1 \to \mathrm{Hilb}_n$, his *moduli space of ADHM sheaves* on $\mathbb{P}^1$, and then provided a direct geometric identification of the DT-theory of $\mathbb{P}^1 \times \mathbb{C}^2$ with the intersection theory of this new moduli space.

The main result of [3], and the jumping board for the paper [6], is a proof of the above-mentioned formula for the J-function of $\mathrm{Hilb}_n(\mathbb{C}^2)$. The following identity involving partitions is obtained as a corollary. Choose a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ and let α, β be indeterminates. For a square $\mathbf{z} = (i,j)$ in $|\lambda|$, we define its *weight* to be $w_{\mathbf{z}} = -(i-1)\alpha - (j-1)\beta$. Then, for each $n \geq 1$ and each partition λ of n, we have

$$\sum_{\mathbf{c}\in\mathcal{C}[\lambda]} (w_{\mathbf{c}} - (\alpha+\beta)) \prod_{\mathbf{z}\in[\lambda]\setminus\{\mathbf{c}\}} \frac{(w_{\mathbf{c}} - w_{\mathbf{z}} - \alpha)(w_{\mathbf{c}} - w_{\mathbf{z}} - \beta)}{(w_{\mathbf{c}} - w_{\mathbf{z}})(w_{\mathbf{c}} - w_{\mathbf{z}} - (\alpha+\beta))} = -n(\alpha+\beta). \tag{3}$$

If λ has ℓ corners, there are ℓ different parts of λ. Let x_ℓ denote the smallest part, $x_{\ell-1} + x_\ell$ the second smallest etc., and $x_1 + x_2 + \ldots + x_\ell$ the largest part. Furthermore, let y_1 be the number of times the largest part appears in λ, y_2 the number of times the second largest part appears, etc. A careful analysis of the cancellations and the substitution of x_i for $x_i\alpha$ and y_i for $y_i\beta$ gives the rational function identity

$$\sum_{k=1}^{\ell} x_k y_k \prod_{p=1}^{k-1} \left(1 + \tfrac{y_p}{x_p+\ldots+x_{k-1}+y_{p+1}+\ldots+y_k}\right) \prod_{q=k+1}^{\ell} \left(1 + \tfrac{x_q}{x_k+\ldots+x_{q-1}+y_{k+1}+\ldots+y_q}\right) = \sum_{1\leq p\leq q\leq \ell} x_q y_p.$$

This is exactly WBR for the staircase shape $(\ell, \ell-1, \ldots, 1)$. See [6] for a more detailed explanation.

3 Bijective proofs of weighted branching formulas

Now we present a bijective proof of WBR, by interpreting both sides as certain sets of arrangements of labels, and then constructing a bijection between two sets of labels.

For the left-hand side of WBR, we are given: special labels x_p, y_q, corresponding to the first summation $\sum_{(p,q)\in[\lambda]} x_p y_q$, and a label x_k for some $i < k \leq \lambda'_j$, or y_l for some $j < l \leq \lambda_i$, in every non-corner square (i,j). Denote by F the resulting arrangement of labels, see Figure 1, left.

We can interpret the special labels x_p, y_q as the starting square (p, q). Furthermore, we can interpret all other labels as arrows: if the label in square (i, j) is x_k, the arrow points to (k, j), and if the label is y_l, the arrow points to (i, l). The arrow from (p, q) points to a square (p', q') in the hook of (p, q), the arrow from (p', q') points to a square (p'', q'') in the hook of (p', q'), etc. Eventually we obtain a *hook walk* which reaches a corner $(r, s) \in \mathcal{C}[\lambda]$ (Figure 1, second drawing). Shade row r and column s. Now we shift the labels in the hook walk and in its projection onto the shaded row and column. If the hook walk has a horizontal step from (i, j) to (i, j'), move the label in (i, j) right and down from (i, j) to (r, j'), and the label from (r, j) up to (i, j). If the hook walk has a vertical step from (i, j) to (i', j), move the label from (i, j) down and right to (i', s), and the label from (i, s) left to (i, j). If the hook walk has a horizontal step from (r, j) to (r, j'), move the label in (r, j) right to (r, j'). If the hook walk has a vertical step from (i, s) to (i', s), move the label in (i, s) down to (i', s). Finally, move the label x_p to (p, s), and the label y_q to (r, q). See Figure 1, third drawing. Denote the resulting arrangement G (Figure 1, right). It turns out that G represents a term on the right-hand side. Furthermore, φ is a bijection.

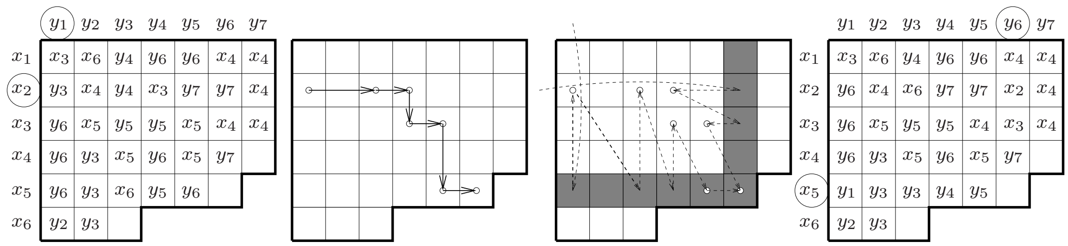


Fig. 1: An arrangement corresponding to the left-hand side of WBR; hook walk; shift of labels; final arrangement.

There are three more identities in the same spirit. To save on space, let us write them down in an abbreviated fashion. If WBR is the identity

$$\left[\sum_{(p,q)\in[\lambda]} x_p y_q\right] \cdot \left[\prod_{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]} *\right] = \sum_{(r,s)\in\mathcal{C}[\lambda]} \left[\prod_{\substack{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]\\ i\neq r, j\neq s}} *\right] \cdot \left[\prod_{i=1}^{r} *\right] \cdot \left[\prod_{j=1}^{s} *\right],$$

then the following identities are also true:

$$\left[\sum_{p=1}^{\ell(\lambda)} x_p\right] \cdot \left[\prod_{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]} *\right] = \sum_{(r,s)\in\mathcal{C}[\lambda]} \left[\prod_{\substack{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]\\ i\neq r, j\neq s}} *\right] \cdot \left[\prod_{i=1}^{r} *\right] \cdot \left[\prod_{j=2}^{s} *\right], \tag{4}$$

$$\left[\sum_{q=1}^{\lambda_1} y_q\right] \cdot \left[\prod_{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]} *\right] = \sum_{(r,s)\in\mathcal{C}[\lambda]} \left[\prod_{\substack{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]\\ i\neq r, j\neq s}} *\right] \cdot \left[\prod_{i=2}^{r} *\right] \cdot \left[\prod_{j=1}^{s} *\right], \tag{5}$$

$$\left[\prod_{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]} *\right] = \sum_{(r,s)\in\mathcal{C}[\lambda]} \left[\prod_{\substack{(i,j)\in[\lambda]\setminus\mathcal{C}[\lambda]\\ i\neq r, j\neq s}} *\right] \cdot \left[\prod_{i=2}^{r} *\right] \cdot \left[\prod_{j=2}^{s} *\right]. \tag{6}$$

The proofs are very similar. We start the hook walk in square $(p,1)$ for (4), $(1,q)$ for (5), and $(1,1)$ for (6). We proceed as in the proof of WBR, except that in the final arrangement, the square $(r,1)$ (respectively, $(1,s)$, respectively, both $(r,1)$ and $(1,s)$) does not get a label.

A direct bijective proof of CWBR shares many characteristics with the bijective proof of WBR. We interpret both left-hand and right-hand sides as labelings of the diagram; we start the bijection with a (variant of the) hook walk; and the hook walk determines a relabeling of the diagram. There are, however, some important differences. First, the walk always starts in the square $(1,1)$. Second, the hook walk can never pass through a square that is not in the same row as an outer corner and the same column as an outer corner. Third, the rule for one step of the hook walk is different from the one in [7]. And finally, there is an extra shift in the relabeling process.

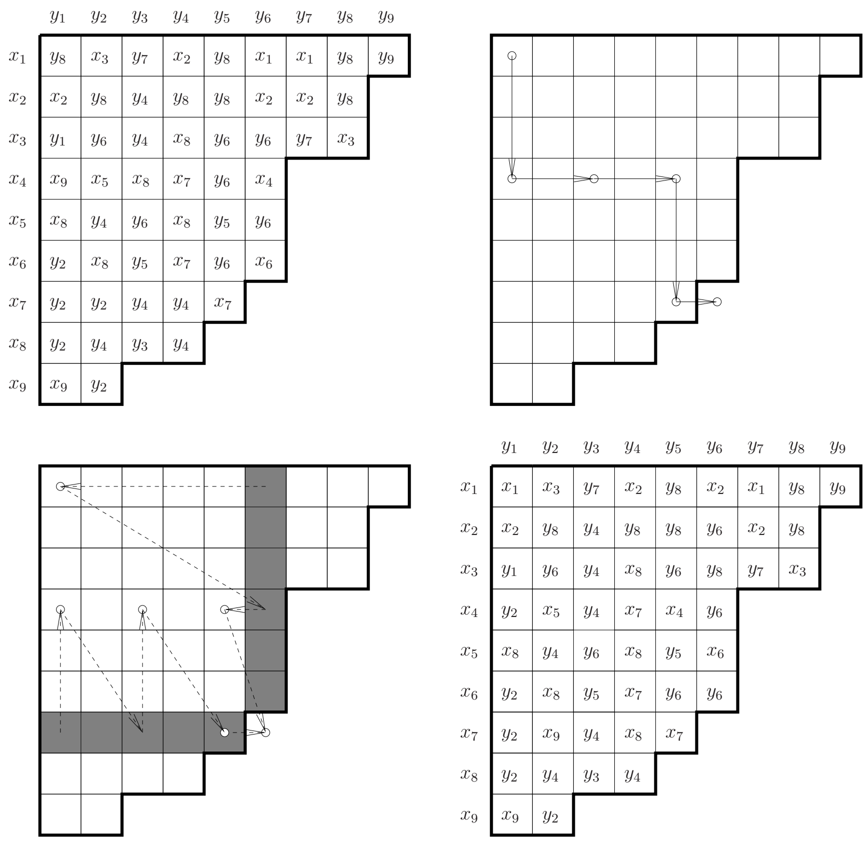


Fig. 2: An example of an arrangement corresponding to the left-hand side of CWBR for $\lambda = 988666542$; hook walk; shift of labels; final arrangement.

For the left-hand side of CWBR, we are given a label x_k for some $i \le k \le \lambda'_j$, or y_l for some $j \le l \le \lambda_i$, for every square $(i,j) \in [\lambda]$. Denote by F the resulting arrangement of n labels (see Figure 2, top left).

Again, we first construct a hook walk. Start in $(1,1)$, and move only through squares which are in the same row as an outer corner and in the same column as an outer corner. The rule is as follows. If the current square is (i,j) and the label of (i,j) in F is x_k for $i \le k \le \lambda'_j$, move to $(i, \lambda_k + 1)$. If the label of (i,j) in F is y_l for $j \le l \le \lambda'_j$, move to $(\lambda'_l + 1, j)$. Note that $i \le k$ implies $\lambda_k \le \lambda_i$ and $j \le l$ implies $\lambda'_l \le \lambda'_j$, so the square we move to is either in $[\lambda]$ or is the outer corner to the right or below (i,j). The process continues until we arrive in an outer corner (r,s), see the top right drawing in Figure 2.

Shade row r and column s. Now we shift the labels in the hook walk and in its projection onto the shaded row and column. If the hook walk has a horizontal step from (i,j) to (i,j'), $i \neq r$, move the label in (i,j) right and down to (r,j'), and the label from (r,j) up to (i,j). If the hook walk has a vertical step from (i,j) to (i',j), $j \neq s$, move the label from (i,j) down and right to (i',s), and the label from (i,s) left to (i,j). If the hook walk has a horizontal step from (r,j) to (r,j'), move the label in (r,j) right to (r,j'). If the hook walk has a vertical step from (i,s) to (i',s), move the label in (i,s) down to (i',s). See Figure 2, bottom left.

After these changes, we have the following situation. If $r = 1$, there is no label in $(1,1)$, and in $(1,s)$ the label is x_k, $1 \le k \le \lambda'_{\lambda_1}$. Move all the labels in row 1 one square to the left. If $s = 1$, there is no label in $(1,1)$, and in $(r,1)$ the label is y_l, $1 \le l \le \lambda_{\ell(\lambda)}$. Move all the labels in column 1 one square up. If $r > 1$ and $s > 1$, there are no labels in $(r,1)$ and $(1,s)$. In (r,s), there are two labels: one of the form x_k for $r \le k \le \lambda'_{s-1}$, and one of the form y_l for $s \le l \le \lambda_{r-1}$. Push all the labels in row r, including x_k in (r,s), one square to the left; and push all labels in column s, including y_l in (r,s), one square up. See Figure 2, bottom right, for the final arrangement, which we denote G. It turns out that the final arrangement represents a term on the right-hand side of CWBR, and the map $F \mapsto G$ is a bijection.

Again, there are variants of the formula with similar bijective proofs. Namely, if CWBR is

$$\prod_{(i,j)\in[\lambda]} * = \sum_{(r,s)\in\mathcal{C}'[\lambda]} \left[\prod_{\substack{(i,j)\in[\lambda]\\ i\neq r, j\neq s}} *\right] \cdot \left[\prod_{i=1}^{r-1} *\right] \cdot \left[\prod_{j=1}^{s-1} *\right],$$

we also have

$$\left[\sum_{p=1}^{\ell(\lambda)} x_p\right] \cdot \left[\prod_{(i,j)\in[\lambda], j\neq 1} *\right] = \sum_{(r,s)\in\mathcal{C}'[\lambda], s\neq 1} \left[\prod_{\substack{(i,j)\in[\lambda]\\ i\neq r, j\neq 1,s}} *\right] \cdot \left[\prod_{i=1}^{r-1} *\right] \cdot \left[\prod_{j=1}^{s-1} *\right], \tag{7}$$

$$\left[\sum_{q=1}^{\lambda_1} y_q\right] \cdot \left[\prod_{(i,j)\in[\lambda], i\neq 1} *\right] = \sum_{(r,s)\in\mathcal{C}'[\lambda], r\neq 1} \left[\prod_{\substack{(i,j)\in[\lambda]\\ i\neq 1,r, j\neq s}} *\right] \cdot \left[\prod_{i=1}^{r-1} *\right] \cdot \left[\prod_{j=1}^{s-1} *\right], \tag{8}$$

$$\left[\sum_{(p,q)\notin[\lambda]} x_p y_q\right] \cdot \left[\prod_{(i,j)\in[\lambda], i,j\neq 1} *\right] = \sum_{(r,s)\in\mathcal{C}'[\lambda], r,s\neq 1} \left[\prod_{\substack{(i,j)\in[\lambda]\\ i\neq 1,r, j\neq 1,s}} *\right] \cdot \left[\prod_{i=1}^{r-1} *\right] \cdot \left[\prod_{j=1}^{s-1} *\right]. \tag{9}$$

The sum on the left-hand side of (9) is over all (i,j) such that $1 \le i \le \ell(\lambda)$, $1 \le j \le \lambda_1$, $(i,j) \notin [\lambda]$. The proofs of these identities are almost identical to the one for CWBR. We start the hook walk in square $(1, \lambda_p + 1)$ (respectively, in $(\lambda'_q + 1, 1)$, respectively, in $(\lambda'_q + 1, \lambda_p + 1)$); we construct the hook walk in exactly the same fashion as before; we perform the relabeling as before; but before the final shift to the left and up by one, we label $(r, \lambda_p + 1)$ (respectively, $(\lambda'_q + 1, s)$, respectively, both) with x_p (respectively, y_q, respectively, both).

4 Weighted hook walks

Choose a partition λ and draw the borders of its diagram in the plane. Now add lines $x = 0$, $x = \ell(\lambda)$, $y = 0$, $y = \lambda_1$; this divides the plane into ten regions $R_1, \ldots, R_{10}$ (R_5 is empty if $\lambda = a^b$ for some a and b). See Figure 3 for an example and the labelings of these regions. Draw the following lines in bold: the half-line $x = 0$, $y \geq \lambda_1$, the half-line $x = \ell(\lambda)$, $y \leq 0$, the half-line $y = 0$, $x \geq \ell(\lambda)$, the half-line $y = \lambda_1$, $x \leq 0$, and the zigzag line separating regions R_1 and R_5.

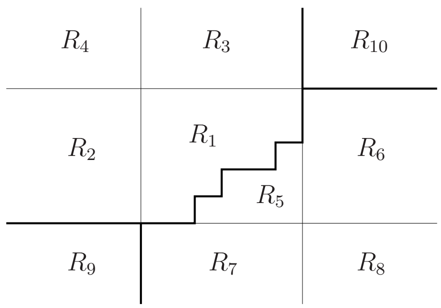


Fig. 3: Division of the plane into regions $R_1, \ldots, R_{10}$ for $\lambda = 66532$, with some lines in bold.

Define a *weighted hook walk* as follows. Choose positive weights $(x_i)_{i=-\infty}^{\infty}$, $(y_j)_{j=-\infty}^{\infty}$ satisfying $\sum_i x_i < \infty$, $\sum_j y_j < \infty$. Select the starting square for the hook walk so that the probability of selecting the square (i, j) is proportional to $x_i y_j$. In each step, move in a vertical or horizontal direction toward the bolded line; in regions R_1, R_2, R_3 and R_4, right or down; in regions R_5, R_6, R_7 and R_8, left or up; in region R_9, right or up; and in region R_{10}, left or down. More specifically, if the current position is (i, j), move to the square (i', j) between (i, j) and the bolded line with probability proportional to $x_{i'}$, and to the square (i, j') between (i, j) and the bolded line with probability proportional to $y_{j'}$. The process stops if we are either in one of the corners of λ (if the initial square was in regions R_1, R_2, R_3 or R_4), one of the outer corners of λ (if the initial square was in regions R_5, R_6, R_7 or R_8), the square $(\ell(\lambda) + 1, 0)$ (if the initial square was in region R_9) or $(0, \lambda_1 + 1)$ (if the initial square was in region R_{10}). These last two possibilities are not particularly interesting.

Below, we give the probabilities of terminating in a particular corner conditional on starting in R_1, R_2, R_3 and R_4, as well as probabilities of ending in a particular outer corner, conditional on starting in R_5, R_6, R_7 and R_8. The most interesting observation is that these probabilities turn out to depend only on $x_1, \ldots, x_{\ell(\lambda)}, y_1, \ldots, y_{\lambda_1}$. As a corollary, we obtain the conditional probabilities in the case where all these values are equal. They represent generalizations of classical results due to Greene, Nijenhuis and Wilf from [13], [14].

We extend the definition of λ_i, λ'_j to all $i, j \in \mathbb{Z}$ in a natural way as follows: for $i \leq 0$, $\lambda_i = \lambda_1$; for $i \geq \ell(\lambda) + 1$, $\lambda_i = 0$; for $j \leq 0$, $\lambda'_j = \ell(\lambda)$; for $j \geq \lambda_1 + 1$, $\lambda'_j = 0$. The following two theorems tell us how to compute probabilities of ending in corners and outer corners.

Theorem 1 *For a corner* $\mathbf{c} = (r,s)$ *of* λ*, denote by* $P(\mathbf{c}|R)$ *the probability that the weighted hook walk terminates in* $\mathbf{c}$*, conditional on the starting point being in* R*. Write*

$$\prod_{rs} = x_r y_s \prod_{i=1}^{r-1}\left(1+\frac{x_i}{x_{i+1}+\ldots+x_r+y_{s+1}+\ldots+y_{\lambda_i}}\right)\cdot\prod_{j=1}^{s-1}\left(1+\frac{y_j}{x_{r+1}+\ldots+x_{\lambda'_j}+y_{j+1}+\ldots+y_s}\right).$$

Then:

(a) $P(\mathbf{c}|R_1) = \frac{1}{\sum_{(p,q)\in[\lambda]} x_p y_q}\cdot\prod_{rs}$

(b) $P(\mathbf{c}|R_2) = \frac{1}{\left(\sum_{p=1}^{\ell(\lambda)} x_p\right)(x_{r+1}+\ldots+x_{\ell(\lambda)}+y_1+\ldots+y_s)}\cdot\prod_{rs}$

(c) $P(\mathbf{c}|R_3) = \frac{1}{\left(\sum_{q=1}^{\lambda_1} y_q\right)(x_1+\ldots+x_r+y_{s+1}+\ldots+y_{\lambda_1})}\cdot\prod_{rs}$

(d) $P(\mathbf{c}|R_4) = \frac{1}{(x_{r+1}+\ldots+x_{\ell(\lambda)}+y_1+\ldots+y_s)(x_1+\ldots+x_r+y_{s+1}+\ldots+y_{\lambda_1})}\cdot\prod_{rs}$

In particular, the sum of each of the above terms over all corners of λ *equals* 1*; note that this proves WBR,* (4), (5) *and* (6). *Also,*

(e) $P(\mathbf{c}) = \frac{1}{\left(\sum_p x_p\right)\cdot\left(\sum_q y_q\right)}\cdot\left(1+\frac{\sum_{p\le 0} x_p}{x_1+\ldots+x_r+y_{s+1}+\ldots+y_{\lambda_1}}\right)\cdot\left(1+\frac{\sum_{q\le 0} y_q}{x_{r+1}+\ldots+x_{\ell(\lambda)}+y_1+\ldots+y_s}\right)\cdot\prod_{rs}.$

Theorem 2 *For an outer corner,* $\mathbf{c} = (r,s)$ *of* λ*, denote by* $P(\mathbf{c}|R)$ *the probability that the weighted hook walk terminates in* $\mathbf{c}$*, conditional on the starting point being in* R*. Write*

$$\prod'_{rs} = \prod_{i=1}^{r-1}\left(1-\frac{x_i}{x_i+\ldots+x_{r-1}+y_s+\ldots+y_{\lambda_i}}\right)\cdot\prod_{j=1}^{s-1}\left(1-\frac{y_j}{x_r+\ldots+x_{\lambda'_j}+y_j+\ldots+y_{s-1}}\right).$$

Then:

(a) $P(\mathbf{c}|R_5) = \frac{(x_r+\ldots+x_{\ell(\lambda)}+y_1+\ldots+y_{s-1})(x_1+\ldots+x_{r-1}+y_s+\ldots+y_{\lambda_1})}{\sum_{(p,q)\notin[\lambda]} x_p y_q}\cdot\prod'_{rs}$

(b) $P(\mathbf{c}|R_6) = \frac{x_r+\ldots+x_{\ell(\lambda)}+y_1+\ldots+y_{s-1}}{\sum_{i=1}^{\ell(\lambda)} x_p}\cdot\prod'_{rs}$

(c) $P(\mathbf{c}|R_7) = \frac{x_1+\ldots+x_{r-1}+y_s+\ldots+y_{\lambda_1}}{\sum_{q=1}^{\lambda_1} y_q}\cdot\prod'_{rs}$

(d) $P(\mathbf{c}|R_8) = \prod'_{rs}$

In particular, the sum of each of the above terms over all outer corners of λ *equals* 1*; note that this proves CWBR,* (7), (8) *and* (9). *Also,*

(e) $P(\mathbf{c}) = \frac{(x_1+\ldots+x_{r-1}+\sum_{q=s}^{\infty} y_q)\cdot(\sum_{p=r}^{\infty} x_p+y_1+\ldots+y_{s-1})}{\left(\sum_p x_p\right)\cdot\left(\sum_q y_q\right)}\cdot\prod'_{rs}.$

Corollary 3 *If $x_1 = \ldots = x_{\ell(\lambda)} = y_1 = \ldots = y_{\lambda_1}$, then we have the following. For a corner $\mathbf{c} = (r, s)$ of λ,*

$$P(\mathbf{c}|R_1) = \frac{f^{\lambda-\mathbf{c}}}{f^\lambda}, \qquad P(\mathbf{c}|R_2) = \frac{nf^{\lambda-\mathbf{c}}}{\ell(\lambda)(\ell(\lambda)-r+s)f^\lambda}$$
$$P(\mathbf{c}|R_3) = \frac{nf^{\lambda-\mathbf{c}}}{\lambda_1(\lambda_1+r-s)f^\lambda}, \qquad P(\mathbf{c}|R_4) = \frac{nf^{\lambda-\mathbf{c}}}{(\ell(\lambda)-r+s)(\lambda_1+r-s)f^\lambda}$$

In particular, the sum of each of the above terms over all corners of λ equals 1.
For an outer corner, $\mathbf{c} = (r, s)$ of λ,

$$P(\mathbf{c}|R_5) = \frac{(\ell(\lambda)-r+s)(\lambda_1+r-s)f^{\lambda+\mathbf{c}}}{(n+1)(\ell(\lambda)\lambda_1-n)f^\lambda}, \qquad P(\mathbf{c}|R_6) = \frac{(\ell(\lambda)-r+s)f^{\lambda+\mathbf{c}}}{(n+1)\ell(\lambda)f^\lambda}$$
$$P(\mathbf{c}|R_7) = \frac{(\lambda_1+r-s)f^{\lambda+\mathbf{c}}}{(n+1)\lambda_1 f^\lambda}, \qquad P(\mathbf{c}|R_8) = \frac{f^{\lambda+\mathbf{c}}}{(n+1)f^\lambda}$$

In particular, the sum of each of the above terms over all outer corners of λ equals 1.

The corollary (for probabilities conditional on starting in $R_2, R_3, \ldots, R_7$) gives six new recursive formulas for numbers of standard Young tableaux. The sums over outer corners have the following interesting interpretation. Recall that the content of a square (i, j) of a diagram $[\lambda]$ is defined as $i - j$.

Corollary 4 *Fix a partition $\lambda \vdash n$. Choose a standard Young tableau of shape λ uniformly at random, and an integer i, $1 \le i \le n+1$ uniformly at random. In the standard Young tableau, increase all integers $\ge i$ by 1, and use the bumping process of the Robinson-Schensted algorithm to insert i in the tableau. Define the random variable X as the content of the square that is added to λ. Then*

$$E(X) = 0, \qquad \mathrm{var}(X) = n.$$

Proof: The bumping process is a bijection $\mathrm{SYT}(\lambda) \times \{1, 2, \ldots, n+1\} \longrightarrow \bigcup_{\mathbf{c}\in\mathcal{C}'[\lambda]} \mathrm{SYT}(\lambda+\mathbf{c})$. This means that the probability that $\mathbf{c}$ is the square added to λ is equal to $\frac{f^{\lambda+\mathbf{c}}}{(n+1)f^\lambda}$. We have

$$(n+1)\lambda_1 f^\lambda = \sum(\lambda_1+r-s)f^{\lambda+\mathbf{c}} = \lambda_1\sum f^{\lambda+\mathbf{c}} + \sum(r-s)f^{\lambda+\mathbf{c}} = (n+1)\lambda_1 f^\lambda + \sum(r-s)f^{\lambda+\mathbf{c}}$$

and therefore $\sum(r-s)f^{\lambda+\mathbf{c}} = 0$, which is equivalent to $E(X) = 0$. On the other hand, we know that

$$(n+1)(\ell(\lambda)\lambda_1 - n)f^\lambda = \sum(\ell(\lambda)-r+s)(\lambda_1+r-s)f^{\lambda+\mathbf{c}} =$$

$$= \ell(\lambda)\lambda_1\sum f^{\lambda+\mathbf{c}} + (\ell(\lambda)-\lambda_1)\sum(r-s)f^{\lambda+\mathbf{c}} - \sum(r-s)^2 f^{\lambda+\mathbf{c}}$$

and so $\sum(r-s)^2 f^{\lambda+\mathbf{c}} = (n+1)nf^\lambda$. Division by $(n+1)f^\lambda$ shows that $\mathrm{var}(X) = n$. $\square$

5 Final remarks

As Knuth wrote in 1973, "Since the hook-lengths formula is such a simple result, it deserves a simple proof ..." (see p. 63 of the first edition of [17], cited also in [26]). Unfortunately, the desired simple proofs have been sorely lacking. It is our hope that Section 3 can be viewed as one such proof.

Surveying the history of the hook length formula is a difficult task, even if one is restricted to purely combinatorial proofs. This extended abstract is too short to even attempt such an endeavor. See [7, §6] for a brief outline, and the references therein.

There are several directions in which our results can be potentially extended. First, it would be interesting to obtain the analogues of our results for *shifted* Young diagrams and Young tableaux, for which there is a analogue of the hook length formula due to Thrall [25] (see also [23]). Similarly, most hook formula results easily extend to trees, and one can try to obtain a weighted analogue in this case as well. However, we are less confident this approach will give new and interesting (or at least non-trivial) formulas. Extending to semi-standard and skew tableaux is another possibility, in which case one would be looking for a weighted analogue of Stanley's hook-content formula [24]. Finally, let us mention several new extensions of the hook length formula recently introduced by Guo-Niu Han in [11, 12].

Acknowledgements

The authors are grateful to Dennis Stanton for pointing out [4] to us and explaining its inner working. MK would like to thank Paul Edelman for several helpful comments on an early draft of [7]. IP would like to thank Persi Diaconis for teaching him Kerov's "segment splitting" algorithm. IP was partially supported by the NSF and the NSA.

References

[1] M. Atiyah, V. Drinfeld, N. Hitchin, and Y. Manin, *Construction of instantons*, Phys. Lett. A **65** (1978), no. 3, 185–187.

[2] A. Bertram, I. Ciocan-Fontanine, and B. Kim, *Gromov-Witten invariants for nonabelian and abelian quotients*, J. of Algebraic Geometry, **17** (2008), 275–294.

[3] I. Ciocan-Fontanine, D. Diaconescu, B. Kim, and D. Maulik, in preparation, 2009.

[4] K. Carde, J. Loubert, A. Potechin and A. Sanborn, Proof of Han's hook expansion conjecture, arXiv:0808.0928.

[5] I. Ciocan-Fontanine, B. Kim, and C. Sabbah, *The abelian/nonabelian correspondence and Frobenius manifolds*, Invent. Math., **171** (2008), 301–343.

[6] (I. Ciocan-Fontanine, M. Konvalinka and I. Pak) *On an identity related to the quantum cohomology of* $\mathrm{Hilb}_n(\mathbb{C}^2)$, preprint (2009)

[7] (I. Ciocan-Fontanine, M. Konvalinka and I. Pak) *The weighted hook length formula*, preprint (2009)

[8] D.-E. Diaconescu, *Moduli of ADHM sheaves and local Donaldson-Thomas theory*, arXiv:0801.0820.

[9] J. S. Frame, G. de B. Robinson and R. M. Thrall, The hook graphs of the symmetric group, *Canad. J. Math.* **6** (1954), 316–325.

[10] A. M. Garsia and M. Haiman, A random q, t-hook walk and a sum of Pieri coefficients, *J. Combin. Theory, Ser. A* **82** (1998), 74–111.

[11] Guo-Niu Han, The Nekrasov-Okounkov hook length formula: refinement, elementary proof, extension and applications, arXiv:0805.1398.

[12] Guo-Niu Han, Discovering hook length formulas by an expansion technique, *Electronic J. Combin.* **15** (2008), RP 133, 41 pp.

[13] C. Greene, A. Nijenhuis and H. S. Wilf, A probabilistic proof of a formula for the number of Young tableaux of a given shape, *Adv. in Math.* **31** (1979), 104–109.

[14] C. Greene, A. Nijenhuis and H. S. Wilf, Another probabilistic method in the theory of Young tableaux, *J. Combin. Theory, Ser. A* **37** (1984), 127–135.

[15] M. Haiman, *Hilbert schemes, polygraphs and the Macdonald positivity conjecture.*, J. Amer. Math. Soc. **14** (2001), no. 4, 941–1006.

[16] S. Kerov, A q-analog of the hook walk algorithm for random Young tableaux, *J. Algebraic Combin.* **2** (1993), 383–396.

[17] D. E. Knuth, *The Art of Computer Programming* (Second edition), Vol. 3, Addison-Wesley, Reading, Massachusetts, 1998.

[18] M. Konvalinka, The weighted hook-length formula II: Complementary formulas, preprint (2009)

[19] H. Nakajima, *Heisenberg algebra and Hilbert schemes of points on projective surfaces*, Ann. of Math. (2) **145** (1997), no. 2, 379–388.

[20] A. Okounkov and R. Pandharipande, *Quantum cohomology of the Hilbert scheme of points in the plane*, math.AG/0411210, 2004.

[21] A. Okounkov and R. Pandharipande, *The local Donaldson-Thomas theory of curves*, arXiv:math/0512573.

[22] D. E. Rutherford, On the relations between the numbers of standard tableaux, *Proc. Edinburgh Math. Soc.* **7** (1942), 51–54.

[23] B. Sagan, *The symmetric group* (Second edition), Springer, New York, 2001.

[24] R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, 2, Cambridge University Press, 1997, 1999.

[25] R. M. Thrall, A combinatorial problem, *Michigan Math. J.* **1** (1952), 81–88.

[26] D. Zeilberger, A short hook-lengths bijection inspired by the Greene–Nijenhuis–Wilf proof, *Discrete Math.* **51** (1984), 101–108.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 150–161

Valuative invariants for polymatroids

Harm Derksen[1†] and Alex Fink[2]

[1] *University of Michigan*
[2] *University of California, Berkeley*

Abstract. Many important invariants for matroids and polymatroids, such as the Tutte polynomial, the Billera-Jia-Reiner quasi-symmetric function, and the invariant $\mathcal{G}$ introduced by the first author, are valuative. In this paper we construct the $\mathbb{Z}$-modules of all $\mathbb{Z}$-valued valuative functions for labeled matroids and polymatroids on a fixed ground set, and their unlabeled counterparts, the $\mathbb{Z}$-modules of valuative invariants. We give explicit bases for these modules and for their dual modules generated by indicator functions of polytopes, and explicit formulas for their ranks. Our results confirm a conjecture of the first author that $\mathcal{G}$ is universal for valuative invariants.

Résumé. Beaucoup des invariants importants des matroïdes et polymatroïdes, tels que le polynôme de Tutte, la fonction quasi-symmetrique de Billera-Jia-Reiner, et l'invariant $\mathcal{G}$ introduit par le premier auteur, sont valuatifs. Dans cet article nous construisons les $\mathbb{Z}$-modules de fonctions valuatives aux valeurs entières des matroïdes et polymatroïdes étiquetés définis sur un ensemble fixe, et leurs équivalents pas étiquetés, les $\mathbb{Z}$-modules des invariants valuatifs. Nous fournissons des bases des ces modules et leurs modules duels, engendrés par fonctions charactéristiques des polytopes, et des formules explicites donnants leurs rangs. Nos résultats confirment une conjecture du premier auteur, que $\mathcal{G}$ soit universel pour les invariants valuatifs.

Keywords: polymatroids, polymatroid polytopes, decompositions, valuations

1 Introduction

Matroids were introduced by Whitney in 1935 (see [22]) as a combinatorial abstraction of linear dependence of vectors in a vector space. Some standard references are [21] and [17]. Polymatroids are multiset analogs of matroids and appeared in the late 1960s (see [9, 13]). There are many distinct but equivalent definitions of matroids and polymatroids, for example in terms of bases, independent sets, flats, polytopes or rank functions. For polymatroids, the equivalence between the various definitions is given in [13]. We will stick to the definition in terms of rank functions:

Definition 1.1 *Suppose that X is a finite set (the* ground set*) and* $\mathrm{rk} : 2^X \to \mathbb{N} = \{0, 1, 2, \dots\}$*, where* 2^X *is the set of subsets of* X*. Then* (X, rk) *is called a* polymatroid *if:*

1. $\mathrm{rk}(\emptyset) = 0$*;*
2. rk *is* weakly increasing*: if* $A \subseteq B$ *then* $\mathrm{rk}(A) \leq \mathrm{rk}(B)$*;*

† Supported by NSF grants DMS 0349019 and DMS 0901298.
() The full version of this paper is arXiv:0908.2988.

3. rk *is* submodular*:* $\mathrm{rk}(A \cup B) + \mathrm{rk}(A \cap B) \leq \mathrm{rk}(A) + \mathrm{rk}(B)$ *for all* $A, B \subseteq X$.

If moreover $\mathrm{rk}(\{x\}) \leq 1$ *for every* $x \in X$*, then* (X, rk) *is called a* matroid*.*

An *isomorphism* $\varphi : (X, \mathrm{rk}_X) \to (Y, \mathrm{rk}_Y)$ is a bijection $\varphi : X \to Y$ such that $\mathrm{rk}_Y \circ \varphi = \mathrm{rk}_X$. Every polymatroid is isomorphic to a polymatroid with ground set $\underline{d} = \{1, 2, \ldots, d\}$ for some nonnegative integer d. The rank of a polymatroid (X, rk) is $\mathrm{rk}(X)$.

Our matroid notations will receive the subscript $_{\mathrm{M}}$, and our polymatroid notations the subscript $_{\mathrm{PM}}$. We will write $_{\mathrm{(P)M}}(d, r)$ when we want to refer to both in parallel.

Let $S_{\mathrm{(P)M}}(d, r)$ be the set of all (poly)matroids with ground set $\underline{d}$ of rank r, A function f on $S_{\mathrm{(P)M}}(d, r)$ is a (poly)matroid *invariant* if $f\big((\underline{d}, \mathrm{rk})\big) = f\big((\underline{d}, \mathrm{rk}')\big)$ whenever $(\underline{d}, \mathrm{rk})$ and $(\underline{d}, \mathrm{rk}')$ are isomorphic. Let $S^{\mathrm{sym}}_{\mathrm{(P)M}}(d, r)$ be the set of isomorphism classes in $S_{\mathrm{(P)M}}(d, r)$. Invariant functions on $S_{\mathrm{(P)M}}(d, r)$ correspond to functions on $S^{\mathrm{sym}}_{\mathrm{(P)M}}(d, r)$. Let $Z_{\mathrm{(P)M}}(d, r)$ and $Z^{\mathrm{sym}}_{\mathrm{(P)M}}(d, r)$ be the $\mathbb{Z}$-modules freely generated by the symbols $\langle \mathrm{rk} \rangle$ for rk in $S_{\mathrm{(P)M}}(d, r)$ and $S^{\mathrm{sym}}_{\mathrm{(P)M}}(d, r)$ respectively. For an abelian group A, every function $f : S^{\mathrm{(sym)}}_{\mathrm{(P)M}}(d, r) \to A$ extends uniquely to a group homomorphism $Z^{\mathrm{(sym)}}_{\mathrm{(P)M}}(d, r) \to A$.

To a (poly)matroid $(\underline{d}, \mathrm{rk})$ one can associate its *base polytope* $Q(\mathrm{rk})$ in $\mathbb{R}^d$ (see Definition 2.2). For $d \geq 1$, the dimension of this polytope is $\leq d - 1$. The indicator function of a polytope $\Pi \subseteq \mathbb{R}^d$ is denoted by $[\Pi] : \mathbb{R}^d \to \mathbb{Z}$. Let $P_{\mathrm{(P)M}}(d, r)$ be the $\mathbb{Z}$-module generated by all $[Q(\mathrm{rk})]$ with $(\underline{d}, \mathrm{rk}) \in S_{\mathrm{(P)M}}(d, r)$. We also define an analogue $P^{\mathrm{sym}}_{\mathrm{P(M)}}(d, r)$ by a certain pushout (see Section 6).

Definition 1.2 *Suppose that* A *is an abelian group. A function* $f : S_{\mathrm{(P)M}}(d, r) \to A$ *is* strongly valuative *if there exists a group homomorphism* $\widehat{f} : P_{\mathrm{(P)M}}(d, r) \to A$ *such that for all* $(\underline{d}, \mathrm{rk}) \in S_{\mathrm{(P)M}}(d, r)$*,*

$$f\big((\underline{d}, \mathrm{rk})\big) = \widehat{f}([Q(\mathrm{rk})]).$$

Many interesting functions on matroids are valuative. Among these is the Tutte polynomial, one of the most important matroid invariants [5, 7]. Other valuative functions on matroids include the quasi-symmetric function $\mathcal{F}$ for matroids of Billera, Jia and Reiner introduced in [3], and the first author's quasi-symmetric function $\mathcal{G}$ introduced in [8]. Speyer's invariant defined in [19] using K-theory is strictly speaking not valuative, but its composition with a certain automorphism of $Z^{\mathrm{sym}}_{*\mathrm{M}}(d, r)$ is valuative. Valuative invariants and additive invariants can be useful for deciding whether a given matroid polytope has a decomposition into smaller matroid polytopes (see the discussion in [3, Section 7]). Matroid polytope decompositions appeared in the work of Lafforgue ([14, 15]) on compactifications of a fine Schubert cell in the Grassmannian associated to a matroid.

It follows from Definition 1.2 that the dual $P_{\mathrm{(P)M}}(d, r)^\vee = \mathrm{Hom}_{\mathbb{Z}}(P_{\mathrm{(P)M}}(d, r), \mathbb{Z})$ is the space of all $\mathbb{Z}$-valued valuative functions on $S_{\mathrm{(P)M}}(d, r)$. Likewise $P^{\mathrm{sym}}_{\mathrm{(P)M}}(d, r)^\vee$ is the space of all $\mathbb{Z}$-valued valuative invariants. Let $p^{\mathrm{(sym)}}_{\mathrm{(P)M}}(d, r)$ be the rank of $P^{\mathrm{(sym)}}_{\mathrm{(P)M}}(d, r)$.

We will give explicit bases for each of the spaces $P_{\mathrm{(P)M}}(d, r)$ and $P^{\mathrm{sym}}_{\mathrm{(P)M}}(d, r)$ and their duals (see Theorems 5.1,6.1, Corollaries 5.3, 6.2). From these we obtain the following formulas:

Theorem 1.3

a. $p^{\mathrm{sym}}_{\mathrm{M}}(d, r) = \binom{d}{r}$ *and* $\displaystyle\sum_{0 \leq r \leq d} p^{\mathrm{sym}}_{\mathrm{M}}(d, r) x^{d-r} y^r = \frac{1}{1 - x - y}$.

b. $p^{\mathrm{sym}}_{\mathrm{PM}}(d,r) = \begin{cases} \binom{r+d-1}{r} & \text{if } d \geq 1 \text{ or } r \geq 1; \\ 1 & \text{if } d = r = 0 \end{cases}$ *and* $\displaystyle\sum_{r=0}^{\infty}\sum_{d=0}^{\infty} p^{\mathrm{sym}}_{\mathrm{PM}}(d,r)x^d y^r = \frac{1-x}{1-x-y}.$

c. $\displaystyle\sum_{0 \leq r \leq d} \frac{p_{\mathrm{M}}(d,r)}{d!} x^{d-r} y^r = \frac{x-y}{xe^{-x} - ye^{-y}}.$

d. $p_{\mathrm{PM}}(d,r) = \begin{cases} (r+1)^d - r^d & \text{if } d \geq 1 \text{ or } r \geq 1; \\ 1 & \text{if } d = r = 0, \end{cases}$ *and* $\displaystyle\sum_{d=0}^{\infty}\sum_{r=0}^{\infty} \frac{p_{\mathrm{PM}}(d,r)x^d y^r}{d!} = \frac{e^x(1-y)}{1-ye^x}.$

In particular, the following theorem, which is a corollary of Corollary 6.3, proves a conjecture of the first author in [8]:

Theorem 1.4 *The invariant $\mathcal{G}$ is universal for all valuative (poly)matroid invariants, i.e., the coefficients of $\mathcal{G}$ span the vector space of all valuative (poly)matroid invariants with values in $\mathbb{Q}$.*

Definition 1.5 *Suppose that $d > 0$. A valuative function $f : S_{(\mathrm{P})\mathrm{M}}(d,r) \to A$ is said to be* additive*, if $f\big((\underline{d}, \mathrm{rk})\big) = 0$ whenever the dimension of $Q(\mathrm{rk})$ is $< d-1$.*

In Sections 8 and 9 we construct bigraded modules $T_{(\mathrm{P})\mathrm{M}}$ and $T^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}$ such that $T_{(\mathrm{P})\mathrm{M}}(d,r)^\vee$ is the space of all additive functions on $S_{(\mathrm{P})\mathrm{M}}(d,r)$ and $T^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r)^\vee$ is the space of all additive invariants. Let $t_{(\mathrm{P})\mathrm{M}}(d,r)$ be the rank of $T_{(P)M}(d,r)$ and $t^{\mathrm{sym}}_{(P)M}(d,r)$ be the rank of $T^{\mathrm{sym}}_{(P)M}(d,r)$. Then we have the following formulas:

Theorem 1.6

a. $\displaystyle\prod_{0 \leq r \leq d} (1 - x^{d-r}y^r)^{t^{\mathrm{sym}}_{\mathrm{M}}(d,r)} = 1 - x - y.$

b. $\displaystyle\prod_{r,d} (1 - x^d y^r)^{t^{\mathrm{sym}}_{\mathrm{PM}}(d,r)} = \frac{1-x-y}{1-y}.$

c. $\displaystyle\sum_{r,d} \frac{t_{\mathrm{M}}(d,r)}{d!} x^{d-r} y^r = \log\left(\frac{x-y}{xe^{-x} - ye^{-y}}\right).$

d. $t_{\mathrm{PM}}(d,r) = \begin{cases} r^{d-1} & \text{if } d \geq 1 \\ 0 & \text{if } d = 0, \end{cases}$ *and* $\displaystyle\sum_{r,d} \frac{t_{\mathrm{PM}}(d,r)}{d!} x^d y^r = \log\big(\frac{e^x(1-y)}{1-ye^x}\big).$

2 Polymatroids and their polytopes

By a *polyhedron* we will mean a finite intersection of closed half-spaces. A *polytope* is a bounded polyhedron. It is convenient to have a polyhedral analogue of polymatroid polytopes, so we make the following definition.

Definition 2.1 *A function $2^X \to \mathbb{Z} \cup \{\infty\}$ is called a* megamatroid *if it has the following properties:*

1. $\mathrm{rk}(\emptyset) = 0$*;*
2. $\mathrm{rk}(X) \in \mathbb{Z}$*;*

3. rk *is* submodular*: if* $\mathrm{rk}(A), \mathrm{rk}(B) \in \mathbb{Z}$*, then* $\mathrm{rk}(A \cup B), \mathrm{rk}(A \cap B) \in \mathbb{Z}$ *and* $\mathrm{rk}(A \cup B) + \mathrm{rk}(A \cap B) \leq \mathrm{rk}(A) + \mathrm{rk}(B)$.

Obviously, every matroid is a polymatroid, and every polymatroid is a megamatroid. The *rank* of a megamatroid (X, rk) is the integer $\mathrm{rk}(X)$. We will use notations for megamatroids analogous to our notations for (poly)matroids but with the subscript $_{\mathrm{MM}}$. We will use the subscript $_{*\mathrm{M}}$ and say "$*$matroid" when we want to refer to megamatroids or polymatroids or matroids in parallel.

Definition 2.2 *For a megamatroid* $(\underline{d}, \mathrm{rk})$*, we define its* base polyhedron $Q(\mathrm{rk})$ *as the set of all* $(y_1, \ldots, y_d) \in \mathbb{R}^d$ *such that* $y_1 + y_2 + \cdots + y_d = \mathrm{rk}(X)$ *and* $\sum_{i \in A} y_i \leq \mathrm{rk}(A)$ *for all* $A \subseteq X$.

If rk is a polymatroid then $Q(\mathrm{rk})$ is a polytope, called the *base polytope* of rk. In fact $Q(\mathrm{rk})$ is always nonempty. The base polytope of a polymatroid $(\underline{d}, \mathrm{rk})$ of rank r is contained in the simplex

$$\Delta_{\mathrm{PM}}(d, r) = \{(y_1, \ldots, y_d) \in \mathbb{R}^d \mid y_1, \ldots, y_d \geq 0,\ y_1 + y_2 + \cdots + y_d = r\}$$

and the base polytope of a matroid $(\underline{d}, \mathrm{rk})$ of rank r is contained in the hypersimplex

$$\Delta_{\mathrm{M}}(d, r) = \{(y_1, \ldots, y_d) \in \mathbb{R}^d \mid 0 \leq y_1, \ldots, y_d \leq 1,\ y_1 + y_2 + \cdots + y_d = r\}.$$

The next theorem generalises a theorem of Gelfand-Goresky-MacPherson-Serganova [10] on matroids.

Theorem 2.3 *A convex polyhedron contained in* $y_1 + \cdots + y_d = r$ *equals* $Q(M)$ *for some megamatroid* M *if and only if for every face* F *of* Π*, the linear hull* $\mathrm{lhull}(F)$ *is of the form* $z + W$ *where* $z \in \mathbb{Z}^d$ *and* W *is spanned by vectors of the form* $e_i - e_j$.

Polymatroid polyhedra are, up to translation, the lattice polytopes among the *generalized permutohedra* of [18] or the *submodular rank tests* of [16].

3 The valuative property

There are essentially two definitions of the valuative property in the literature, which we will refer to as the *strong valuative* and the *weak valuative* properties. The equivalence of these definitions is shown in [12] and [20] when valuations are defined on sets of polyhedra closed under intersection. We show their equivalence for megamatroid polytopes, which are not such a set.

Definition 3.1 *A* megamatroid polyhedron decomposition *is a decomposition* $\Pi = \Pi_1 \cup \Pi_2 \cup \cdots \cup \Pi_k$ *such that* $\Pi, \Pi_1, \ldots, \Pi_k$ *are megamatroid polyhedra, and* $\Pi_i \cap \Pi_j$ *is empty or contained in a proper face of* Π_i *and of* Π_j *for all* $i \neq j$.

A megamatroid polyhedron decomposition $\Pi = \Pi_1 \cup \cdots \cup \Pi_k$ is a *(poly)matroid polytope decomposition* if $\Pi, \Pi_1, \ldots, \Pi_k$ are (poly)matroid polytopes.

For a megamatroid polyhedron decomposition $\Pi = \Pi_1 \cup \Pi_2 \cup \cdots \cup \Pi_k$ we define $\Pi_I = \bigcap_{i \in I} \Pi_i$ if $I \subseteq \{1, 2, \ldots, k\}$, and $\Pi_\emptyset = \Pi$. Define

$$m_{\mathrm{val}}(\Pi; \Pi_1, \ldots, \Pi_k) = \sum_{I \subseteq \{1,2,\ldots,k\}} (-1)^{|I|} m_I \in Z_{\mathrm{MM}}(d, r),$$

where $m_I = \langle \mathrm{rk}^I \rangle$ if rk^I is the megamatroid with $Q(\mathrm{rk}^I) = \Pi_I$, and $m_I = 0$ if $\Pi_I = \emptyset$.

Definition 3.2 *A homomorphism of abelian groups* $f : Z_{*\mathrm{M}}(d,r) \to A$ *is called* weakly valuative *if for every megamatroid polyhedron decomposition* $\Pi = \Pi_1 \cup \Pi_2 \cup \cdots \cup \Pi_k$ *we have*

$$f(m_{\mathrm{val}}(\Pi; \Pi_1, \ldots, \Pi_k)) = 0.$$

For a polyhedron Π in $\mathbb{R}^d$, let $[\Pi]$ denote its indicator function. Define $P_{\mathrm{MM}}(d,r)$ as the $\mathbb{Z}$-module generated by all $[Q(\mathrm{rk})]$, where rk lies in $S_{\mathrm{MM}}(d,r)$. There is a natural $\mathbb{Z}$-module homomorphism

$$\Psi_{*\mathrm{M}} : Z_{*\mathrm{M}}(d,r) \to P_{*\mathrm{M}}(d,r)$$

such that $\Psi_{*\mathrm{M}}(\langle \mathrm{rk} \rangle) = [Q(\mathrm{rk})]$ for all $\mathrm{rk} \in S_{*\mathrm{M}}(d,r)$.

Definition 3.3 *A homomorphism of groups* $f : Z_{*\mathrm{M}}(d,r) \to A$ *is* strongly valuative *if there exists a group homomorphism* $\widehat{f} : P_{*\mathrm{M}}(d,r) \to A$ *such that* $f = \widehat{f} \circ \Psi_{*\mathrm{M}}$.

The map $\Psi_{*\mathrm{M}}$ has the weak valuative property, which shows that the strong valuative property implies the weak valuative property. In fact the two valuative properties are equivalent, and in view of this we may speak of *the* valuative property.

Theorem 3.4 *A homomorphism* $f : Z_{*\mathrm{M}}(d,r) \to A$ *of abelian groups is weakly valuative if and only if it is strongly valuative.*

4 Decompositions into cones

A chain of length $k =: \mathrm{length}(\underline{X})$ in $\underline{d}$ is $\underline{X} : \emptyset \subset X_1 \subset \cdots \subset X_{k-1} \subset X_k = \underline{d}$ (here $\subset$ denotes proper inclusion). If $d > 0$ then every chain has length ≥ 1, but for $d = 0$ there is exactly one chain, namely $\emptyset = \underline{0}$, and this chain has length 0. For a chain $\underline{X}$ of length k and a k-tuple $\underline{r} = (r_1, r_2, \ldots, r_k) \in (\mathbb{Z} \cup \{\infty\})^k$, we define a megamatroid polyhedron

$$R_{\mathrm{MM}}(\underline{X}, \underline{r}) = \Big\{ (y_1, \ldots, y_d) \in \mathbb{R}^d \;\Big|\; \sum_{i=1}^{d} y_i = r_k,\ \forall j \sum_{i \in X_j} y_i \leq r_j \Big\}.$$

We will always use the conventions $r_0 = 0$, $X_0 = \emptyset$. Note that the polytopes $R_{\mathrm{MM}}(\underline{X}, \underline{r})$ are full-dimensional cones in $\{\sum_{y=1}^{d} y_i = r\}$.

The next theorem is an analogue of the Brianchon-Gram Theorem [4, 11] for megamatroid polytopes.

Theorem 4.1 *For any megamatroid* $\mathrm{rk} : 2^{\underline{d}} \to \mathbb{Z} \cup \{\infty\}$ *we have*

$$[Q(\mathrm{rk})] = \sum_{\underline{X}} (-1)^{d - \mathrm{length}(\underline{X})} [R_{\mathrm{MM}}(\underline{X}, (\mathrm{rk}(X_1), \ldots, \mathrm{rk}(X_k)))].$$

Example 4.2 *Consider the case where* $d = 3$ *and* $r = 3$, *and* rk *is defined by* $\mathrm{rk}(\{1\}) = \mathrm{rk}(\{2\}) = \mathrm{rk}(\{3\}) = 2$, $\mathrm{rk}(\{1,2\}) = \mathrm{rk}(\{2,3\}) = \mathrm{rk}\{(1,3\}) = 3$, $\mathrm{rk}(\{1,2,3\}) = 4$. *The right of Figure 1 depicts the decomposition using the Brianchon-Gram theorem of a polytope* $Q_\varepsilon(\mathrm{rk})$, *which is defined by a certain perturbation of the inequalities defining* $Q(\mathrm{rk})$. *Note how the summands in the decomposition correspond to the faces of* $Q_\varepsilon(\mathrm{rk})$. *In the limit where the perturbation approaches* 0, $Q_\varepsilon(\mathrm{rk})$ *tends to* $Q(\mathrm{rk})$ *and we obtain the left of Figure 1. This is exactly the decomposition in Theorem 4.1. In this decomposition, the summands do not correspond bijectively to the faces of* $Q(\mathrm{rk})$.

The dashed triangle is the triangle defined by $y_1, y_2, y_3 \geq 0$, $y_1 + y_2 + y_3 = 4$.

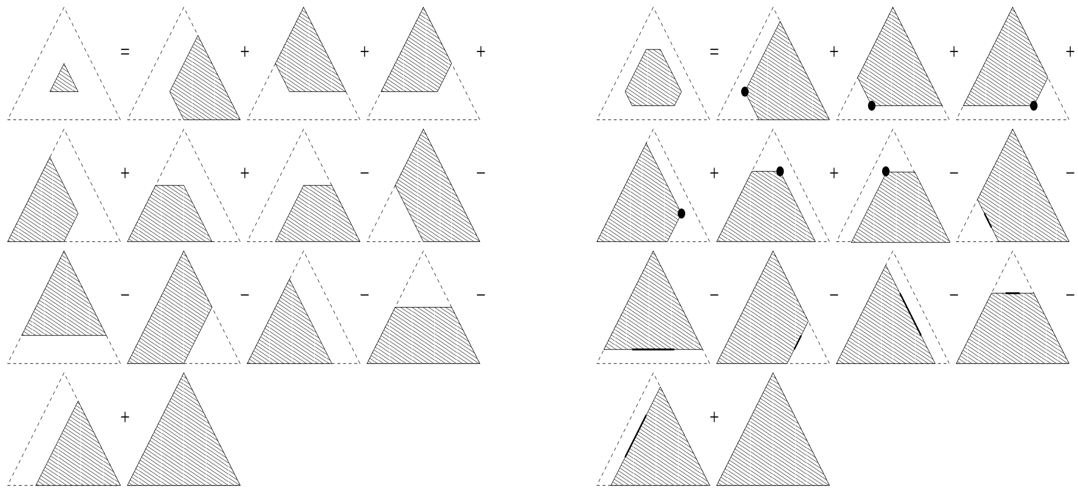


Fig. 1: At left, a decomposition of $Q(\mathrm{rk})$ as in Theorem 4.1. At right, a decomposition of a polytope $Q_\epsilon(\mathrm{rk})$.

5 Valuations

Suppose that $d \geq 1$. Let $\mathfrak{p}_{\mathrm{MM}}(d,r)$ be the set of all pairs $(\underline{X},\underline{r})$ such that $\underline{X}$ is a chain of length k $(1 \leq k \leq d)$ and $\underline{r} = (r_1, r_2, \dots, r_k)$ is an integer vector with $r_k = r$. We define $R_{\mathrm{(P)M}}(\underline{X},\underline{r}) = R_{\mathrm{MM}}(\underline{X},\underline{r}) \cap \Delta_{\mathrm{(P)M}}(d,r)$. If $R_{\mathrm{(P)M}}(\underline{X},\underline{r})$ is nonempty, then it is a (poly)matroid base polytope. Define $\mathfrak{p}_{\mathrm{PM}}(d,r) \subseteq \mathfrak{p}_{\mathrm{MM}}(d,r)$ as the set of all pairs $(\underline{X},\underline{r})$ with $0 \leq r_1 < \cdots < r_k = r$. Let $\mathfrak{p}_{\mathrm{M}}(d,r)$ denote the set of all pairs $(\underline{X},\underline{r}) \in \mathfrak{p}_{\mathrm{MM}}(d,r)$ such that $\underline{r} = (r_1, \dots, r_k)$ for some k $(1 \leq k \leq d)$,

$$0 \leq r_1 < r_2 < \cdots < r_k = r$$

and

$$0 < |X_1| - r_1 < |X_2| - r_2 < \cdots < |X_{k-1}| - r_{k-1} \leq |X_k| - r_k = d - r.$$

For $d = 0$, we define $\mathfrak{p}_{\mathrm{MM}}(0,r) = \mathfrak{p}_{\mathrm{PM}}(0,r) = \mathfrak{p}_{\mathrm{M}}(0,r) = \emptyset$ for $r \neq 0$ and $\mathfrak{p}_{\mathrm{MM}}(0,0) = \mathfrak{p}_{\mathrm{PM}}(0,0) = \mathfrak{p}_{\mathrm{M}}(0,0) = \{(\emptyset \subseteq \underline{0}, ())\}$.

Theorem 5.1 *The group $P_{*\mathrm{M}}(d,r)$ is freely generated by the basis $\{[R_{*\mathrm{M}}(\underline{X},\underline{r})] \mid (\underline{X},\underline{r}) \in \mathfrak{p}_{*\mathrm{M}}(d,r)\}$.*

Note that the basis of this theorem is a generating set by Theorem 4.1.

Suppose that $\underline{X}$ is a chain of length k and $\underline{r} = (r_1, \dots, r_k)$ is an integer vector with $r_k = r$. Define a homomorphism $s_{\underline{X},\underline{r}} : Z_{\mathrm{MM}}(d,r) \to \mathbb{Z}$ by

$$s_{\underline{X},\underline{r}}(\mathrm{rk}) = \begin{cases} 1 & \text{if } \mathrm{rk}(X_j) = r_j \text{ for } j = 1,2,\dots,k, \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 5.2 *The homomorphism $s_{\underline{X},\underline{r}}$ is valuative.*

Theorem 5.3 *The group $P_{\mathrm{(P)M}}(d,r)^\vee$ is freely generated by the basis $\{s_{\underline{X},\underline{r}} : (\underline{X},\underline{r}) \in \mathfrak{p}_{\mathrm{(P)M}}(d,r)\}$.*

If $\underline{X}$ is not a maximal chain, then $s_{\underline{X},\underline{r}}$ is a linear combination of functions of the form $s_{\underline{X}',\underline{r}'}$ where $\underline{X}'$ is a maximal chain. The set of such functions $s_{\underline{X}',\underline{r}'}$ appeared as the coordinates of the function H defined in [1, §6], which was introduced there as a labeled analogue of $\mathcal{G}$.

6 Valuative invariants

Let $Y_{*\mathrm{M}}(d,r)$ be the group generated by all $\langle \mathrm{rk}\rangle - \langle \mathrm{rk}\circ\sigma\rangle$ where $\mathrm{rk} : 2^{\underline{d}} \to \mathbb{Z}\cup\{\infty\}$ is a $*$matroid of rank r and σ is a permutation of $\underline{d}$. We define $Z^{\mathrm{sym}}_{*\mathrm{M}}(d,r) = Z_{*\mathrm{M}}(d,r)/Y_{*\mathrm{M}}(d,r)$. Let $\pi_{*\mathrm{M}} : Z_{*\mathrm{M}}(d,r) \to Z^{\mathrm{sym}}_{*\mathrm{M}}(d,r)$ be the quotient homomorphism. If $\mathrm{rk}_X : 2^X \to \mathbb{Z}\cup\{\infty\}$ is any $*$matroid, then we can choose a bijection $\varphi : \underline{d} \to X$, where d is the cardinality of X. Let $r = \mathrm{rk}_X(X)$. The image of $\langle \mathrm{rk}_X \circ\varphi\rangle$ in $Z^{\mathrm{sym}}_{*\mathrm{M}}(d,r)$ does not depend on φ, and will be denoted by $[\mathrm{rk}_X]$. The $*$matroids (X,rk_X) and (Y,rk_Y) are isomorphic if and only if $[\mathrm{rk}_X] = [\mathrm{rk}_Y]$. So we may think of $Z^{\mathrm{sym}}_{*\mathrm{M}}(d,r)$ as the free group generated by all isomorphism classes of rank r $*$matroids on sets with d elements.

Let $P^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r)$ be the pushout of the diagram

$$\begin{tikzcd} Z_{(\mathrm{P})\mathrm{M}}(d,r) \arrow[r, "\pi_{(\mathrm{P})\mathrm{M}}"] \arrow[d, "\Psi_{(\mathrm{P})\mathrm{M}}"'] & Z^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r) \arrow[d, dotted, "\Psi^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}"] \\ P_{(\mathrm{P})\mathrm{M}}(d,r) \arrow[r, dotted, "\rho_{(\mathrm{P})\mathrm{M}}"'] & P^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r) \end{tikzcd} \quad . \tag{1}$$

Then the dual space $P^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r)^\vee$ is exactly the set of all $\mathbb{Z}$-valued valuative (poly)matroid invariants. Define $\mathfrak{p}^{\mathrm{sym}}_{*\mathrm{M}}(d,r)$ as the set of all pairs $(\underline{X},\underline{r}) \in \mathfrak{p}_{*\mathrm{M}}(d,r)$ such that every X_j is of the form $\{1,2,\dots,i\}$.

Theorem 6.1 *The $\mathbb{Z}$-module $P^{\mathrm{sym}}_{\star\mathrm{M}}(d,r)$ is freely generated by all $\rho_{\star\mathrm{M}}([R_{\star\mathrm{M}}(\underline{X},\underline{r})])$ with $(\underline{X},\underline{r}) \in \mathfrak{p}^{\mathrm{sym}}_{\star\mathrm{M}}(d,r)$.*

The matroid polytopes $R_{\mathrm{M}}(\underline{X},\underline{r})$ with $(\underline{X},\underline{r}) \in \mathfrak{p}^{\mathrm{sym}}_{\star\mathrm{M}}(d,r)$ are the polytopes of *Schubert matroids*. Schubert matroids were first described by Crapo [6], and have since arisen in several contexts, prominent among these being the stratification of the Grassmannian into Schubert cells [2, §2.4].

For $(\underline{X},\underline{r}) \in \mathfrak{p}^{\mathrm{sym}}_{\mathrm{MM}}(d,r)$, define a homomorphism $s^{\mathrm{sym}}_{\underline{X},\underline{r}} : Z_{\mathrm{MM}}(d,r) \to \mathbb{Z}$ by

$$s^{\mathrm{sym}}_{\underline{X},\underline{r}} = \sum_{\sigma\underline{X}} s_{\sigma\underline{X},\underline{r}}$$

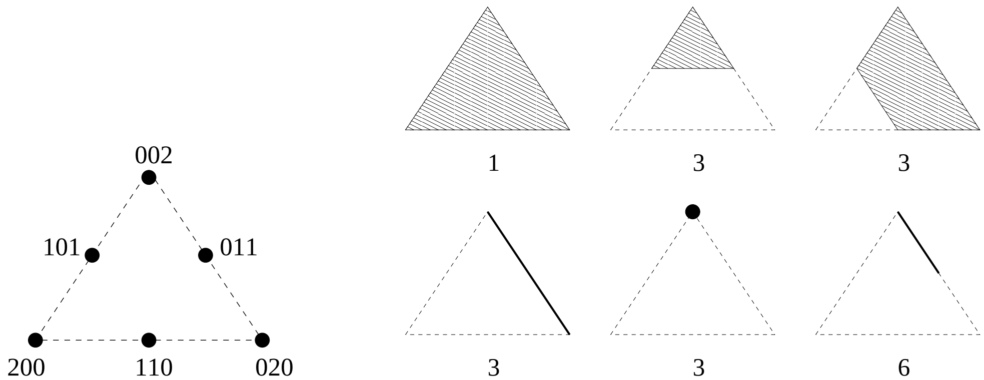
where the sum is over all chains $\sigma\underline{X}$ in the orbit of $\underline{X}$ under the action of the symmetric group.

Theorem 6.2 *The $\mathbb{Q}$-vector space $P^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r)^\vee \otimes_{\mathbb{Z}} \mathbb{Q}$ of valuations $Z^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r) \to \mathbb{Q}$ has a basis given by the functions $s^{\mathrm{sym}}_{\underline{X},\underline{r}}$ for $(\underline{X},\underline{r}) \in \mathfrak{p}^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r)$.*

For a sequence $\alpha = (\alpha_1,\dots,\alpha_d)$ of nonnegative integers with $|\alpha| = \sum_i \alpha_i = r$, we define $u_\alpha = s_{\underline{X},\underline{r}} : Z^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r) \to \mathbb{Z}$, where $\underline{X}_i = \underline{i}$ for $i = 1,2,\dots,r$ and $\underline{r} = (\alpha_1, \alpha_1+\alpha_2, \dots, \alpha_1+\cdots+\alpha_d)$.

Corollary 6.3 *The $\mathbb{Q}$-vector space $P^{\mathrm{sym}}_{\mathrm{PM}}(d,r)^\vee \otimes_{\mathbb{Z}} \mathbb{Q}$ of valuations $Z^{\mathrm{sym}}_{\mathrm{PM}}(d,r) \to \mathbb{Q}$ has a $\mathbb{Q}$-basis given by the functions u_α, where α runs over all sequences $(\alpha_1,\dots,\alpha_d)$ of nonnegative integers with $|\alpha| = r$.*

Corollary 6.4 *The $\mathbb{Q}$-vector space $P^{\mathrm{sym}}_{\mathrm{M}}(d,r)^\vee \otimes_{\mathbb{Z}} \mathbb{Q}$ of valuations $Z^{\mathrm{sym}}_{\mathrm{M}}(d,r) \to \mathbb{Q}$ has a $\mathbb{Q}$-basis given by all functions u_α where α runs over all sequences $(\alpha_1,\dots,\alpha_d) \in \{0,1\}^d$ with $|\alpha| = r$.*



$\underline{X}:\ \{1,2,3\}$ $\underline{r} =\ (2)$	$\underline{X}:\ \{1,2\} \subset \{1,2,3\}$ $\underline{r} =\ (1,2)$	$\underline{X}:\ \{1\} \subset \{1,2,3\}$ $\underline{r} =\ (1,2)$
$\underline{X}:\ \{1\} \subset \{1,2,3\}$ $\underline{r} =\ (0,2)$	$\underline{X}:\ \{1,2\} \subset \{1,2,3\}$ $\underline{r} =\ (0,2)$	$\underline{X}:\ \{1\} \subset \{1,2\} \subset \{1,2,3\}$ $\underline{r} =\ (0,1,2)$

Fig. 2: A polymatroid example of Theorems 5.1 and 6.1.

Example 6.5 *Figures 2 and 3 are examples of both Theorems 5.1 and 6.1 for matroids in the cases $(d,r) = (3,2)$ and $(d,r) = (4,2)$, respectively. At top left are the polyhedra $\Delta_{(\mathrm{P})\mathrm{M}}(d,r)$, containing all the (poly)matroid base polytopes. At top right are the polytopes $R(\underline{X},\underline{r})$ for $(\underline{X},\underline{r}) \in \mathfrak{p}^{\mathrm{sym}}_{(\mathrm{P})\mathrm{M}}(d,r)$, and at bottom the corresponding pairs $(\underline{X},\underline{r})$.*

The symmetric group Σ_d acts on Δ_d by permuting the coordinates. If Σ_d acts on the generators $R(\underline{X},\underline{r})$ with $(\underline{X},\underline{r}) \in \mathfrak{p}_{(\mathrm{P})\mathrm{M}}(d,r)$, then we get all $R(\underline{X},\underline{r})$ with $(\underline{X},\underline{r}) \in \mathfrak{p}_{(\mathrm{P})\mathrm{M}}(d,r)$. In the figure, we have written under each polytope the cardinality of its Σ_d-orbit.

7 Hopf algebra structures

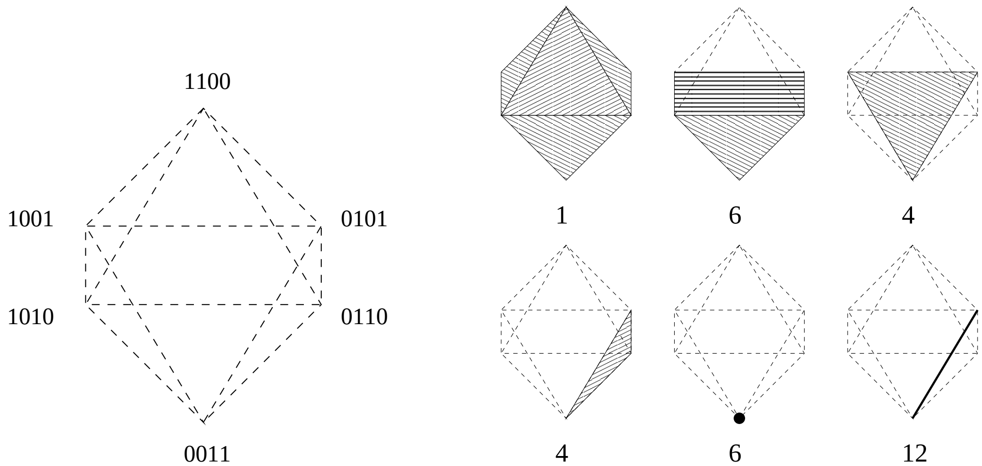
Define $Z_{*\mathrm{M}} = \bigoplus_{d,r} Z_{*\mathrm{M}}(d,r)$, and in a similar way define $Z^{\mathrm{sym}}_{*\mathrm{M}}$, $P_{*\mathrm{M}}$, and $P^{\mathrm{sym}}_{*\mathrm{M}}$. We can view $Z^{\mathrm{sym}}_{*\mathrm{M}}$ as the $\mathbb{Z}$-module freely generated by all isomorphism classes of $*$matroids. In this section we will only speak of the megamatroid objects; in every case, there are analogous matroid and polymatroid objects, which are substructures.

If $\mathrm{rk}_1 : 2^{\underline{d}} \to \mathbb{Z} \cup \{\infty\}$ and $\mathrm{rk}_2 : 2^{\underline{e}} \to \mathbb{Z} \cup \{\infty\}$ then we define $\mathrm{rk}_1 \boxplus \mathrm{rk}_2 : 2^{\underline{d+e}} \to \mathbb{Z} \cup \{\infty\}$ by

$$(\mathrm{rk}_1 \boxplus \mathrm{rk}_2)(A) = \mathrm{rk}_1(A \cap \underline{d}) + \mathrm{rk}_2(\{i \in \underline{e} \mid d+i \in A\})$$

for any set $A \subseteq \underline{d+e}$. Note that $\boxplus$ is not commutative. We have a multiplication $\nabla : Z_{\mathrm{MM}} \otimes_{\mathbb{Z}} Z_{\mathrm{MM}} \to Z_{\mathrm{MM}}$ defined by $\nabla(\langle \mathrm{rk}_1 \rangle \otimes \langle \mathrm{rk}_2 \rangle) = \langle \mathrm{rk}_1 \boxplus \mathrm{rk}_2 \rangle$, which makes $Z_{\mathrm{MM}}(d,r)$ into an associative (noncommutative) ring with 1. The multiplication also respects the bigrading of $Z_{\mathrm{MM}}(d,r)$. The unit $\eta : \mathbb{Z} \to Z_{\mathrm{MM}}(d,r)$ is given by $1 \mapsto \langle \mathrm{rk}_0 \rangle$ where $\mathrm{rk}_0 : 2^{\underline{0}} \to \mathbb{Z} \cup \{\infty\}$ is the unique megamatroid defined by $\mathrm{rk}(\emptyset) = 0$.

Next, we define a comultiplication for Z_{MM}. Suppose that $X = \{i_1, i_2, \ldots, i_d\}$ is a set of integers with $i_1 < \cdots < i_d$ and $\mathrm{rk} : 2^X \to \mathbb{Z} \cup \{\infty\}$ is a megamatroid. We define a megamatroid $\widehat{\mathrm{rk}} : 2^{\underline{d}} \to \mathbb{Z} \cup \{\infty\}$



$\underline{X}: \quad \{1,2,3,4\}$ $\underline{r} = \quad (2)$	$\underline{X}: \quad \{1,2\} \subset \{1,2,3,4\}$ $\underline{r} = \quad (1,2)$	$\underline{X}: \quad \{1,2,3\} \subset \{1,2,3,4\}$ $\underline{r} = \quad (1,2)$
$\underline{X}: \quad \{1\} \subset \{1,2,3,4\}$ $\underline{r} = \quad (0,2)$	$\underline{X}: \quad \{1,2\} \subset \{1,2,3,4\}$ $\underline{r} = \quad (0,2)$	$\underline{X}: \quad \{1\} \subset \{1,2,3\} \subset \{1,2,3,4\}$ $\underline{r} = \quad (0,1,2)$

Fig. 3: A matroid example of Theorems 5.1 and 6.1.

by $\widehat{\mathrm{rk}}(A) = \mathrm{rk}(\{i_j \mid j \in A\})$. If $\mathrm{rk} : 2^X \to \mathbb{Z} \cup \{\infty\}$ is a megamatroid and $B \subseteq A \subseteq X$ then we define $\mathrm{rk}_{A/B} : 2^{A \setminus B} \to \mathbb{Z} \cup \{\infty\}$ by $\mathrm{rk}_{A/B}(C) = \mathrm{rk}(B \cup C) - \mathrm{rk}(B)$ for all $C \subseteq A \setminus B$. We also define $\mathrm{rk}_A := \mathrm{rk}_{A/\emptyset}$ and $\mathrm{rk}_{/B} = \mathrm{rk}_{X/B}$.

We now define $\Delta : Z_{\mathrm{MM}} \to Z_{\mathrm{MM}} \otimes_{\mathbb{Z}} Z_{\mathrm{MM}}$ by

$$\Delta(\langle \mathrm{rk} \rangle) = \sum_{A \subseteq \underline{d};\ \mathrm{rk}(A) < \infty} \langle \widehat{\mathrm{rk}_A} \rangle \otimes \langle \widehat{\mathrm{rk}_{/A}} \rangle.$$

where A runs over all subsets of $\underline{d}$ for which $\mathrm{rk}(A)$ is finite. This comultiplication is coassociative, but not cocommutative. If $\mathrm{rk} : 2^{\underline{d}} \to \mathbb{Z} \cup \{\infty\}$ is a megamatroid, then the counit is defined by $\epsilon(\langle \mathrm{rk} \rangle) = 1$ (if $d = 0$), 0 (otherwise). We omit here the definition of the antipode S.

It is well-known that $Z_{\mathrm{M}}^{\mathrm{sym}}$ has the structure of a Hopf algebra over $\mathbb{Z}$. In fact we have that $Z_{\mathrm{PM}}^{\mathrm{sym}}$ has a Hopf algebra structure, with $Z_{\mathrm{M}}^{\mathrm{sym}}$ as a Hopf subalgebra. This structure is defined analogously to the one on Z_{MM} above, replacing each megamatroid by its isomorphism class: e.g. multiplication is given by the direct sum of megamatroids, and is now commutative. The map π_{MM} of (1) is a Hopf algebra morphism.

The space P_{MM} inherits a Hopf algebra structure from Z_{MM}. Most of this structure can be defined in the expected fashion, but the coproduct requires some care. We define $\Delta : P_{\mathrm{MM}} \to P_{\mathrm{MM}} \otimes P_{\mathrm{MM}}$ by

$$\Delta([R_{\mathrm{MM}}(\underline{X}, \underline{r})]) = \sum_{i=0}^{k} [R_{\mathrm{MM}}(\widehat{\underline{X}_i, \underline{r}_i})] \otimes [R_{\mathrm{MM}}(\widehat{\underline{X}^i, \underline{r}^i})].$$

Since the $R_{\mathrm{MM}}(\underline{X}, \underline{r})$ with $(\underline{X}, \underline{r}) \in \mathfrak{p}_{\mathrm{MM}} = \bigcup_{d,r} \mathfrak{p}_{\mathrm{MM}}(d, r)$ form a basis of P_{MM}, this is sufficient to linearly extend. From Theorem 4.1 one can check that $(\Psi_{\mathrm{MM}} \otimes \Psi_{\mathrm{MM}}) \otimes \Delta = \Delta \circ \Psi_{\mathrm{MM}}$.

The Hopf algebra structure on P_{MM} naturally induces a Hopf algebra structure on $P^{\mathrm{sym}}_{\mathrm{MM}}$ such that ρ_{MM} and $\Psi^{\mathrm{sym}}_{\mathrm{MM}}$ are Hopf algebra homomorphisms.

8 Additive valuations

For $0 \leq e \leq d$ we define $P_{*\mathrm{M}}(d,r,e) \subseteq P_{*\mathrm{M}}(d,r)$ as the span of all $[\Pi]$ where $\Pi \subseteq \mathbb{R}^d$ is a $*$matroid polytope of dimension $\leq d-e$. We have $P_{*\mathrm{M}}(0,r,0) = P_{*\mathrm{M}}(0,r)$ and $P_{*\mathrm{M}}(d,r,1) = P_{*\mathrm{M}}(d,r)$ for $d \geq 1$. These subgroups form a filtration

$$\cdots \subseteq P_{*\mathrm{M}}(d,r,2) \subseteq P_{*\mathrm{M}}(d,r,1) \subseteq P_{*\mathrm{M}}(d,r,0) = P_{*\mathrm{M}}(d,r).$$

Define $\overline{P}_{*\mathrm{M}}(d,r,e) := P_{*\mathrm{M}}(d,r,e)/P_{*\mathrm{M}}(d,r,e+1)$, and $T_{\star\mathrm{M}}(d,r) = \overline{P}_{\star\mathrm{M}}(d,r,1)$. The image of $[Q(M)]$ in $T_{*\mathrm{M}}(d,r)$ is zero if and only if M is *connected*. The associated graded algebra $\overline{P}_{*\mathrm{M}} = \bigoplus_{d,r,e} \overline{P}_{*\mathrm{M}}(d,r,e)$ has an induced Hopf algebra structure.

Define

$$\overline{P}_{*\mathrm{M}}(\underline{X}) = \bigoplus_{r_1,r_2,\ldots,r_e \in \mathbb{Z}} T_{*\mathrm{M}}(|X_1|,r_1) \otimes \cdots \otimes T_{*\mathrm{M}}(|X_e|,r_e).$$

There is a group homomorphism $\overline{\phi}_{\underline{X}} : \overline{P}_{*\mathrm{M}}(\underline{X}) \to \overline{P}_{*\mathrm{M}}(d,r,e)$. which takes the classes of a list of $*$matroids to the class of their direct sum. The next theorem essentially asserts a unique decomposition of $*$matroids into connected components.

Theorem 8.1 *We have the isomorphism*

$$\left(\sum_{\underline{X}} \overline{\phi}_{\underline{X}}\right) : \bigoplus_{\substack{\underline{X}=(X_1,X_2,\ldots,X_e) \\ \underline{d}=X_1 \sqcup X_2 \sqcup \cdots \sqcup X_e; X_1,\ldots,X_e \neq \emptyset}} \overline{P}_{*\mathrm{M}}(\underline{X}) \to \bigoplus_{r \in \mathbb{Z}} \overline{P}_{*\mathrm{M}}(d,r,e) \tag{2}$$

If $d \geq 1$, let $\mathfrak{t}_{\mathrm{PM}}(d,r)$ be the set of all pairs $(\underline{X},\underline{r}) \in \mathfrak{p}_{\mathrm{PM}}(d,r)$ such that $r_1 > 0$, and $d \notin X_{k-1}$, where k is the length of $\underline{X}$. Similarly, if $d \geq 2$, let $\mathfrak{t}_{\mathrm{M}}(d,r)$ be the set of all pairs $(\underline{X},\underline{r}) \in \mathfrak{t}_{\mathrm{M}}(d,r)$ such that $r_1 > 0$, $|X_{k-1}| - r_{k-1} < d-r$, and $d \notin X_{k-1}$.

Theorem 8.2 *The group $T_{(\mathrm{P})\mathrm{M}}(d,r)$ is freely generated by all $[R_{(\mathrm{P})\mathrm{M}}(\underline{X},\underline{r})]$ with $(\underline{X},\underline{r}) \in \mathfrak{t}_{(\mathrm{P})\mathrm{M}}(d,r)$.*

9 Additive invariants

The algebra $P^{\mathrm{sym}}_{\star M}$ also has a natural filtration:

$$\cdots \subseteq P^{\mathrm{sym}}_{\star\mathrm{M}}(d,r,2) \subseteq P^{\mathrm{sym}}_{\star\mathrm{M}}(d,r,1) \subseteq P^{\mathrm{sym}}_{\star M}(d,r,0) = P^{\mathrm{sym}}_{\star\mathrm{M}}(d,r).$$

Here $P^{\mathrm{sym}}_{\star\mathrm{M}}(d,r,e)$ is spanned by the indicator functions of all $*$matroid base polytopes of rank r and dimension $d-e$. Define $\overline{P}^{\mathrm{sym}}_{\star\mathrm{M}}(d,r,e) = P^{\mathrm{sym}}_{*\mathrm{M}}(d,r,e)/P^{\mathrm{sym}}_{*\mathrm{M}}(d,r,e+1)$. Let $\overline{P}^{\mathrm{sym}}_{\star\mathrm{M}} = \bigoplus_{d,r,e} \overline{P}^{\mathrm{sym}}_{\star\mathrm{M}}(d,r,e)$ be the associated graded algebra.

Define $T^{\mathrm{sym}}_{\star\mathrm{M}} = \bigoplus_{d,r} \overline{P}^{\mathrm{sym}}_{\star\mathrm{M}}(d,r,1)$. The following theorem follows from Theorem 8.1.

Theorem 9.1 *The algebra $\overline{P}^{\mathrm{sym}}_{\star\mathrm{M}}$ is the free symmetric algebra $S(T^{\mathrm{sym}}_{\star\mathrm{M}})$ on $T^{\mathrm{sym}}_{\star\mathrm{M}}$, and there exists an isomorphism*

$$S^e(T^{\mathrm{sym}}_{\star\mathrm{M}}) \cong \bigoplus_{d,r} \overline{P}^{\mathrm{sym}}_{\star\mathrm{M}}(d,r,e). \tag{3}$$

Corollary 9.2 *The algebra $P^{\mathrm{sym}}_{\star\mathrm{M}}$ is a polynomial ring over $\mathbb{Z}$.*

10 Invariants as elements in free algebras

Let
$$(P^{\mathrm{sym}}_{*\mathrm{M}})^{\#} := \bigoplus_{d,r} P^{\mathrm{sym}}_{*\mathrm{M}}(d,r)^{\vee}$$
be the *graded* dual of $P^{\mathrm{sym}}_{*\mathrm{M}}$. Let $\mathfrak{m}_{\star\mathrm{M}} = \bigoplus_{d,r} P^{\mathrm{sym}}_{*\mathrm{M}}(d,r,1)$. Then we have $\mathfrak{m}^2_{\star\mathrm{M}} = \bigoplus_{d,r} P^{\mathrm{sym}}_{*\mathrm{M}}(d,r,2)$ and $T^{\mathrm{sym}}_{*\mathrm{M}} = \mathfrak{m}_{*\mathrm{M}}/\mathfrak{m}^2_{*\mathrm{M}}$. The graded dual $\mathfrak{m}^{\#}_{\star\mathrm{M}}$ can be identified with
$$(P^{\mathrm{sym}}_{\star\mathrm{M}})^{\#}/P^{\mathrm{sym}}_{*\mathrm{M}}(0,0) \cong \bigoplus_{d=1}^{\infty} \bigoplus_{r} P^{\mathrm{sym}}_{\star\mathrm{M}}(d,r)^{\vee}.$$

So $\mathfrak{m}^{\#}_{\mathrm{PM}} \otimes_{\mathbb{Z}} \mathbb{Q}$ will be identified with the ideal $(u_0, u_1, \dots)$ of $\mathbb{Q}\langle u_0, u_1, \dots\rangle$ and $\mathfrak{m}^{\#}_{\mathrm{M}} \otimes_{\mathbb{Z}} \mathbb{Q}$ will be identified with the ideal (u_0, u_1) of $\mathbb{Q}\langle u_0, u_1\rangle$. The graded dual $(T^{\mathrm{sym}}_{\mathrm{PM}})^{\#} \otimes_{\mathbb{Z}} \mathbb{Q}$ is a subalgebra (without 1) of the ideal $(u_0, u_1, \dots)$, and $(T^{\mathrm{sym}}_{\mathrm{PM}})^{\#} \otimes_{\mathbb{Z}} \mathbb{Q}$ is a subalgebra of (u_0, u_1).

Theorem 10.1 *Let $u_0, u_1, u_2, \dots$ be indeterminates, where u_i has bidgree $(1,i)$. We have the following isomorphisms of bigraded associative algebras over $\mathbb{Q}$:*

a. *The space $(P^{\mathrm{sym}}_{M})^{\vee} \otimes_{\mathbb{Z}} \mathbb{Q}$ is isomorphic to $\mathbb{Q}\langle\langle u_0, u_1\rangle\rangle$, the completion (in power series) of the free associative algebra generated by u_0, u_1.*

b. *The space $(P^{\mathrm{sym}}_{\mathrm{PM}})^{\vee} \otimes_{\mathbb{Z}} \mathbb{Q}$ is isomorphic to $\mathbb{Q}\langle\langle u_0, u_1, u_2, \dots\rangle\rangle$.*

c. *The space $(T^{\mathrm{sym}}_{\mathrm{M}})^{\vee} \otimes_{\mathbb{Z}} \mathbb{Q}$ is isomorphic to $\mathbb{Q}\{\!\{u_0, u_1\}\!\}$, the completion of the free Lie algebra generated by u_0, u_1.*

d. *The space $(T^{\mathrm{sym}}_{\mathrm{PM}})^{\vee} \otimes_{\mathbb{Z}} \mathbb{Q}$ is isomorphic to $\mathbb{Q}\{\!\{u_0, u_1, u_2, \dots\}\!\}$.*

Proposition 10.2 *The Hopf algebra $P^{\mathrm{sym}}_{\mathrm{PM}} \otimes_{\mathbb{Z}} \mathbb{Q}$ is isomorphic to the ring $QSym$ of quasi-symmetric functions over $\mathbb{Q}$.*

If we identify $P^{\mathrm{sym}}_{\mathrm{PM}} \otimes_{\mathbb{Z}} \mathbb{Q}$ with $QSym$, then $\mathcal{G}$ is equal to $\Psi^{\mathrm{sym}}_{\mathrm{PM}}$.

Acknowledgements

The second author is grateful to David Speyer for collaboration which led to a proof of Theorem 1.4 before he got in touch with the first author.

References

[1] F. Ardila, A. Fink, F. Rincón, *Valuations for matroid polytope subdivisions*, Canadian Math. Bulletin **13**, to appear.

[2] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler. *Oriented Matroids.* Encyclopedia of Mathematics and Its Applications, vol. 46. Cambridge University Press, Cambridge, 1993.

[3] L.J. Billera, N. Jia, V. Reiner, *A quasi-symmetric function for polymatroids*, to appear in the European Journal of Combinatorics, `arXiv:math/0606646`

[4] C. J. Brianchon, *Théorème nouveau sur les polyèdres convexes*, J. École Polytechnique **15** (1837), 317–319.

[5] T. H. Brilawski, *The Tutte-Grothendieck ring*, Algebra Universalis **2** (1972), 375–388.

[6] H. Crapo, *Single-element extensions of matroids*, J. Res. Nat. Bur. Standards Sect. B **69B** (1965), 55–65.

[7] H. Crapo, *The Tutte polynomial*, Aequationes Math. **3** (1969), 211–229.

[8] H. Derksen, *Symmetric and quasi-symmetric functions associated to polymatroids*, preprint, `arXiv:0801.4393`.

[9] J. Edmonds, *Submodular functions, matroids, and certain polyhedra*, in: R. Guy, H. Hanani, N. Sauer, J. Schonheim (eds.), *Combinatorial Structures and Their Applications*, Gordon and Breach, New York, 1970, 69–87.

[10] I. M. Gel'fand, M. Goresky, R. MacPherson, V. Serganova, *Combinatorial geometries, convex polyhedra and Schubert cells*, Adv. in Math. **63** (1987), 301–316.

[11] J. P. Gram, *Om rumvinklerne i et polyeder*, Tidsskrift for Math. (Copenhagen) (3) **4** (1874), 161–163.

[12] H. Groemer, *On the extension of additive functionals on classes of convex sets*, Pacific J. Math. **75** (1978), no. 2, 397–410.

[13] J. Herzog, T. Hibi, *Discrete polymatroids*, J. Algebraic Combinatorics **16** (2002), no. 2, 239–268.

[14] L. Lafforgue, *Pavages des simplexes, schémas de graphes recollés et compactification des* $\mathrm{PGL}_r^{n+1}/\mathrm{PGL}_r$, Invent. Math. **136** (1999), no. 1, 233–271.

[15] L. Lafforgue, Chirurgie des Grassmanniennes, CRM Monograph Series **19**, AMS, Providence, RI, 2003.

[16] J. Morton, L. Pachter, A. Shiu, B. Sturmfels, O. Wienand, *Convex rank tests and semigraphoids*, `arXiv:math/0702564v2`.

[17] J. G. Oxley, *Matroid theory*, Oxford University Press, New York, 1992.

[18] A. Postnokiv, *Permutohedra, associahedra, and beyond*, Int. Math. Res. Notices (2009), doi:10.1093/imrn/rnn153.

[19] D. Speyer, *A matroid invariant via the K-theory of the Grassmannian*, to appear in Advances in Math., `arXiv:math/0603551`.

[20] W. Volland, *Ein Fortsetzungssatz für additive Eipolyederfunktionale im euklidischen Raum*, Arch. Math. **8** (1957), 144–149.

[21] D. J. A. Welch, *Matroid theory*, London Mathematical Society Monographs **8**, Academic Press, London, New York, 1976.

[22] H. Whitney, *On the abstract properties of linear independence*, Amer. J. of Math. **57** (1935), 509–533.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 162–173

A bijection between (bounded) dominant Shi regions and core partitions

Susanna Fishel[1] and Monica Vazirani[2]

[1] *School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-1804, USA*
[2] *University of California, Davis, Department of Mathematics, One Shields Ave, Davis, CA 95616-8633, USA*

Abstract.

It is well-known that Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ count the number of dominant regions in the Shi arrangement of type A, and that they also count partitions which are both n-cores as well as $(n+1)$-cores. These concepts have natural extensions, which we call here the m-Catalan numbers and m-Shi arrangement. In this paper, we construct a bijection between dominant regions of the m-Shi arrangement and partitions which are both n-cores as well as $(mn+1)$-cores.

We also modify our construction to produce a bijection between bounded dominant regions of the m-Shi arrangement and partitions which are both n-cores as well as $(mn-1)$-cores. The bijections are natural in the sense that they commute with the action of the affine symmetric group.

Résumé.

Il est bien connu que les nombres de Catalan $C_n = \frac{1}{n+1}\binom{2n}{n}$ comptent non seulement le nombre de régions dominantes dans le Shi arrangement de type A mais aussi les partitions qui sont à la fois n-coeur et $(n+1)$-coeur. Ces concepts ont des extensions naturelles, que nous appelons ici les nombres m-Catalan et le m-Shi arrangement. Dans cet article, nous construisons une bijection entre régions dominantes du m-Shi arrangement et les partitions qui sont à la fois n-coeur et $(nm+1)$-coeur.

Nous modifions également notre construction pour produire une bijection entre régions dominantes bornées du m-Shi arrangement et les partitions qui sont à la fois n-coeur et $(mn-1)$-coeur. Ces bijections sont naturelles dans le sens où elles commutent avec l'action du groupe affine symétrique.

Keywords: cores, symmetric group, Shi arrangement, Catalan numbers

1 Introduction

Let Δ be the root system of type A_{n-1}, with Weyl group W, and let m be a positive integer. Then let $\mathcal{S}_n^m$ be the arrangement of hyperplanes $H_{\alpha,k} = \{x \mid \langle \alpha \mid x \rangle = k \text{ for } -m+1 \leq k \leq m \text{ and } \alpha \in \Delta^+\}$. $\mathcal{S}_n^m$ is the mth extended Shi arrangement of type A_{n-1}, called here the m-Shi arrangement.

In Fishel and Vazirani (2010, 2009), the authors constructed and analyzed bijections between certain regions of $\mathcal{S}_n^m$ and certain n-cores. In this extended abstract, we summarize the results from both papers. We will first construct and discuss a bijection between dominant regions of $\mathcal{S}_n^m$, and partitions that are

n-cores as well as $(mn+1)$-cores. We will then modify the construction to give a direct bijection between *bounded* dominant regions and partitions which are simultaneously n-cores and $(mn-1)$-cores.

Our bijection is W-equivariant in the following sense. In each connected component of $\mathcal{S}_n^m$ there is exactly one m-minimal alcove, the alcove closest to the fundamental alcove $\mathcal{A}_0$, and in each bounded connected component there is exactly one m-maximal alcove, the alcove farthest from the fundamental alcove $\mathcal{A}_0$. Since the affine Weyl group W acts freely and transitively on the set of alcoves, there is a natural way to associate an element $w \in W = \widehat{\mathfrak{S}}_n$ to any alcove $w^{-1}\mathcal{A}_0$, and to the m-minimal and m-maximal alcoves in particular. There is also a natural action of $\widehat{\mathfrak{S}}_n$ on partitions, whereby the orbit of the empty partition $\emptyset$ is precisely the n-cores. We will show that when w is associated to an m-minimal alcove, then $w\emptyset$ is an $(mn+1)$-core as well as an n-cores and that all such $(mn+1)$-cores that are also n-cores can be obtained this way. We will also show the analogous result for m-maximal alcoves and $(mn-1)$-cores that are also n-cores.

Roughly speaking, to each n-core λ we can associate an integer vector $\vec{n}(\lambda)$ whose entries sum to zero, as in Garvan et al. (1990). When λ is also an $(mn+1)$-core, these entries satisfy certain inequalities. On the other hand, these are precisely the inequalities that describe when a dominant alcove is m-minimal. We λ is an $(mn-1)$-core, the inequalites which must be satisfied by the entries of the vector exactly describe when a dominant alcove is m-maximal.

As a consequence, we show an n-core λ is automatically an $(mn+1)$-core if $\varepsilon_i(\lambda) \leq m$ for all $0 \leq i < n$, where $\varepsilon_i(\lambda)$ counts the number of removable boxes of residue i. We also show the related result, that an n-core λ is automatically an $(mn-1)$-core if $\varphi_i(\lambda) \leq m$ for all $0 \leq i < n$, where $\varphi_i(\lambda)$ counts how many addable boxes of residue i the partition λ has.

The article is organized as follows. In Section 2 we introduce notation and recall facts about Coxeter groups, root systems of type A, and inversion sets for elements of the affine symmetric group. Section 3 explains how the position of $w^{-1}\mathcal{A}_0$ relative to our system of affine hyperplanes is captured by the action of w on affine roots and that m-minimality and m-maximality can each be expressed by certain inequalities on the entries of $w(0,0,\ldots,0)$. In Section 4 we review facts about core partitions and in particular remind the reader how to associate an element of the root lattice to each core. Our main theorems, the bijection between dominant regions of the m-Shi arrangement and special cores and the bijection between bounded dominant regions of the m-Shi arrangement and other special cores, is in Section 5. Section 6 describes the effect of a related bijection on m-minimal and m-maximal alcoves. In Section 7, we derive further results that refine our bijection between alcoves and cores and that involve Narayana numbers. We also characterize alcove walls in terms of addable and removable boxes.

2 Preliminaries

Please also see Fishel and Vazirani (2010, 2009). Let Δ be the root system for type A_{n-1}, with Weyl group the symmetric group $\mathfrak{S}_n$. Let $\widetilde{\Delta}$ be the affine root system of type $A_{n-1}^{(1)}$, with null root δ, and with Weyl group the affine symmetric group $\widehat{\mathfrak{S}}_n$. See Kac (1990) for more details. Δ spans a Euclidean space V with inner product $\langle\,|\,\rangle$. Let $Q \subseteq V$ denote the root lattice for Δ. Let m be a positive integer. The *m-Shi arrangement* is the collection of hyperplanes

$$\mathcal{S}_n^m = \{H_{\alpha,k} \mid \alpha \in \Delta^+, -m < k \leq m\},$$

where $H_{\alpha,k} = \{v \in V \mid \langle v \mid \alpha\rangle = k\}$. This arrangement can be defined for all types; here we are concerned with type A.

The arrangement dissects V into connected components we call regions. We refer to regions which are in the dominant chamber of V as *dominant* regions. Each connected component of $V \setminus \bigcup_{\substack{\alpha\in\Delta^+\\ k\in\mathbb{Z}}} H_{\alpha,k}$ is called an alcove and the *fundamental alcove* is denoted $\mathcal{A}_0$.

We denote the (closed) half spaces $H^+_{\alpha,k} = \{v \in V \mid \langle v \mid \alpha\rangle \geq k\}$ and $H^-_{\alpha,k} = \{v \in V \mid \langle v \mid \alpha\rangle \leq k\}$. Note $\mathcal{A}_0$ is the interior of $H^-_{\theta,1} \cap \bigcap_{i=1}^{n-1} H^+_{\alpha_i,0}$ and the dominant chamber is $\bigcap_{i=1}^{n-1} H^+_{\alpha_i,0}$.

The affine symmetric group $\widehat{\mathfrak{S}}_n$ acts on V (preserving Q) via affine linear transformations, and acts freely and transitively on the set of alcoves. We thus identify each alcove $\mathcal{A}$ with the unique $w \in \widehat{\mathfrak{S}}_n$ such that $\mathcal{A} = w\mathcal{A}_0$. We also note that we may express any $w \in \widehat{\mathfrak{S}}_n$ as $w = ut_\gamma$ for unique $u \in \mathfrak{S}_n$, $\gamma \in Q$, or equivalently $w = t_{\gamma'}u$ where $\gamma' = u(\gamma)$. If we embed V into $\mathbb{R}^n$ by mapping α_i to $\varepsilon_i - \varepsilon_{i+1}$, note that $\gamma' = w(0,\ldots,0)$.

We also remind the reader that when w^{-1} is a minimal length right coset representative for $\widehat{\mathfrak{S}}_n/\mathfrak{S}_n$, then we may write $w^{-1} = t_{\gamma'}u$ where $u \in \mathfrak{S}_n$ and γ' is in the dominant chamber.

For $w \in \widehat{\mathfrak{S}}_n$, we define the inversion set $\mathrm{Inv}(w) = \{\alpha \in \widetilde{\Delta^+} \mid w(\alpha) \in \widetilde{\Delta^-}\}$. Notice that the length $\ell(w) = |\mathrm{Inv}(w)|$ for $w \in \widehat{\mathfrak{S}}_n$ is just the minimal number of affine hyperplanes separating $w^{-1}\mathcal{A}_0$ from $\mathcal{A}_0$. We will need the following well-known proposition and corollary, both describing $\mathrm{Inv}(w)$ and both proved in Fishel and Vazirani (2010).

Proposition 2.1. *Let $w \in \widehat{\mathfrak{S}}_n$ and $\alpha + k\delta \in \widetilde{\Delta^+}$. Then $\alpha + k\delta \in \mathrm{Inv}(w)$ iff $w^{-1}\mathcal{A}_0 \subseteq H^+_{-\alpha,k}$*

Corollary 2.2. *Suppose w is a minimal length left coset representative for $\widehat{\mathfrak{S}}_n/\mathfrak{S}_n$. Then $\mathrm{Inv}(w)$ consists only of roots of the form $-\alpha + k\delta$, $k \in \mathbb{Z}_{>0}, \alpha \in \Delta^+$. Further, if $-\alpha + k\delta \in \mathrm{Inv}(w)$ and $k > 1$ then $-\alpha + (k-1)\delta \in \mathrm{Inv}(w)$.*

3 m-minimal and m-maximal alcoves

We say an alcove $w\mathcal{A}_0$ is m-minimal if it is the unique alcove in its region such that $\ell(w)$ is smallest. Such alcoves are termed "representative alcoves" by Athanasiadis. We can identify each connected component of the complement of the m-Shi arrangement with its unique m-minimal alcove.

If the region is bounded, we can also identify it with the unique alcove $w'\mathcal{A}_0$ contained in it such that $\ell(w')$ is largest. In this situation we will say the alcove $w'\mathcal{A}_0$ is *m-maximal.* Note that for unbounded regions, no such alcove exists.

See Figure 5 below for a picture of the m-maximal alcoves of type A_2 for $m = 1, 2$.

The following proposition is useful. For a given alcove, it characterizes the affine hyperplanes containing its walls and which simple reflections flip it over those walls (by the right action). It can be found in Shi (1987) in slightly different notation.

Proposition 3.1. *Suppose $w\mathcal{A}_0 \subseteq H^+_{\alpha,k}$ but $ws_i\mathcal{A}_0 \subseteq H^-_{\alpha,k}$*

1. *Then $w(\alpha_i) = \alpha - k\delta$.*
2. *Let $\beta = w^{-1}(0,\ldots,0) \in V$. Then $\langle \beta \mid \alpha_i\rangle = -k$.*

Using the coordinates of $V \subseteq \mathbb{R}^n$, we note $k = \gamma_{u(i)} - \gamma_{u(i+1)}$, where $w = t_\gamma u$.

Remark 3.2. *Note, if $w\mathcal{A}_0$ is m-minimal, then whenever $k \in \mathbb{Z}_{\geq 0}$ and $w\mathcal{A}_0 \subseteq H^+_{\alpha,k}$ but $ws_i\mathcal{A}_0 \subseteq H^-_{\alpha,k}$ then we must have $k \leq m$ in the case $\alpha > 0$ and $k \leq m-1$ in the case $\alpha < 0$.*

It is easy to see that this condition is not only necessary but sufficient to describe when $w\mathcal{A}_0$ is m-minimal. Together with Proposition 2.1, Proposition 3.1 says that when $\alpha_i \in \mathrm{Inv}(w)$ and $w(\alpha_i) = \alpha - k\delta$ then $k \leq m$, and for $\beta = w^{-1}(0, \ldots, 0)$ that $\langle \beta \mid \alpha_i \rangle \geq -m$.

Applying Remark 3.2 to positive α and alcoves in the dominant chamber, we get the following corollary.

Corollary 3.3. *Suppose $w\mathcal{A}_0$ is in the dominant chamber and m-minimal.*

1. *If $w\mathcal{A}_0 \subseteq H^+_{\alpha,k}$ but $ws_i\mathcal{A}_0 \subseteq H^-_{\alpha,k}$ for some $\alpha \in \Delta^+, k \in \mathbb{Z}_{\geq 0}$, then $k \leq m$.*
2. *Let $\beta = w^{-1}(0, \ldots, 0)$. Then $\langle \beta \mid \alpha_i \rangle \geq -m$, for all i, and in particular $\langle \beta \mid \theta \rangle \leq m + 1$.*

Proof. The first statement follows directly from Proposition 3.1 and Remark 3.2. To conclude that the second statement holds for all i, note that if $k \leq 0$ then automatically $k \leq m$. □

It is possible to make a remark analogous to Remark 3.2 for the case of m-maximal alcoves and we derive a corollary analogous to Corollary 3.3

Corollary 3.4. *Suppose $w\mathcal{A}_0$ is in the dominant chamber and m-maximal.*

1. *If $w\mathcal{A}_0 \subseteq H^-_{\alpha,k}$ but $ws_i\mathcal{A}_0 \subseteq H^+_{\alpha,k}$ for some $\alpha \in \Delta^+, k \in \mathbb{Z}_{\geq 0}$, then $k \leq m$.*
2. *Let $\beta = w^{-1}(0, \ldots, 0)$. Then $\langle \beta \mid \alpha_i \rangle \leq m$, for all i, and in particular $\langle \beta \mid \theta \rangle \geq -m + 1$.*

4 Core partitions and their abacus diagrams

In this section we review some well-known facts about n-cores and review the useful tool of the abacus construction. Details can be found in James and Kerber (1981).

There is a well-known bijection $\mathcal{C} : \{n\text{-cores}\} \to Q$ that commutes with the action of $\widehat{\mathfrak{S}}_n$. One can use the $\widehat{\mathfrak{S}}_n$-action to define the bijection, or describe it directly from the combinatorics of partitions via the work of Garvan-Kim-Stanton's $\vec{n}$-vectors in Garvan et al. (1990) or of Lascoux (2001), or as described in terms of balanced abaci as in Berg et al. (2009). Here, we will recall the description from Berg et al. (2009) as well as remind the reader of the $\widehat{\mathfrak{S}}_n$-action on n-cores.

We identify a partition $\lambda = (\lambda_1, \ldots, \lambda_r)$ with its Young diagram, the array of boxes with coordinates $\{(i, j) \mid 1 \leq j \leq \lambda_i\}$. We say the box $(i, j) \in \lambda$ has *residue* $j - i \bmod n$, and in that case, we often refer to it as a $(j - i \bmod n)$-box. Its *hook length* $h^\lambda_{(i,j)}$ is $1+$ the number of boxes to the right of and below (i, j).

An *n-core* is a partition λ such that $n \nmid h^\lambda_{(i,j)}$ for all $(i, j) \in \lambda$.

We say a box is *removable* from λ if its removal results in a partition. Equivalently its hook length is 1. A box not in λ is *addable* if its union with λ results in a partition.

Claim 4.1. *Let λ be an n-core. Suppose λ has a removable i-box. Then it has no addable i-boxes. Likewise, if λ has an addable i-box it has no removable i-boxes.*

$\widehat{\mathfrak{S}}_n$ acts transitively on the set of n-cores as follows. Let λ be an n-core. Then

$$s_i\lambda = \begin{cases} \lambda \cup \text{all addable } i\text{-boxes} & \exists \text{ any addable } i\text{-box}, \\ \lambda \setminus \text{all removable } i\text{-boxes} & \exists \text{ any removable } i\text{-box}, \\ \lambda & \text{else.} \end{cases}$$

It is easy to check $s_i\lambda$ is an n-core.

4.1 Abacus diagrams

We can associate to each partition λ its abacus diagram. When λ is an n-core, its abacus has a particularly nice form, and then can be used to construct an element of Q. Each partition $\lambda = (\lambda_1, \ldots, \lambda_r)$ is determined by its hook lengths in the first column, the $\beta_k = h^{\lambda}_{(k,1)}$.

An *abacus diagram* is a diagram, with entries from $\mathbb{Z}$ arranged in n columns labeled $0, 1, \ldots, n-1$, called *runners*. The horizontal cross-sections or rows will be called *levels* and runner k contains the entry labeled by $rn + k$ on level r where $-\infty < r < \infty$. We draw the abacus so that each runner is vertical, oriented with $-\infty$ at the top and ∞ at the bottom, and we always put runner 0 in the leftmost position, increasing to runner $n-1$ in the rightmost position. Entries in the abacus diagram may be circled; such circled elements are called *beads*. Entries which are not circled will be called *gaps*. We shall say two abaci are equivalent if they differ by adding a constant to all entries. Note, in this case we must cyclically permute the runners so that runner 0 is leftmost. Given a partition λ its abacus is any abacus diagram equivalent to the one obtained by placing beads at entries $\beta_k = h^{\lambda}_{(k,1)}$ and all $j \in \mathbb{Z}_{<0}$.

Remark 4.2. *It is well-known that λ is an n-core if and only if its abacus is flush, that is to say whenever there is a bead at entry j there is also a bead at $j - n$.*

We define the *balance number* of an abacus to be the sum over all runners of the largest level in that runner which contains a bead. We say that an abacus is *balanced* if its balance number is zero. Note that there is a unique abacus which represents a given n-core λ for each balance number. Given a flush abacus, that is, the abacus of an n-core λ, we can associate to it the vector whose i^{th} entry is the largest level in runner $i-1$ which contains a bead. The sum of the entries in this vector is the balance number of the abacus. When the abacus is balanced, we will call this vector $\vec{n}(\lambda)$, in keeping with the notation of Garvan et al. (1990). We note that $\vec{n}(\lambda) \in Q$, when we identify Q with $\{(a_1, \ldots, a_n) \in \mathbb{Z}^n \mid \sum_i a_i = 0\}$.

We recall the following claim, which can be found in Berg et al. (2009).

Claim 4.3. *The map $\lambda \mapsto \vec{n}(\lambda)$ is an $\widehat{\mathfrak{S}}_n$-equivariant bijection $\{n\text{-cores}\} \to Q$.*

We recall here results of Anderson (2002), which describe the abacus of an n-core that is also a t-core, for t relatively prime to n. When $t = mn - 1$, this takes a particularly nice form.

Proposition 4.4 (Anderson)**.** *Let λ be an n-core. Suppose t is relatively prime to n. Let $M = nt - n - t$. Consider the grid of points $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ with $0 \le x \le n-1$, $0 \le y$ labelled by $M - xt - yn$. Circle a point in this grid if and only if its label is obtained from the first column hooklengths of λ or its label is in $\mathbb{Z}_{<0}$. Then λ is a t-core if and only if the following three conditions hold.*

1. *All beads in the abacus of λ are at entries $\le M$, in other words at (x, y) with $0 \le x \le n-1$, $0 \le y$;*
2. *The circled points in the grid are upwards flush, in other words if (x, y) is circled, so is $(x, y-1)$;*
3. *The circled points in the grid are flush to the right, in other words if (x, y) is circled and $x \le n-2$, so is $(x+1, y)$.*

Note that the columns of this grid are exactly the runners of λ's abacus, written out of order, with each runner shifted up or down relative to its new left neighbor. The runners have also been truncated, which is irrelevant given condition (1) above. This shifting is performed exactly so labels in the same row are congruent $\bmod t$. This explains why the circles must be flush to the right as well as upwards flush.

We will now analyze the special cases $t = mn+1$ and $t = mn-1$ to derive the conditions for a n-core to be a $mn \pm 1$-core.

Corollary 4.5. *Let λ be an n-core.*

1. *Then λ is an $(mn+1)$-core if and only if $\langle \vec{n}(\lambda) \mid \alpha_i \rangle \geq -m$ for $0 < i < n$ and $\langle \vec{n}(\lambda) \mid \theta \rangle \leq m+1$.*
2. *Let λ be an n-core. Then λ is an $(mn-1)$-core if and only if $\langle \vec{n}(\lambda) \mid \alpha_i \rangle \leq m$ for $0 < i < n$ and $\langle \vec{n}(\lambda) \mid \theta \rangle \geq -m+1$.*

Proof. In the notation of Proposition 4.4, in the special case $t = mn+1$, the columns of the grid are the runners of λ's abacus, written in reverse order. Furthermore, each runner has been shifted m units down relative to its new left neighbor. So the condition of being flush to the right on Anderson's grid is given by requiring on the abacus that if the largest circled entry on runner $i+1$ is at level r then runner i must have a circled entry at level $r-m$. In other words, if $(a_1, \ldots, a_n) = \vec{n}(\lambda)$, then we require $a_i + m - a_{i+1} \geq 0$, i.e. $\langle \vec{n}(\lambda) \mid \alpha_i \rangle \geq -m$ for $0 < i < n$. Recall the 0^{th} and $(n-1)^{\mathrm{st}}$ and runners must also have this relationship (adding a constant to all entries in the abacus cyclically permutes the runners). This condition becomes $a_n + 1 + m - a_1 \geq 0$, i.e. $\langle \vec{n}(\lambda) \mid \theta \rangle \leq m+1$.

In the other special case, $t = mn-1$, the columns of the grid are now the runners of λ's abacus, cyclically shifted so the 0-runner is now rightmost versus leftmost. Otherwise, the analysis is the same. □

5 The bijection between cores and alcoves

Let Φ be the map

$$\begin{aligned} \{n\text{-cores}\} &\to \{\text{alcoves in the dominant chamber}\} \\ w\emptyset &\mapsto w^{-1}\mathcal{A}_0, \end{aligned}$$

which is $\widehat{\mathfrak{S}}_n$-equivariant, except for the minor technicality that the action on cores is a left action, but we take the right action on alcoves when discussing the Shi arrangement.

Theorem 5.1. *The map $\Phi : w\emptyset \mapsto w^{-1}\mathcal{A}_0$ for w a minimal length left coset representative of $\widehat{\mathfrak{S}}_n/\mathfrak{S}_n$ induces a bijection between the set of n-cores that are also $(mn+1)$-cores and the set of m-minimal alcoves, which are in the dominant chamber of V.*

Theorem 5.2. *The map $\Phi : w\emptyset \mapsto w^{-1}\mathcal{A}_0$ for w a minimal length left coset representative of $\widehat{\mathfrak{S}}_n/\mathfrak{S}_n$ induces a bijection from the set of n-cores that are also $(mn-1)$-cores to the set of m-maximal alcoves in the dominant chamber.*

The first bijection is pictured below in Figure 1 and the second in Figure 2.

6 A bijection on alcoves

Although they are not an ingredient in the main theorem of this paper, the following theorems build on the work of Section 3. They describe what the bijection $w\mathcal{A}_0 \mapsto w^{-1}\mathcal{A}_0$ does to the m-minimal and m-maximal alcoves. In particular, we do not limit ourselves to dominant alcoves.

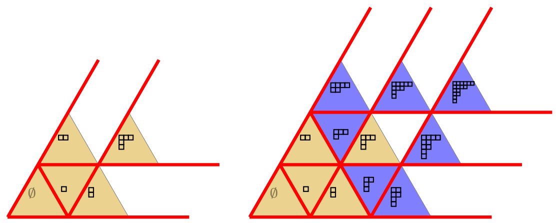


Fig. 1: m-minimal alcoves $w^{-1}\mathcal{A}_0$ in the dominant chamber of the m-Shi arrangement of type A_2, filled with the 3-core partition $w\emptyset$. On the left ($m = 1$), they are also 4-cores, and on the right ($m = 2$), they are also 7-cores.

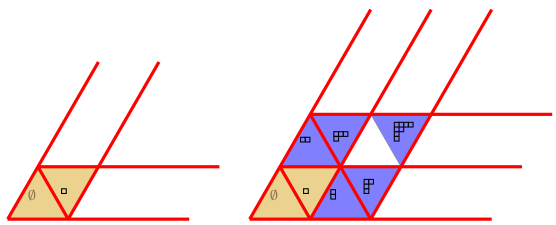


Fig. 2: m-maximal alcoves $w^{-1}\mathcal{A}_0$ in the dominant chamber of the m-Shi arrangement of type A_2, filled with the 3-core partition $w\emptyset$. On the left ($m = 1$), they are also 2-cores, and on the right ($m = 2$), they are also 5-cores.

6.1 Effect on m-minimal alcoves

Define $\mathfrak{A}_m$ to be the m-dilation of $\mathcal{A}_0$:

$$\mathfrak{A}_m = \{v \in V \mid \langle v \mid \alpha_i \rangle \geq -m, \, \langle v \mid \theta \rangle \leq m+1\}.$$

Note that the set of alcoves in $\mathfrak{A}_m$ is in bijection with $Q/(mn+1)Q$. Furthermore, it is easy to see by translating (by $m\rho = \frac{m}{2}\sum_{\alpha \in \Delta^+} \alpha$) that $Q \cap \mathfrak{A}_m$ is in bijection with $Q \cap \overline{(mn+1)\mathcal{A}_0}$. It is the latter that is discussed in Lemma 7.4.1 of Haiman (1994) and studied in Athanasiadis (2005) (technically for the co-root lattice $Q^\vee$). Taking the latter bijection into account, the second statement of Theorem 6.1 below appears in Theorem 4.2 of Athanasiadis (2005).

Theorem 6.1. *1. The map $w\mathcal{A}_0 \mapsto w^{-1}\mathcal{A}_0$ restricts to a bijection between alcoves in the region $\mathfrak{A}_m$ and m-minimal alcoves.*

2. The map $w(0, \ldots, 0) \mapsto w^{-1}\mathcal{A}_0$ restricts to a bijection between $Q \cap \mathfrak{A}_m$ and m-minimal alcoves in the dominant chamber.

Proof. Observe $\mathfrak{A}_m = H^-_{\theta,m+1} \cap \bigcap_{i=1}^{n-1} H^+_{\alpha_i,-m}$ can be viewed as an m-dilation of (the closure of) $\mathcal{A}_0 \subseteq H^-_{\theta,1} \cap \bigcap_{i=1}^{n-1} H^+_{\alpha_i,0}$.

The second statement follows directly from Corollary 3.3.

A proof of the first statement can be given that is very similar to that of Propositions 3.1 and 2.1. In Fishel and Vazirani (2010) we use those propositions to prove it.

□

The first part of the bijection is illustrated in Figures 3, and 4, by comparing Figure 4 to Figure 3.

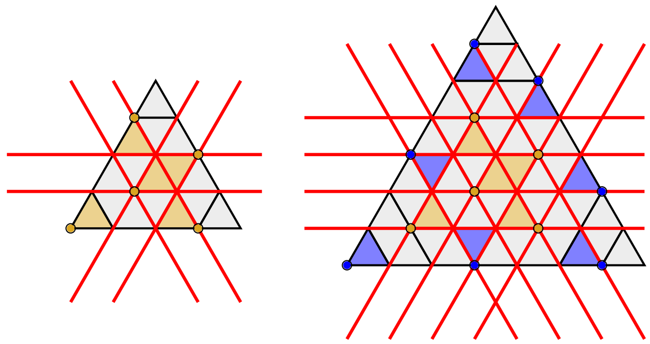


Fig. 3: $w\mathcal{A}_0$ for the m-minimal alcoves $w^{-1}\mathcal{A}_0$ in Figure 4 below, $m = 1, 2$. Note $w\mathcal{A}_0 \subseteq \mathfrak{A}_m$. Each $\gamma \in Q$ is in precisely one yellow/blue alcove, so this illustrates the second statement of Theorem 6.1.

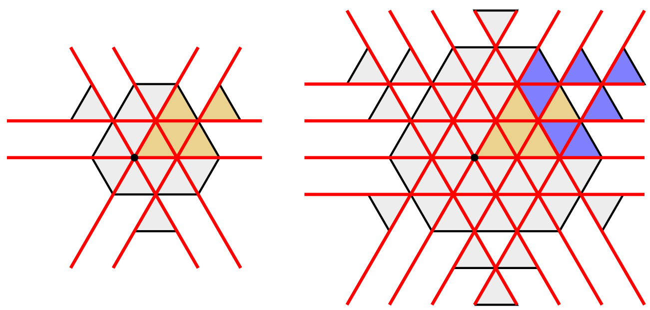


Fig. 4: m-minimal alcoves in the m-Shi arrangement for $m = 1$ ($m = 2$). Dominant alcoves are shaded yellow (and/or blue, respectively).

6.2 Effect on m-maximal alcoves

Let

$$\mathfrak{a}_m = \{v \in V \mid \langle v \mid \alpha_i \rangle \leq m \text{ for } 1 \leq i < n, \langle v \mid \theta \rangle \geq -m + 1\}.$$

Theorem 6.2. *1. The map $w\mathcal{A}_0 \mapsto w^{-1}\mathcal{A}_0$ restricts to a bijection between alcoves in the region $\mathfrak{a}_m$ and m-maximal alcoves.*

2. The map $w(0, \ldots, 0) \mapsto w^{-1}\mathcal{A}_0$ restricts to a bijection between $Q \cap \mathfrak{a}_m$ and m-maximal alcoves in the dominant chamber.

The proof is similar to the proof of Theorem 6.1.

The bijection is illustrated below, the first part comparing Figure 5 to Figure 6, and the second part from restricting our attention to the lattice points.

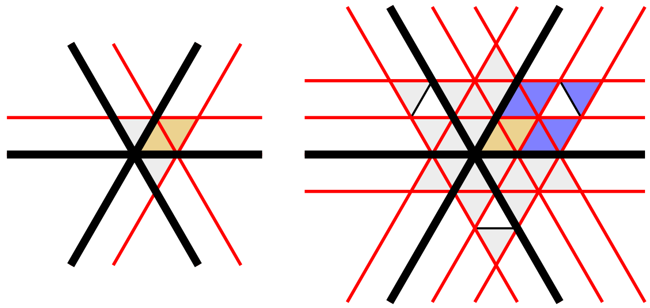


Fig. 5: m-maximal alcoves in the m-Shi arrangement for $m = 1$ ($m = 2$). Dominant alcoves are shaded yellow (and/or blue, respectively), whereas other m-maximal alcoves are shaded gray.

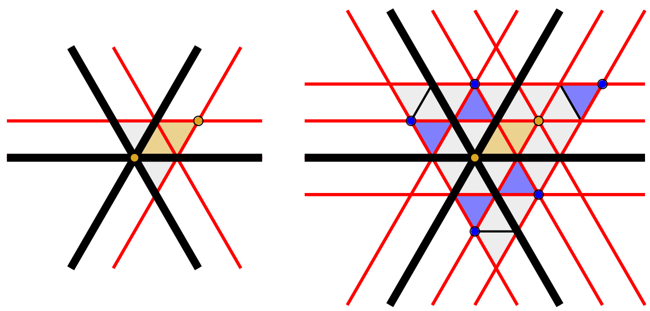


Fig. 6: $w\mathcal{A}_0$ for the m-maximal alcoves $w^{-1}\mathcal{A}_0$ in Figure 5 above, $m = 1, 2$. Note $\bigcup w\bar{\mathcal{A}}_0 = \mathfrak{a}_m$. Each $\gamma = w(0,\dots,0) \in Q \cap \mathfrak{a}_m$ is in precisely one yellow/blue alcove, so this illustrates the second statement of Theorem 6.2.

7 Alcove walls and addable and removable boxes for cores

In this section, we show how certain alcove walls correspond to addable and removable boxes in cores. We characterize the regions counted by the Narayana numbers in terms of their corresponding cores and explain an analagous result for bounded regions.

We will use some ideas from the theory of crystal graphs. For those readers familiar with the realization of the basic crystal $B(\Lambda_0)$ of $\widehat{\mathfrak{sl}}_n$ as having nodes parameterized by n-regular partitions,

$$s_i\lambda = \begin{cases} \tilde{f}_i^{\langle h_i, \mathrm{wt}(\lambda)\rangle}(\lambda) & \langle h_i, \mathrm{wt}(\lambda)\rangle \geq 0 \\ \tilde{e}_i^{-\langle h_i, \mathrm{wt}(\lambda)\rangle}(\lambda) & \langle h_i, \mathrm{wt}(\lambda)\rangle \leq 0, \end{cases}$$

where

$$\mathrm{wt}(\lambda) = \Lambda_0 - \sum_{(x,y)\in\lambda} \alpha_{y-x \bmod n}, \tag{7.1}$$

and h_i is the co-root corresponding to α_i.

Then the n-cores are exactly the $\widehat{\mathfrak{S}}_n$-orbit on the highest weight node, which is the empty partition $\emptyset$.

It is well-known that $s_i\lambda = \mu$ iff $s_i\mathrm{wt}(\lambda) = \mathrm{wt}(\mu)$ where the action of $\widehat{\mathfrak{S}}_n$ on the weight lattice is given by

$$s_i(\gamma) = \gamma - \langle \gamma \mid \alpha_i \rangle \alpha_i.$$

We refer the reader to Chapters 5,6 of Kac (1990) for details on the affine weight lattice, definition of Λ_0 and so on. For computational purposes, all we need remind the reader of is that $\langle \Lambda_0 \mid \alpha_i \rangle = \delta_{i,0}$ and $\langle \alpha_0 \mid \alpha_i \rangle = 2\delta_{i,0} - \delta_{i,1} - \delta_{i,n-1}$.

It is useful to recall the following notation from the theory of crystal graphs. In the case s_i removes k boxes of residue i from the core λ, write $\varepsilon_i(\lambda) = k, \varphi_i(\lambda) = 0$. In the case s_i adds r boxes to λ to obtain μ, write $\varepsilon_i(\lambda) = 0, \varphi_i(\lambda) = r$.

7.1 Narayana numbers

In this section, we add another set to the list in Theorem 1.2 of Athanasiadis (2005) of combinatorial objects counted by generalized Narayana numbers. We further refine the enumeration of n-cores λ which are also $(mn+1)$-cores. This refinement produces the m-Narayana numbers, or generalized Narayana numbers, $N_n^m(k)$, which are defined in Definition 7.4 below. Recall that the (k,l)-box of the n-core λ is referred to as an i-box if it has residue $i = l - k \bmod n$. Our refinement here is to count the number of n-cores λ which are also $(mn+1)$-cores by the number of residues i such that λ has exactly m removable i-boxes.

Remark 7.1. *Equation* (7.1) *says that if s_i removes k boxes (of residue i) from λ, or adds $-k$ boxes to λ to obtain μ, then* $\mathrm{wt}(\mu) = s_i(\mathrm{wt}(\lambda)) = \mathrm{wt}(\lambda) - k\alpha_i$. *In either case,* $\mathrm{wt}(\mu) = \mathrm{wt}(\lambda) + (\varphi_i(\lambda) - \varepsilon_i(\lambda))\,\alpha_i$.

A straightforward rephrasing of Proposition 3.1 is then:

Proposition 7.2. *Let λ be an n-core, $k \in \mathbb{Z}_{>0}$, and $w \in \widehat{\mathfrak{S}}_n$ of minimal length such that $w\emptyset = \lambda$. Fix $0 \le i < n$. The following are equivalent:*

1. *λ has k many removable i-boxes; in particular $|s_i\lambda| = |\lambda| - k$ as the action of s_i removes those i-boxes.*
2. *$\langle \vec{n}(\lambda) \mid \alpha_i \rangle = -k$ for $i \neq 0$, $\quad \langle \vec{n}(\lambda) \mid \theta \rangle = k+1$ for $i = 0$,*
3. *$w^{-1}\mathcal{A}_0 \subseteq H^+_{\alpha,k}$, $\quad w^{-1}s_i\mathcal{A}_0 \subseteq H^-_{\alpha,k}$ where $w^{-1}(\alpha_i) = \alpha - k\delta$.*

When we rephrase Corollary 3.3 in this context, it says:

Proposition 7.3 (Corollary 3.3 restated)**.** *Suppose $\lambda = w\emptyset$ is the n-core associated to the dominant alcove $\mathcal{A} = w^{-1}\mathcal{A}_0$ via the bijection Φ of Section 5.*

Then $\mathcal{A}$ is m-minimal if and only if whenever λ has exactly k removable boxes of residue i then $k \le m$. (And in this case, λ is also an $(mn+1)$-core.)

7.1.1 A refinement

Proposition 7.2 thus gives us another combinatorial interpretation of the m-Narayana numbers, as in Athanasiadis (2005).

Definition 7.4. *The k^{th} m-Narayana number of type A is*

$$N_n^m(k) = \frac{1}{nm+1}\binom{n-1}{n-k-1}\binom{mn+1}{n-k}.$$

Recall from Athanasiadis (2005) that $N_n^m(k)$ is the number of dominant regions of the m-Shi arrangement which have exactly k hyperplanes $H_{\alpha,m}$ separating them from $\mathcal{A}_0$ such that $H_{\alpha,m}$ contains a wall of the region.

In other words, for fixed k, we count how many m-minimal alcoves $\mathcal{A} = w^{-1}\mathcal{A}_0$ satisfy that for exactly k positive roots α, there exists an i such that $w^{-1}\mathcal{A}_0 \subseteq H^+_{\alpha,m}$ but $w^{-1}s_i\mathcal{A}_0 \subseteq H^-_{\alpha,m}$. It is clear that

$$\sum_{k\geq 0} N_n^m(k) = m\text{-Catalan number}$$

since each dominant m-minimal alcove gets counted once.

By Proposition 7.2 above, $N_n^m(k)$ equivalently counts how many n-cores λ that are also $(mn+1)$-cores have exactly k distinct residues i such that λ has precisely m removable i-boxes.

Corollary 7.5. *Let $N_n^m(k)$ denote the m-Narayana number of type A_{n-1}. Then*

$$\begin{aligned} N_n^m(k) = |\{\lambda \mid \lambda \textit{ is an } n\textit{-core and } (mn+1)\textit{-core and } \exists K \subseteq \mathbb{Z}/n\mathbb{Z} \\ \textit{with } |K| = k \textit{ such that } \lambda \textit{ has exactly } m \textit{ removable boxes} \\ \textit{of residue } i \textit{ iff } i \in K\}|. \end{aligned}$$

7.2 Bounded regions

We rephrase Equation (7.1) and Proposition 3.1 again, this time for m-maximal alcoves and bounded regions.

Recall Remark 7.1. In the context of m-maximal alcoves and bounded regions, Proposition 3.1 becomes

Proposition 7.6. *Let λ be an n-core, $k \in \mathbb{Z}_{>0}$, and $w \in \widehat{\mathfrak{S}}_n$ of minimal length such that $w\emptyset = \lambda$. Fix $0 \leq i < n$. The following are equivalent*

1. $\varphi_i(\lambda) = k$,
2. $\langle \vec{n}(\lambda) \mid \alpha_i \rangle = k$ *for* $i \neq 0$, $\quad \langle \vec{n}(\lambda) \mid \theta \rangle = -k+1$ *for* $i = 0$,
3. $w^{-1}\mathcal{A}_0 \subseteq H^-_{\alpha,k}$, $\quad w^{-1}s_i\mathcal{A}_0 \subseteq H^+_{\alpha,k}$ *where* $w^{-1}(\alpha_i) = -\alpha + k\delta$.

When we rephrase Corollary 3.4 in this context, it says:

Suppose $\lambda = w\emptyset$ is the n-core associated to the dominant alcove $\mathcal{A} = w^{-1}\mathcal{A}_0$. Then $\mathcal{A}$ is m-maximal iff whenever λ has exactly k addable boxes of residue i then $k \leq m$. (And in this case, λ is also an $(mn-1)$-core.)

As a consequence, note an n-core λ is automatically an $(mn-1)$-core if $\varphi_i(\lambda) \leq m$ for all $0 \leq i < n$.

Athanasiadis and Tzanaki (2006) define $h_k^+(\Delta, m)$, $0 \leq k < n$ as the number of bounded dominant regions of $\mathcal{S}_n^m$ for which exactly $n-1-k$ hyperplanes of the form $H_{\alpha,m}$, $\alpha \in \Delta^+$ are walls (i.e. support a facet) of that region and do *not* separate it from the fundamental alcove $\mathcal{A}_0$.

By the definition of m-maximal, we can replace a bounded region by its unique m-maximal alcove and consider its walls instead. In other words, to calculate $h_k^+(\Delta, m)$, we count how many m-maximal alcoves $\mathcal{A} = w^{-1}\mathcal{A}_0$ satisfy that for exactly $n-1-k$ positive roots α, there exists an i such that $w^{-1}\mathcal{A}_0 \subseteq H^-_{\alpha,m}$ but $w^{-1}s_i\mathcal{A}_0 \subseteq H^+_{\alpha,m}$.

By Proposition 7.6 above, $h_k^+(\Delta, m)$ equivalently counts how many n-cores λ that are also $(mn-1)$-cores have exactly $n-1-k$ distinct residues i such that λ has precisely m addable i-boxes.

Acknowledgements

Both authors wish to thank AIM and the SQuaREs program where this work was started. We thank Nathan Reading who pointed out that the number of bounded dominant regions in the complement of the m-Shi arrangement is counted by $\frac{1}{s+t}\binom{s+t}{t}$ for appropriate s, t, and by Anderson (2002) this number also counts the number of partitions that are s-cores as well as t-cores. We thank Mark Haiman for pointing out parts of Section 6 to us.

References

J. Anderson. Partitions which are simultaneously t_1- and t_2-core. *Discrete Math.*, 248(1-3):237–243, 2002. ISSN 0012-365X.

C. A. Athanasiadis. On a refinement of the generalized Catalan numbers for Weyl groups. *Trans. Amer. Math. Soc.*, 357(1):179–196 (electronic), 2005. ISSN 0002-9947.

C. A. Athanasiadis and E. Tzanaki. On the enumeration of positive cells in generalized cluster complexes and Catalan hyperplane arrangements. *J. Algebraic Combin.*, 23(4):355–375, 2006. ISSN 0925-9899.

C. Berg, B. Jones, and M. Vazirani. A bijection on core partitions and a parabolic quotient of the affine symmetric group. *J. Combin. Theory Ser. A*, 116(8):1344–1360, 2009. ISSN 0097-3165.

S. Fishel and M. Vazirani. A bijection between bounded dominant Shi regions and core partitions, 2009.

S. Fishel and M. Vazirani. A bijection between dominant Shi regions and core partitions. *European J. Combin.*, 2010. doi: 10.1016/j.ejc.2010.05.014.

F. Garvan, D. Kim, and D. Stanton. Cranks and t-cores. *Invent. Math.*, 101(1):1–17, 1990. ISSN 0020-9910.

M. D. Haiman. Conjectures on the quotient ring by diagonal invariants. *J. Algebraic Combin.*, 3(1):17–76, 1994. ISSN 0925-9899.

G. James and A. Kerber. *The representation theory of the symmetric group*, volume 16 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Co., Reading, Mass., 1981. ISBN 0-201-13515-9. With a foreword by P. M. Cohn, With an introduction by Gilbert de B. Robinson.

V. G. Kac. *Infinite-dimensional Lie algebras*. Cambridge University Press, Cambridge, third edition, 1990. ISBN 0-521-37215-1; 0-521-46693-8.

A. Lascoux. Ordering the affine symmetric group. In *Algebraic combinatorics and applications (Gößweinstein, 1999)*, pages 219–231. Springer, Berlin, 2001.

J. Y. Shi. Alcoves corresponding to an affine Weyl group. *J. London Math. Soc. (2)*, 35(1):42–55, 1987. ISSN 0024-6107.

Linear Systems on Tropical Curves

Christian Haase[1]†, Gregg Musiker[2]‡, and Josephine Yu[3]

[1] *Math. Inst., FU Berlin*
[2] *Massachusetts Institute of Technology, Department of Mathematics, Cambridge, MA*
[3] *Georgia Institute of Technology, School of Mathematics, Atlanta, GA*

Abstract. A tropical curve Γ is a metric graph with possibly unbounded edges, and tropical rational functions are continuous piecewise linear functions with integer slopes. We define the complete linear system $|D|$ of a divisor D on a tropical curve Γ analogously to the classical counterpart. We investigate the structure of $|D|$ as a cell complex and show that linear systems are quotients of tropical modules, finitely generated by vertices of the cell complex. Using a finite set of generators, $|D|$ defines a map from Γ to a tropical projective space, and the image can be modified to a tropical curve of degree equal to $\deg(D)$. The tropical convex hull of the image realizes the linear system $|D|$ as a polyhedral complex.

Résumé. Une courbe tropicale Γ est un graphe métrique pouvant contenir des arêtes infinies, et une fonction rationnelle tropicale est une fonction continue linéaire par morceaux à pentes entières. Le système linéaire complet $|D|$ d'un diviseur D sur une courbe tropicale Γ est défini de façon analogue au cas classique. Nous étudions la structure de $|D|$ en tant que complexe cellulaire et montrons que les systèmes linéaires sont des quotients de modules tropicaux engendrés par un nombre fini de sommets du complexe. Etant donné un ensemble fini de générateurs, $|D|$ définit une application de Γ vers un espace projectif tropical, dont l'image peut être modifiée en une courbe tropicale de degré égal à $\deg(D)$. L'enveloppe convexe tropicale de l'image réalise le système linéaire $|D|$ en tant que complexe polyédral.

Keywords: tropical curves, divisors, linear systems, canonical embedding, chip-firing games, tropical convexity

1 Introduction

An abstract tropical curve Γ is a connected metric graph with possibly unbounded edges. A *divisor* D on Γ is a formal finite $\mathbb{Z}$-linear combination $D = \sum_{x\in\Gamma} D(x) \cdot x$ of points of Γ. The *degree* of a divisor is the sum of the coefficients, $\sum_x D(x)$. The divisor is *effective* if $D(x) \geq 0$ for all $x \in \Gamma$; in this case we write $D \geq 0$. We call $\mathrm{supp}(D) = \{x \in \Gamma \,:\, D(x) \neq 0\}$ the support of the divisor D.

A *(tropical) rational function* f on Γ is a continuous function $f : \Gamma \to \mathbb{R}$ that is piecewise-linear on each edge with finitely many pieces and integral slopes. The *order* $\mathrm{ord}_x(f)$ of f at a point $x \in \Gamma$ is the sum of outgoing slopes at x. The *principal divisor* associated to f is

$$(f) := \sum_{x\in\Gamma} \mathrm{ord}_x(f) \cdot x.$$

A point $x \in \Gamma$ is called a *zero* of f if $\mathrm{ord}_x(f) > 0$ and a *pole* of f if $\mathrm{ord}_x(f) < 0$. We call two divisors D and D' linearly equivalent and write $D \sim D'$ if $D - D' = (f)$ for some f. For any divisor D on Γ, let $R(D)$ be the set of all rational functions f on Γ such that the divisor $D + (f)$ is effective, and $|D| = \{D' \geq 0 : D' \sim D\}$, the *linear system* of D. Let $\mathbb{1}$ denote the set of constant functions on Γ.

The set $R(D)$ is naturally embedded in the set $\mathbb{R}^\Gamma$ of all real-valued functions on Γ, and $|D|$ is a subset of the d^{th} symmetric product of Γ where $d = \deg(D)$. The map $R(D)/\mathbb{1} \to |D|$ given by $f \mapsto D + (f)$

†Supported by Emmy Noether grant HA 4383/1 of the German Research Foundation (DFG).
‡,3 Supported by an NSF Mathematical Sciences Postdoctoral Fellowship.

is a homeomorphism from $R(D)/\mathbb{1}$ to $|D|$. It was shown in [GK08, MZ06] that $|D|$ is a cell complex, so is $R(D)/\mathbb{1}$. Our aim is to study the combinatorial and algebraic structure of this object $R(D)$.

In Section 2 we give definitions and state linear equivalence in terms of weighted chip firing moves, which are continuous analogues of the chip firing games on finite graphs. In Section 3 we show that $R(D)$ is a finitely generated tropical semi-module and describe a generating set. In Section 4, we study the cell complex structure of $|D|$. We show that the vertex set of $|D|$ coincides with the generating set of $R(D)$ described in Section 3. We give a triangulation of the link of each cell as the order complex of a poset of possible weighted chip firing moves.

Any finite set $\mathcal{F}$ of linearly equivalent divisors induces a map $\phi_{\mathcal{F}}$ from the abstract curve to a tropical projective space. This map is described in Section 5. If $\mathcal{F}$ generates $R(D)$, we show that the tropical convex hull of the image of this map is homeomorphic to $|D|$. The image of this map $\phi_{\mathcal{F}}$ can be naturally modified to an embedded tropical curve.

2 Metric graphs, rational functions, and chip-firing

A *metric graph* Γ is a complete connected metric space such that each point $x \in \Gamma$ has a neighborhood U_x isometric to a star-shaped set of valence $\mathrm{val}(x) \geq 1$ endowed with the path metric. To be precise, a star-shaped set of valence v is a set of the form

$$S(v,r) = \{z \in \mathbb{C} : z = te^{2\pi ik/v} \text{ for some } 0 \leq t < r \text{ and } k \in \mathbb{Z}\}.$$

The points $x \in \Gamma$ with valence different from 2 are precisely those where Γ fails to look locally like an open interval. Accordingly, we refer to a point of valence 2 as a *smooth* point.

Let $V(\Gamma)$ be any finite nonempty subset of Γ such that $V(\Gamma)$ contains all of the points with $\mathrm{val}(x) \neq 2$. Then $\Gamma \setminus V(\Gamma)$ is a finite disjoint union of open intervals. For a metric graph Γ, we say that a choice of such $V(\Gamma)$ gives rise to a *model* $G(\Gamma)$ for Γ. Each edge has a nonzero length inherited from the metric space Γ.

Let $V_0(\Gamma) = \{x \in \Gamma : \mathrm{val}(x) \neq 2\}$, where val denotes the valence of a vertex of $V(\Gamma)$. Unless Γ is a circle, $V_0(\Gamma)$ gives a model. For some of our applications, we may choose a model whose vertex set is strictly bigger than $V_0(\Gamma)$. However unless otherwise specified, the reader may assume that $G(\Gamma)$ denotes the coarsest model and that a *vertex* is an element of $V_0(\Gamma)$.

A *tropical curve* is a metric graph in which the leaf edges may have length ∞. A leaf edge is an edge adjacent to a one-valent vertex. Note that we add a "point at infinity" for each unbounded edge. A tropical rational function on a tropical curve may attain values $\pm\infty$ at points at infinity.

We will use the term *subgraph* in a topological sense, that is, as a compact subset of a tropical curve Γ with a finite number of connected components. For a subgraph $\Gamma' \subset \Gamma$ and a positive real number l, the *chip firing move* $\mathrm{CF}(\Gamma', l)$ by a (not necessarily connected) subgraph is the tropical rational function $\mathrm{CF}(\Gamma', l)(x) = -\min(l, \mathrm{dist}(x, \Gamma'))$. It is constant 0 on Γ', has slope -1 in the l-neighborhood of Γ' directed away from Γ', and it is constant $-l$ on the rest of the graph. We will sometimes refer to an effective divisor D as a *chip configuration*. For example, for $D = c_1 \cdot x_1 + \cdots + c_n \cdot x_n$, we say that there are c_i chips at the point $x_i \in \Gamma$. The total number of chips is the *degree* of the divisor. We say that a subgraph $\Gamma' \subset \Gamma$ can *fire* if for each boundary point of Γ' there are at least as many chips as the number of edges pointing out of Γ'. In other words, Γ' can fire if the divisor $D + (\mathrm{CF}(\Gamma', l))$ is effective for some positive real number l. The chip configuration $D + (\mathrm{CF}(\Gamma', l))$ is then obtained from D by moving one chip from the boundary of Γ' along each edge out of Γ' by distance l. Here we assume that l was chosen to be small enough so that the chips do not pass through each other or pass through a non-smooth point.

We will now show that these chip firing moves are enough to move between linearly equivalent divisors (Proposition 3 below). To this end, call a tropical rational function f a *weighted chip firing move* if there are two disjoint (not necessarily connected) proper closed subgraphs Γ_1 and Γ_2 such that the complement $\Gamma \setminus (\Gamma_1 \cup \Gamma_2)$ consists only of open line segments and such that f is constant on Γ_1 and Γ_2 and linear (smooth) with integer slopes on the complement.

A weighted chip firing move f can also be thought of as a combinatorial transformation that acts on chip configurations. Such transformations move chips from the boundary of Γ_2 along the open line segments in the complement $\Gamma \setminus (\Gamma_1 \cup \Gamma_2)$. (Here we assume w.l.o.g. that $f(\Gamma_2) > f(\Gamma_1)$.) During this process, a law of conservation of momentum holds so that a stack of m chips that move together will only move a distance of l/m. The numbers l and m can be different on each component of the complement. Note that a (simple) chip firing move $\mathrm{CF}(\Gamma', l)$ with small l is a special case of a weighted chip firing move when all the slopes are 0 or ± 1. The following two lemmas make the connection between $R(D)$ and chip firing games.

Lemma 1. *A weighted chip firing move is an (ordinary) sum of chip firing moves (plus a constant).*

Lemma 2. *Every tropical rational function is an (ordinary) sum of chip firing moves (plus a constant).*

Note that even if we start with a tropical rational function $f \in R(D)$, the sequence of weighted chip firing moves $f_1, \ldots, f_n$ for which $f = f_1 + \cdots + f_n$ may not be in $R(D)$, i.e. the divisors $D + (f_i)$ may not be effective although $D + (f)$ is. The following proposition follows easily from the two previous lemmas.

Proposition 3. *Two divisors are linearly equivalent if and only if one can be attained from the other using chip firing moves.*

Studying linear equivalence of divisors is partially motivated by a certain rank function satisfying tropical Riemann-Roch. In particular, the *rank* $r(D)$ of a divisor D is the maximum integer r such that $|D - E| \neq \emptyset$ for all degree-r divisors E. The Riemann-Roch Theorem [GK08, MZ06] (based on work of [BN07]) , which is the same for classical and tropical geometry, says that

$$r(D) - r(K - D) = \deg D + 1 - g, \tag{RR}$$

where g is the genus of tropical curve Γ, and the canonical divisor of Γ, K, is defined in Section 4.2.

3 Extremals and Generators of $R(D)$

The *tropical semiring* $(\mathbb{R}, \oplus, \odot)$ is the set of real numbers $\mathbb{R}$ with two tropical operations:

$$a \oplus b = \max(a, b), \text{ and } a \odot b = a + b.$$

The space $R(D)$ is naturally a subset of the space $\mathbb{R}^\Gamma$ of real-valued functions on Γ. For $f, g \in \mathbb{R}^\Gamma$, and $a \in \mathbb{R}$, the functions $f \oplus g$ and $a \odot f$ are defined by taking tropical sums and tropical products pointwise.

Lemma 4. *The space $R(D)$ is a tropical semi-module, i.e. it is closed under tropical addition and tropical scalar multiplication.*

Tropical semi-modules in $\mathbb{R}^n$ are also called *tropically convex sets* [DS04]. Since $R(D+(f)) = R(D)+f$, the tropical algebraic structure of $R(D)$ does not depend on the choice of the representative D. An element $f \in R(D)$ is called *extremal* if for any $g_1, g_2 \in R(D)$, $f = g_1 \oplus g_2 \implies f = g_1$ or $f = g_2$. Any generating set of $R(D)$ must contain all extremals up to tropical scalar multiplication.

Lemma 5. *A tropical rational function f is an extremal of $R(D)$ if and only if there are not two proper subgraphs Γ_1 and Γ_2 covering Γ (i.e. $\Gamma_1 \cup \Gamma_2 = \Gamma$) such that each can fire on $D + (f)$.*

A *cut set* of a graph Γ is a set of points $A \subset \Gamma$ such that $\Gamma \setminus A$ is not connected. A *smooth cut set* is a cut set consisting of smooth points (2-valent points). Note that being a smooth cut set depends only on the topology of Γ and is not affected by the choice of model $G(\Gamma)$.

Theorem 6. *Let $\mathcal{S}$ be the set of rational functions $f \in R(D)$ such that the support of $D + (f)$ does not contain a smooth cut set. Then*

(a) $\mathcal{S}$ contains all the extremals of $R(D)$,

(b) $\mathcal{S}$ is finite modulo tropical scaling, and

(c) $\mathcal{S}$ generates $R(D)$ as a tropical semi-module.

For the proof of (b) we need a boundedness lemma that improves the bound in [GK08, Lemma 1.8].

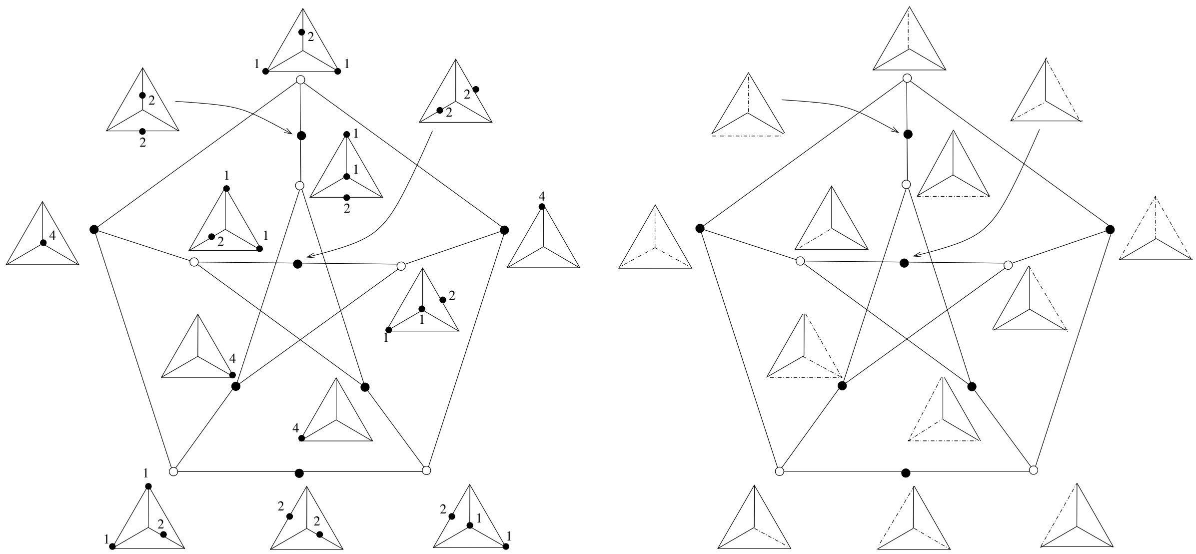


Fig. 1: (Left): The linear system $|K|$ for the tropical curve $\Gamma = K_4$, the complete graph on four vertices with edges of equal length, and the canonical divisor K. The 13 divisors shown here, together with K, correspond to the elements of $\mathcal{S}$ that generate $R(K)$, from Theorem 6. The seven black dots correspond to the extremals. See Example 10.

(Right): The link of the canonical divisor in the canonical class, where Γ is the complete graph on four vertices, with arbitrary edge lengths. This graph is also the order complex of the firing poset. The firing subgraphs in Γ are shown by solid lines. See Example 26. Compare with Figure 2 in [AK06].

Lemma 7. *For $D \geq 0$ every slope of $f \in R(D)$ is bounded by* $\deg D$.

of Theorem 6. (a) Suppose $f \notin \mathcal{S}$, then $D + (f)$ splits Γ into two subgraphs Γ_1 and Γ_2. Both of these graphs can fire, and the union of their closures is the entire Γ, so by Lemma 5, f is not an extremal.

(b) Let $f \in \mathcal{S}$. The support of $D + (f)$ meets the interior of each edge in at most one point, because two points on the same edge form a smooth cut set. Removing the set of edges meeting the support of $D + (f)$ does not disconnect Γ, and so the remaining edges contain a spanning tree of Γ. There are finitely many spanning trees in a graph and finitely many possible slopes for each edge in this spanning tree because of Lemma 7. Therefore, the number of possible values of f on vertices of Γ is finite modulo tropical scaling. (Here, vertices are non-smooth points. If Γ is a circle, then fix any point as a vertex.) On each non-tree edge, knowing the values and the slopes of f at the two end points uniquely determines f since all the chips of $D + (f)$ must fall on the same point of a given edge. We conclude that $\mathcal{S}$ is finite modulo tropical scaling.

(c) Let f be an arbitrary function in $R(D)$. We need to show that f can be written as a finite tropical sum of elements of $\mathcal{S}$. Let $N(f)$ be the number of smooth points in $\operatorname{supp}(D + (f))$. If f is not already in $\mathcal{S}$, then there is a smooth cut set A and two components Γ_1 and Γ_2. Let g_1 and g_2 be the weighted chip firing moves that fire all chips on their boundaries as far as possible. Then $f = (f + g_1) \oplus (f + g_2)$. Repeating this decomposition terminates after a finite number of steps because $0 \leq N(f+g_i) < N(f)$ for each $i = 1, 2$. □

Proposition 8. *Any finitely generated tropical sub-semimodule M of $\mathbb{R}^\Gamma$ is generated by the extremals.*

Corollary 9. *The tropical semimodule $R(D)$ is generated by the extremals. This generating set is minimal and unique up to tropical scalar multiplication.*

The set of extremals can be obtained from $\mathcal{S}$ by removing the elements not satisfying the condition in Lemma 5.

Example 10. Let Γ be a tropical curve with the complete graph on 4 vertices with equal edge lengths as a model. Consider the canonical divisor K, that is the divisor with value 1 on the four vertices and zero elsewhere. The canonical divisor is defined in general in Section 4.2. Then the set $\mathcal{S}$ from Theorem 6 consists of 14 elements, 7 of which are extremals. The linear system $|K|$ is the cone over the Petersen graph.

If the edge lengths of the complete graph are not all equal, then the set $\mathcal{S}$ may be different from this. We describe the local cell complex structure of $R(K)$ near K in the next section, in Example 20. See Figure 1.

4 Cell complex structure of $|D|$

As seen in the previous section, $R(D) \subset \mathbb{R}^\Gamma$ is finitely generated as a tropical semi-module or a tropical polytope. However, it is not a polyhedral complex in the ordinary sense. For example, let Γ be the line segment $[0,1]$, and D be the point 1. Then $R(D)$ is the tropical convex hull of $f, g \in \mathbb{R}^\Gamma$ where $f(x) = x$ and $g(x) = 0$. Although $R(D)$ is one-dimensional, it does not contain the usual line segment between any two points in it. Letting $\mathbb{1}$ denote the constant function taking the value 1 at all points, we consider functions in $R(D)$ modulo addition of $\mathbb{1}$, i.e. translation.

Lemma 11. *The set $R(D)/\mathbb{1}$ does not contain any nontrivial ordinary convex sets.*

Recall that $R(D)/\mathbb{1}$, i.e. $R(D)$ modulo tropical scaling can be identified with the linear system $|D| := \{D + (f) : f \in R(D)\}$ via the map $f \mapsto D + (f)$. In what follows, elements of $|D|$ and elements of projectivized $R(D)$, i.e. $R(D)/\mathbb{1}$, will be used interchangeably.

A choice of model $G(\Gamma)$ induces a polytopal cell decomposition of $\operatorname{Sym}^d \Gamma$, the d^{th} symmetric product of Γ. Andreas Gathmann and Michael Kerber [GK08] as well as Grigory Mikhalkin and Ilia Zharkov [MZ06] describe $|D|$ as a cell complex $|D|_{G(\Gamma)} \subset \operatorname{Sym}^d \Gamma$. Let us coordinatize this construction.

We identify each open edge $e \in E$ with the interval $(0, \ell(e))$ thereby giving the edge a direction, and we identify $\operatorname{Sym}^k e$ with the open simplex $\{x \in \mathbb{R}^k : 0 < x_1 < \ldots < x_k < \ell(e)\}$. A cell of $|D|$ is indexed by the following discrete data:

- $d_v \in \mathbb{Z}$ for every vertex $v \in V$,
- a composition (i.e. an ordered partition) $d_e = d_e^{(1)} + \cdots + d_e^{(r_e)}$ for every edge e of Γ, and
- an integer m_e for every edge e of Γ.

Then, a divisor D' belongs to that cell if

- $d_v = D'(v)$ for all $v \in V$,
- D' is given on e by $\sum_i d_e^{(i)} x_i$ for $0 < x_1 < \ldots < x_{r_e} < \ell(e)$, and
- the slope of f at the start of edge e is m_e, where f is such that $(f) + D = D'$.

The intersection of $|D|$ with an open cell of $\operatorname{Sym}^d \Gamma$ is a union of cells of $|D|$.

This cell complex structure depends on the choice of the model $G(\Gamma)$, but not on the choice of representative divisor D in the linear system $|D|$. In particular, choosing a finer model amounts to subdividing the cell complex $|D|$, and choosing a different divisor $D' = D + (g)$ amounts to changing the integer slopes at the starting points on the edges by the slopes of g, but this does not change the cells. Whenever we talk about a cell complex structure of $|D|$, we are impliciting assuming a model $G(\Gamma)$. Unless Γ is a circle, there is a unique coarsest model with the least number of vertices.

Example 12. Let Γ be a circle (for example a single vertex v with a loop edge e attached). Consider D to be the divisor $3v$. As we analyze in Example 17, $|D|$ contains two 2-cells in this case. The elements of both cells are divisors $D' = x + y + z$ with distinct points x, y, and z on the interior of e. However the two 2-cells differ from one another by the slope of the function f (defined by $D' = D + (f)$) at v. The outgoing slopes of f at v are given by $[-2, -1]$ for one 2-cell and by $[-1, -2]$ for the other. This example shows that the combinatorial type of the divisor D' – the cell of $\operatorname{Sym}^d \Gamma$ containing D' – does not determine the cell of $|D|$ containing D'. The different cells of $|D|$ in one cell of $\operatorname{Sym}^d \Gamma$ are indexed by the slopes of f.

Proposition 13. *For $D' \in |D|$, let $I_{D'}$ be the set of points in the support of D' that lie in the interior of edges. Then the dimension of the carrier of D' is one less than the number of connected components of $\Gamma \backslash I_{D'}$.*

Here, the carrier of D' is the cell containing D' in its interior. Recall that Γ is connected, and note that being in the interior of an edge depends on the model $G(\Gamma)$.

Theorem 14. *Let G be a model for Γ, and let $\mathcal{S}_G$ be the set of functions $f \in R(D)$ such that the support of $D + (f)$ does not contain an interior cut set (i.e. a cut set consisting of points in interior of edges in the model G). Then*

(a) $\mathcal{S}_G$ contains the set $\mathcal{S}$ from Theorem 6,

(b) $\mathcal{S}_G$ is finite modulo tropical scaling, and

(c) $\mathcal{S}_G = \{f \in R(D) : D + (f)$ is a vertex of $|D|\}$.

Proof. The statement (a) follows from definitions since points in the interior of edges (for any model) are smooth, and the statement (b) can be shown in the exact same way as Theorem 6(b). By the previous proposition, any element of $\mathcal{S}_G$ has dimension 0. This shows (c). ☐

This shows in particular that the cell complex $|D|$ has finitely many vertices. If the model G is the coarsest one, i.e. the vertices of G are non-smooth points of Γ, then $\mathcal{S}_G = \mathcal{S}$. If Γ is a circle, then there is no unique coarsest model.

Proposition 15. *Each closed cell in the cell complex is finitely-generated as a tropical semi-module by its vertices. In particular, it is tropically convex.*

Example 16. (Line Segment) Any tree is a genus zero tropical curve. Like genus zero algebraic curves, two divisors on a tree are linearly equivalent if and only if they have the same degree d. The simplest tree is a line segment consisting of an edge between two vertices, v_1 and v_2. In this case, $|D|$ is a d-simplex. The vertices of $|D|$ correspond to ordered pairs $[d_1, d_2]$ summing to d associated to the chip configuration at v_1 and v_2.

Example 17. (Circle) A circle is the only tropical curve where the canonical divisor K is 0. Let Γ be homeomorphic to a circle and let D be of degree 3. Then $D \sim 3x$ for some point $x \in \Gamma$. The coarsest cell structure of $R(D)$ is a triangle, but it is not realized by any model on Γ because Γ does not have a unique coarsest model. If the model contains only one vertex v and $D \sim 3v$, then $R(D)$ is a triangle subdivided by a median; see Figure 2. In particular $|D|$ contains four 0-cells, five 1-cells, and two 2-cells. If the model $G(\Gamma)$ consists of a vertex u such that $D \not\sim 3u$, then the cell complex structure would be different. If the model $G(\Gamma)$ consists of 3 equally spaced vertices v_1, v_2, v_3, and $D \sim 3v_1$, then $R(D)$ is isomorphic as a polyhedral complex to the barycentric subdivision of a triangle.

Example 18. (Circle with higher degree divisor) Let Γ be a circle graph with only a single vertex v and a single edge e, a loop based at v. Let $D = dv$; then the linear system $|D|$ is a cone over a cell complex, which we denote as $P_d(circle)$, which has an f-vector given by the following:

$$\text{The number of } i\text{−cells of } P_d(circle) = f_i = (i+1)\binom{d}{i+2}.$$

Consequently, the f-vector for $|D|$ is given by

$$\begin{cases} \binom{d}{2} + 1 \text{ if } i = 0 \\ (i+1)\binom{d}{i+2} + i\binom{d}{i+1} \text{ if } i \geq 1 \end{cases}.$$

To see how to get these f-vectors, we note that a divisor $D' \sim dv$ corresponds to a tropical rational function f such that $dv + (f) = D'$. One such f is the zero function, this corresponds to the cone point. Each other tropical rational function is parameterized by an increasing sequence of integer slopes $(a_1, \ldots, a_{i+2})$ such that $a_1 < 0$, $a_{i+2} > 0$, and $a_{i+2} - a_1 \leq d$. The first slope must be negative and the last slope must be

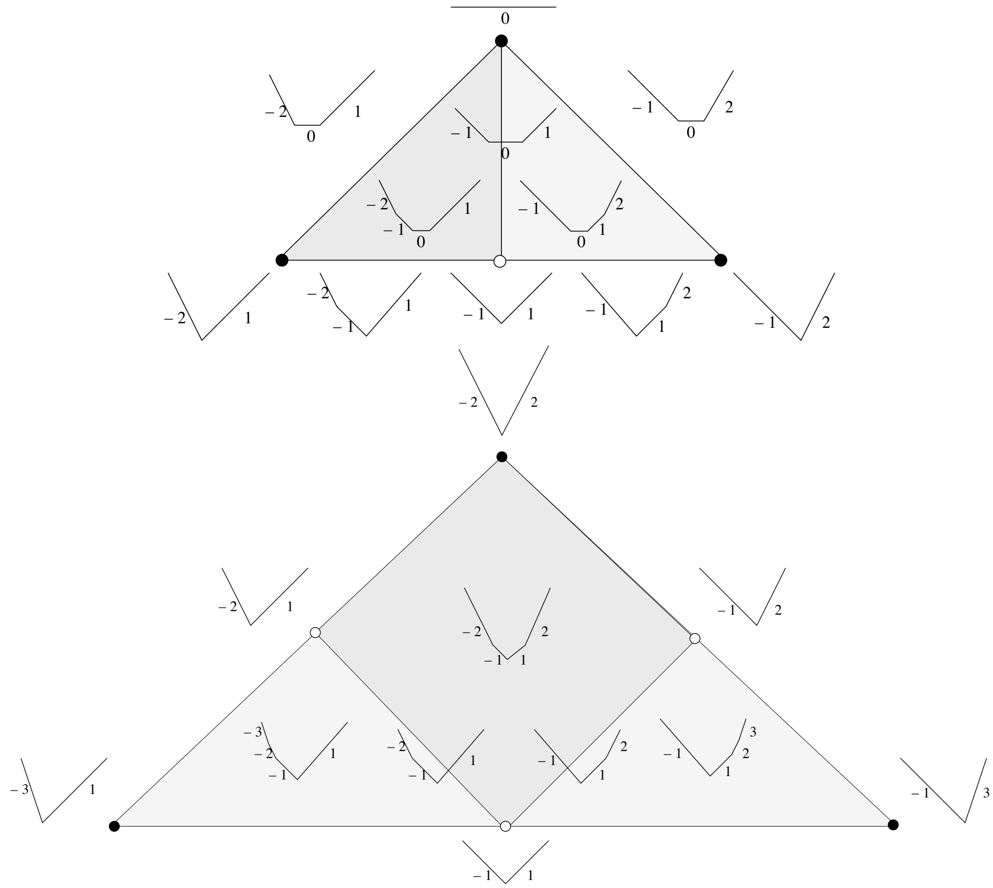


Fig. 2: (Top): The polyhedral cell complex $R(3v)/\mathbb{1}$ on $\Gamma = S^1$. The three black vertices are the extremals, and they correspond to the three divisors which are linearly equivalent to $3v$ and have the form $3w$. We have presented S^1 as the line segment $[0,1]$ with points 0 and 1 identified.

(Bottom): The polyhedral cell complex $R(4v)/\mathbb{1}$ on $\Gamma = S^1$ is a subdivided tetrahedron, a cone over this subdivided triangle with the cone-point corresponding to the constant function. (The labels of most 1-cells are suppressed, but may be read off from the incident vertices or 2-cells.) The cone-point plus the three black vertices are the extremals.

positive so that the values of f at the two ends of the loop e agree. The cells not incident to the cone point yield the cell complex $P_d(circle)$, and are given by sequences $(a_1, \dots, a_{i+2})$ such that all $a_i \neq 0$. To finish the computation of the f-vector for $P_d(circle)$, we pick an ordered pair $[j,k]$ with $j,k \geq 1$ and $j+k = i+2$ to denote the number of negative and positive a_k's, respectively. After setting $a_1 = -\ell$, we note that the number of ways to pick the remaining negative a_k's is given by $\binom{\ell-1}{j-1}$, and the number of ways to pick a subset of positive a_k's such that $a_{i+2} - a_1 \leq d$ is given by $\binom{d-\ell}{k}$. Summing over possible ℓ, and using a standard identity involving binomial coefficients (for instance see [BQ03, Identity 136]), we obtain $\binom{d}{i+2}$ such tropical rational functions for each $[j,k]$. Since there are $i+2$ such $[j,k]$'s, we get the above number of i-cells not incident to the cone point. For the case of $d=4$, see Figure 2.

Example 19. (Circle. Cell structure of $|D|$ as a simplex) In Examples 17 and 18, we saw that having to choose a model, even one with only one vertex, gives $|D|$ a cell structure of a subdivided simplex. Moreover, different choices of models, even if they contain only one vertex each, may give combinatorially different cell complex structures for $|D|$. We wish to describe $|D|$ as a simplex.

First, let us look at the embedding of $|D|$ in the symmetric product of the tropical curve. Let Γ be the circle

$\mathbb{R}/\mathbb{Z}$, and $D = d \cdot [0]$ be a divisor of degree d. The embedding of $|D|$ in $\mathrm{Sym}^d \Gamma = \mathrm{Sym}^d(\mathbb{R}/\mathbb{Z})$ is given by

$$\{x \in (0,1]^d \,:\, 0 < x_1 \leq x_2 \leq \ldots x_d \leq 1,\ \ x_1 + x_2 + \cdots + x_d \in \mathbb{Z}\}.$$

To see this, first consider a tropical rational function g on the line segment $[0,1]$ with $(g) = x_1 + x_2 + \cdots + x_d - d \cdot 0$ and $g(1) = 0$. Then $g(0) = x_1 + x_2 + \cdots + x_d$. If $g(0) \in \mathbb{Z}$, then adding g and a function l with constant slope $g(0)$ on $[0,1]$ gives a tropical rational function $f = g + l$ on the circle with $(f) + D = x_1 + x_2 + \cdots + x_d$. It is easy to check that any $f \in R(D)$ can be obtained this way. Although this description gives $|D|$ a uniform coordinate system, this does not give us a cell complex structure.

In fact, $|D|$ can be realized as a $(d-1)$-dimensional simplex, on d vertices. There is a unique set of d points $v_1, v_1, \ldots, v_d$ in Γ such that $D \sim dv_i$ for all $i = 1, \ldots, d$. These d points are equally spaced along Γ. The extremals of $R(D)$ are

$$\mathcal{E} = \{f \in R(D) : (f) + D = d \cdot v_i \text{ for some } i = 1, 2, \ldots, d\}.$$

Consider the $(d-1)$-dimensional simplex on vertices $V = \{dv_1, dv_2, \ldots, dv_d\}$, that is, the simplicial complex containing a $(k-1)$-dimensional cell for any k subset of V. We would like to stratify $|D|$ into these cells. For any divisor $D' \in |D|$, elements in the same cell as D' are obtained from D' by weighted chip firing moves that do not change the cyclically-ordered composition $d = a_1 + a_2 + \cdots + a_k$ associated to divisor $a_1x_1 + a_2x_2 + \cdots + a_kx_k$ where $x_1, x_2, \ldots, x_k$ are distinct and cyclically ordered along the circle (with a fixed orientation). The complement of the support of $D' = a_1x_1 + a_2x_2 + \cdots + a_kx_k$ consists of k segments. For each of these segments, there is a unique extremal in $R(D')$ that is maximal and constant on it. These k extremals of $R(D')$, which are naturally identified with extremals of $R(D)$, are precisely the vertices of the cell of D' and their convex hull is the cell of D'.

Example 20. (K_4 continued) As in Example 10, consider the graph K_4 with equal edge lengths and the canonical divisor K. The canonical divisor is defined in general in Section 4.2. The coarsest cell structure of $|K|$ consists of 14 vertices and topologically is the cone over the Petersen graph shown in Figure 1. The cone point is the canonical divisor K. The "cones" over the 3 subdivided edges of the Petersen graph are quadrangles. The maximal cells of $|K|$ consist of 12 triangles and 3 quadrangles. In particular, $|K|$ is not simplicial. The quadrangle obtained from "coning" over the bottom edge of the Petersen graph is shown in Figure 3.

4.1 Local structure of a cell complex

If B is a cell complex and x is a point in B, then the $\mathrm{link}(x, B)$ denotes the cell complex obtained by intersecting B with a sufficiently small sphere centered at x. We will define a triangulation of $\mathrm{link}(D, |D|)$ which is finer than the cell structure. Note that $|D|$ and $|D'|$ are isomorphic as cell complexes, so $\mathrm{link}(D, |D|) \cong \mathrm{link}(D, |D'|)$ for any $D' \sim D$.

Let $D' \in \mathrm{link}(D, |D|)$ and f be a rational function such that $D' = D + (f)$. Let $h_0 > h_1 > \cdots > h_n$ be the values taken on by f on the set of points that are either vertices of Γ or where f is not smooth. Notice that h_0 and h_n are maximum and minimum values of f, respectively. Since $D + (f) \in \mathrm{link}(D, |D|)$, we may assume that $h_0 - h_n$ is sufficiently small. Let $G = (\Gamma_0 \subset \Gamma_1 \subset \cdots \subset \Gamma_n = \Gamma)$ be a chain of subgraphs of Γ where $\Gamma_i = \{x \in \Gamma : f(x) \geq h_i\}$.

Let $G' = (\Gamma'_1 \subset \Gamma'_2 \subset \cdots \subset \Gamma'_n = \Gamma)$ be the chain of *compactified* graphs, where Γ'_i is the union of edges of Γ_i that are between two vertices of Γ. Each cell can be subdivided by specifying more combinatorial data: the chain G' obtained this way and the slopes at the non-smooth points. We call this the *fine subdivision*.

For an effective divisor D, we can naturally associate the *firing poset* $\mathcal{P}_D$ as follows. An element of $\mathcal{P}_D$ is a weighted chip firing move without the information about the length, i.e. it is a closed subgraph $\Gamma' \subset \Gamma$ together with an integer c_e for each out-going direction e of Γ' such that for each point $x \in \Gamma'$ we have $\sum c_e \leq D(x)$ where the sum on the left is taken over the all outgoing directions e from x and $D(x)$ denotes the coefficient of x in D. We say that $(\Gamma', c') \leq (\Gamma'', c'')$ if $\Gamma' \subset \Gamma''$ and $c'_e \geq c''_e$ for each common outgoing direction e of Γ' and Γ''.

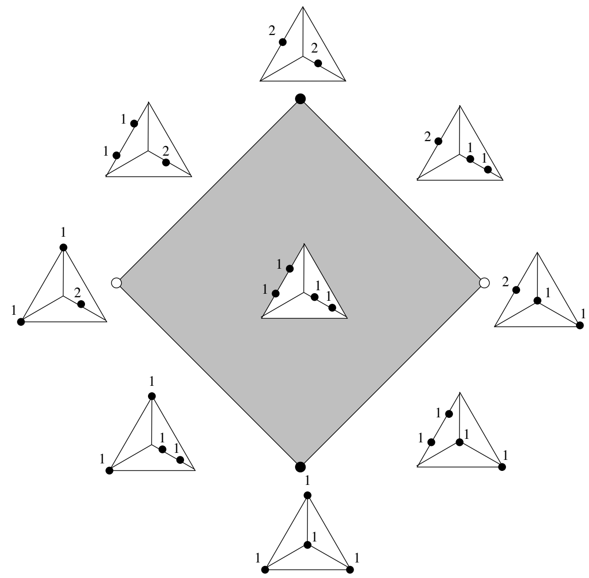


Fig. 3: A non-simplicial cell in the linear system $|K|$ for the complete graph on four vertices with edges of equal length.

Theorem 21. *The fine subdivision of the link of a divisor D in its linear system $|D|$ is a geometric realization of the order complex of the firing poset $\mathcal{P}_D$.*

Proof. By the discussion above, a cell in a fine subdivision $\mathrm{link}(D, |D|)$ corresponds to a unique chain in the firing poset. For any chain in the firing poset, we can construct an element in $\mathrm{link}(D, |D|)$ by performing the weighted chip firing moves in the order given by the chain, starting from the smallest element. The element constructed this way defines a cell in the fine subdivision. □

Note that the link of an element in $|D|$ does not depend on the precise location of the chips, but on the combinatorial data of the location. In other words, changing the edge lengths, without changing which edges the chips are on, does not affect the combinatorial structure of the link.

This Theorem, along with Proposition 13 allows us to explicitly describe the 1-cells incident to a 0-cell D' of $|D|$. For this, we need to define a specific subset of the weighted chip-firing moves. In particular, we call a weighted chip-firing move f (which is constant on Γ_1 and Γ_2) to be *doubly-connected* if Γ_1 and Γ_2 are both connected subgraphs.

Proposition 22. *Given $D' \in |D|$, and a model G such that $\mathrm{supp}(D') \subset V(G)$ (so that D' is a 0-cell in $|D|$), the 1-cells incident to D' correspond to the set of doubly-connected weighted chip-firing moves that are legal on chip configuration D' (up to combinatorial type).*

Proof. Let f be a weighted chip-firing move which is legal at D' that is constant on Γ_1 and Γ_2 such that $f(\Gamma_2) = f(\Gamma_1) - \epsilon$ for small $\epsilon > 0$. Then D'', defined as $D' + (f)$ has a chip on each of the line segments L_i connecting Γ_1 and Γ_2. Then the dimension of the corresponding cell of D'' is one if and only if Γ_1 and Γ_2 are both connected. □

4.2 Bergman subcomplex of $|K|$

Now we analyze the linear systems of an important family of divisors. The *canonical divisor* K on Γ is

$$K := \sum_{x \in \Gamma} (\mathrm{val}(x) - 2) \cdot x.$$

Vertices of valence two do not contribute to this sum so the divisor K is independent of the choice of model.

Let M be a matroid on a ground set E. The *Bergman fan* of M is the set of $w \in \mathbb{R}^E$ such that w attains its maximum at least twice on each circuit C of M. The only matroids considered here are *cographic matroids* of graphs. For a graph G with edge set E, the cographic matroid is the matroid on the ground set E whose dependent sets are cuts of G, i.e. the sets of edges whose complement is disconnected. The *Bergman complex* is the cell complex obtained by intersecting the Bergman fan with a sphere centered at the origin. The following result will be useful to us later.

Theorem 23. *[AK06]*

1. *The Bergman complex (with its fine subdivision) is a geometric realization of the order complex of the lattice of flats of M.*
2. *The Bergman fan is pure of codimension rank(M).*

Note that adding or removing parallel elements does not change the simplicial complex structure of the Bergman complex because the lattice of flats remains unchanged up to isomorphism. In particular, if G_1 and G_2 are two graphs, forming two models of the same tropical curve, then the corresponding cographic matroids have isomorphic Bergman complexes.

Lemma 24. *A subset of edges of a graph forms a flat of the cographic matroid if and only if its complement is a union of circuits of the graph.*

Suppose Γ has genus at least one but K_Γ is not effective. Let Γ' be the subgraph of Γ obtained by removing all the leaf edges recursively. Then the canonical divisor K' of Γ' is effective, and we can apply the following arguments for K' in Γ' or Γ.

Theorem 25. *The fine subdivision of* $\operatorname{link}(K, |K|)$ *contains the fine subdivision of the Bergman complex* $B(M^*(\Gamma))$ *as a subcomplex.*

Proof. The complement of a flat is a union of cocircuits, so the lattice of flats is isomorphic to the lattice of unions of cocircuits, ordered by reverse-inclusion. The cocircuits of the cographic matroid are the circuits of the graph. For the canonical divisor K, the proper union of circuits can always fire. Hence the proper part of the poset of union of circuits is a subposet of the firing poset, and so is the proper part of the lattice of flats. □

The Bergman complex may be a proper subcomplex of the link because there may be subgraphs that can fire on the canonical divisor but that are not union of circuits, e.g. two triangles connected by an edge in the graph of a triangular prism. Moreover, if Γ is not trivalent, there may be vertices that can fire more than one chip on each edge, so the firing poset may be strictly larger and so can the dimension of the order complex.

Example 26. (K_4 continued)

Let Γ be a tropical curve with the complete graph on four vertices as a model, with arbitrary edge lengths. Consider the canonical divisor K. In this case, the firing poset coincides with the lattice of unions of circuits, which is anti-isomorphic to the lattice of flats. Hence the link of the canonical divisor is isomorphic to the Bergman complex of the cographic matroid on the complete graph. Since the complete graph on four vertices is self-dual, its co-Bergman complex is the space of trees on five taxa, which is the Petersen graph [AK06].

See Figure 1. In the case when all edge lengths are equal, the quadrangles of $|K|$ described in Example 20 are subdivided in this fine subdivision of the $\operatorname{link}(K, |K|)$. Note that the link of the canonical divisor stays the same when we vary the edge lengths, while the generators and cell structure of $R(K)$ may change.

5 The induced map and projective embedding of a tropical curve

A finite set $\mathcal{F} = (f_1, \dots, f_r) \subset R(D)$ induces a map $\phi_{\mathcal{F}} : \Gamma \to \mathbb{TP}^{r-1}$, defined as $\phi_{\mathcal{F}}(x) = (f_1(x), \dots, f_r(x))$ for each $x \in \Gamma$. This is a map into $\mathbb{TP}^{r-1}$ rather than $\mathbb{R}^r$ as we take $\mathcal{F}$ to be defined up to translation by $\mathbb{1}$.

Theorem 27. *Let $\langle \mathcal{F} \rangle \subset R(D)$ be the tropical sub-semimodule of $R(D)$ generated by $\mathcal{F}$. Then $\langle \mathcal{F} \rangle / \mathbb{1}$ is homeomorphic to the tropical convex hull of the image of $\phi_{\mathcal{F}}$. In particular, if $\mathcal{F}$ generates $R(D)$, then $|D|$ is homeomorphic to the tropical convex hull of $\phi_{\mathcal{F}}(\Gamma)$.*

The *tropical convex hull* of a set is the tropical semi-module generated by the set.

Proof. The intuition behind this theorem is the result from [DS04] that the tropical convex hull of the rows of a matrix is isomorphic to the tropical convex hull of the columns. Here, the matrix $M_{\mathcal{F}}$ in question has entry $f_i(x)$ in row i and column x. As in [DS04], we define a convex set

$$P_{\mathcal{F}} = \{(y,z) \in (\mathbb{R}^r \times \mathbb{R}^\Gamma)/(\mathbb{1}, -\mathbb{1}) \,:\, y_i + z(x) \geq f_i(x)\}.$$

Let $B_{\mathcal{F}}$ be the union of bounded faces of $P_{\mathcal{F}}$, i.e. $B_{\mathcal{F}}$ contains points in the boundary of $P_{\mathcal{F}}$ that do not lie in the relative interior of an unbounded face of $P_{\mathcal{F}}$ in $(\mathbb{R}^r \times \mathbb{R}^\Gamma)/(\mathbb{1}, -\mathbb{1})$. We will show that $B_{\mathcal{F}}$ projects bijectively onto $\langle \mathcal{F} \rangle / \mathbb{1} \subset \mathbb{R}^\Gamma/\mathbb{1}$ on the one hand, and to $\operatorname{tconv} \phi_{\mathcal{F}}(\Gamma) \subset \mathbb{TP}^{r-1}$ on the other, establishing a homeomorphism. As in [DS04], we associate a *type* to $(y,z) \in P_{\mathcal{F}}$ as follows:

$$\operatorname{type}(y,z) := \{(i,x) \in [r] \times \Gamma : y_i + z(x) = f_i(x)\}.$$

In other words, a type is a collection of defining hyperplanes that contains (y,z), so elements in the relative interior of the same face have the same type. The recession cone of $P_{\mathcal{F}}$ is $\{(y,z) \in (\mathbb{R}^r \times \mathbb{R}^\Gamma)/(\mathbb{1}, -\mathbb{1}) : y_i + z(x) \geq 0\}$, which is the quotient of the positive orthant in $(\mathbb{R}^r \times \mathbb{R}^\Gamma)$ by $(\mathbb{1}, -\mathbb{1})$. Hence, a point $(y,z) \in P_{\mathcal{F}}$ lies in $B_{\mathcal{F}}$ if and only if we cannot add arbitrary positive multiples of any coordinate direction to it while staying in the same face of $P_{\mathcal{F}}$, which means keeping the same type. This holds if and only if

(1) The projection of $\operatorname{type}(y,z)$ onto $[r]$ is surjective, and

(2) The projection of $\operatorname{type}(y,z)$ onto Γ is surjective.

For $(y,z) \in P_{\mathcal{F}}$, these two conditions are equivalent respectively to

(1′) $y_i = \max\{f_i(x) - z(x) : x \in \Gamma\}$ for all $i \in [r]$, i.e. $y = M_{\mathcal{F}} \odot -z$, and

(2′) $z(x) = \max\{f_i(x) - y_i : i \in [r]\}$ for all $x \in \Gamma$, i.e. $z = -y \odot M_{\mathcal{F}}$.

where $M_{\mathcal{F}}$ is the $[r] \times \Gamma$ matrix with entry $f_i(x)$ in row i and column x, and $\odot$ is tropical matrix multiplication. These two conditions respectively imply that the projections of $B_{\mathcal{F}}$ onto $\mathbb{R}^\Gamma/\mathbb{1}$ and $\mathbb{R}^r/\mathbb{1}$ are one-to-one.

On the other hand, let $z \in \langle \mathcal{F} \rangle$, then $z = (u_1 \odot f_1) \oplus \cdots \oplus (u_r \odot f_r) = u \odot M_{\mathcal{F}}$ for some $u \in \mathbb{R}^r$ such that $z \geq u_i \odot f_i$ for each $i = 1, 2, \ldots, r$. Let $y \in \mathbb{R}^r$ such that $y_i = \min\{c : z \geq -c \odot f_i\}$ for $i = 1, 2, \ldots, r$; then $z = -y \odot M_{\mathcal{F}}$, so (y,z) satisfies (2′). Moreover, by construction, $-y_i \odot f_i(x) = z(x)$ for some x, so (y,z) satisfies (1). Thus $(y,z) \in B_{\mathcal{F}}$ and the set $B_{\mathcal{F}}$ projects surjectively onto $\langle \mathcal{F} \rangle / \mathbb{1} \subset \mathbb{R}^\Gamma/\mathbb{1}$. The image under the projection onto $\mathbb{R}^r/\mathbb{1}$ is the tropical convex hull of $\operatorname{image}(\phi_{\mathcal{F}})$, and the homeomorphism follows. □

Remark 28. All of the bounded faces of the convex set $P_{\mathcal{F}}$ are in fact vertices. If the union of bounded faces $B_{\mathcal{F}}$ contained a non-trivial line segment, then its projection $\langle \mathcal{F} \rangle / \mathbb{1}$ would as well, contradicting Lemma 11.

Example 29 (Circle, degree 3 divisor)**.** Let Γ be a circle of circumference 3, identified with $\mathbb{R}/3\mathbb{Z}$ and let D be the degree 3 divisor $[0] + [1] + [2]$. Let $f_0, f_1, f_2 \in R(D)$ be the extremals corresponding to divisors $3 \cdot [0], 3 \cdot [1]$, and $3 \cdot [2]$ respectively, and suppose $f_i([i]) = -1$ for each $i = 0, 1, 2$. Then the image of Γ under $\phi_{\mathcal{F}}$, for $\mathcal{F} = (f_0, f_1, f_2)$ is a union of three line segments between the points

$$\phi_{\mathcal{F}}([0]) = (-1,0,0), \quad \phi_{\mathcal{F}}([1]) = (0,-1,0), \quad \phi_{\mathcal{F}}([2]) = (0,0,-1) \quad \text{in } \mathbb{TP}^3.$$

In this case, the (max-) tropical convex hull of the image of $\phi_{\mathcal{F}}$ coincides with the usual convex hull and is a triangle. However, it is not the tropical convex hull of any proper subset of $\operatorname{image}(\phi_{\mathcal{F}})$. In particular, $|D|$ is not a tropical polytope, i.e. it is not the tropical convex hull of a finite set of points.

We know from [DS04] that tropically convex sets are contractible.

Corollary 30. *The sets $|D|$ and $R(D)$ are contractible.*

Tropical linear spaces are tropically convex [Spe08], so any tropical linear space containing the image $\phi_{\mathcal{F}}(\Gamma)$ must also contain its tropical convex hull.

Corollary 31. *Any tropical linear space in $\mathbb{TP}^{r-1}$ containing $\phi_{\mathcal{F}}(\Gamma)$ has dimension at least dim$(\langle \mathcal{F} \rangle)$.*

6 Conclusions and Open Questions

In this paper, we presented a number of properties of $|D|$ including verification that it is finitely generated as a tropical semi-module. We also provided some tools for explicitly understanding $|D|$ as a polyhedral cell complex such as a formula for the dimension of the face containing a given point, as well as applications such as using $|D|$ to embed an abstract tropical curve into tropical projective space.

There are many ways to continue this research for the future. It is quite tantalizing to investigate how the Baker-Norine rank of a divisor compares with the geometry and combinatorics of the associated linear system as a polyhedral cell complex. Also, is there any relation between $r(D)$ and the minimal number of generators of $R(D)$? How does the structure of $|D|$ change as we continuously move one point in the support of D or if we change the edge lengths of our metric graph while keeping the combinatorial type of the graph fixed?

In the case of finite graphs, i.e. divisors whose support lies within the set of vertices of the graph, can we combinatorially describe the associated linear systems? For example, is there a stabilization or an associated Ehrhart theory that one can use to count the sizes of such linear systems? Lastly, what other results from classical algebraic curve theory carry over to the theory of metric graphs (or tropical curves) and vice-versa?

Acknowledgements

The work was inspired by discussions at the Oberwolfach Workshop on Tropical Geometry that took place in December 2007. We thank Matt Baker, Michael Kerber, David Speyer, and Bernd Sturmfels for helpful discussions. We also thank the anonymous referees of the 2010 FPSAC conference and of Mathematische Zeitschrift for the helpful editorial suggestions. This is an abridged version of [HMY]. Please see [HMY] for omitted proofs and further topics including parameterized tropical curves, canonical embeddings, and tropical Jacobians.

References

[AK06] Federico Ardila and Caroline Klivans. "The Bergman complex of a matroid and phylogenetic trees," *Journal of Combinatorial Theory, Series B* **96** (2006) 38–49.

[BF06] Matthew Baker and Xander Faber. "Metrized Graphs, Laplacian Operators, and Electrical Networks," *Quantum graphs and their applications. Contemp. Math.,* **415**, Amer. Math. Soc., Providence, RI, (2006) 15-33.

[BN07] Matthew Baker and Serguei Norine. "Riemann-Roch and Abel-Jacobi theory on a finite graph," *Advances in Math.* **215** **(2)** (2007) 766-788.

[BQ03] Arthur Benjamin and Jennifer Quinn. *Proofs that really count. The art of combinatorial proof.* volume 27 of *The Dolciani Mathematical Expositions*. Mathematical Association of America, Washington, DC, 2003.

[DS04] Mike Develin and Bernd Sturmfels. Tropical convexity. *Doc. Math.*, 9:1–27, erratum 205–206, 2004.

[GK08] Andreas Gathmann and Michael Kerber. A Riemann-Roch theorem in tropical geometry. *Math. Z.*, 259(1):217–230, 2008.

[HMY] C. Haase, G. Musiker, and J. Yu, Linear systems on tropical curves, eprint, `arXiv:0909.3685`, 2009.

[MZ06] Grigory Mikhalkin and Ilia Zharkov. Tropical curves, their Jacobians and theta functions. *Curves and abelian varieties*, 203–230, Contemp. Math., 465, Amer. Math. Soc., Providence, RI, 2008.

[Spe08] David E Speyer. Tropical linear spaces. SIAM J. Discrete Math. **22** (2008), no. 4, 1527–1558.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 186–197

The biHecke monoid of a finite Coxeter group

Florent Hivert[1], Anne Schilling[2], and Nicolas M. Thiéry[2,3]

[1] *LITIS (EA 4108), Université de Rouen, Avenue de l'Université BP12 76801 Saint-Etienne du Rouvray, France and Institut Gaspard Monge (UMR 8049)*
[2] *Department of Mathematics, University of California, One Shields Avenue, Davis, CA 95616, U.S.A.*
[3] *Univ Paris-Sud, Laboratoire de Mathématiques d'Orsay, Orsay, F-91405; CNRS, Orsay, F-91405, France*

Abstract. For any finite Coxeter group W, we introduce two new objects: its cutting poset and its biHecke monoid. The cutting poset, constructed using a generalization of the notion of blocks in permutation matrices, almost forms a lattice on W. The construction of the biHecke monoid relies on the usual combinatorial model for the 0-Hecke algebra $H_0(W)$, that is, for the symmetric group, the algebra (or monoid) generated by the elementary bubble sort operators. The authors previously introduced the Hecke group algebra, constructed as the algebra generated simultaneously by the bubble sort and antisort operators, and described its representation theory. In this paper, we consider instead the *monoid* generated by these operators. We prove that it admits $|W|$ simple and projective modules. In order to construct the simple modules, we introduce for each $w \in W$ a combinatorial module T_w whose support is the interval $[1, w]_R$ in right weak order. This module yields an algebra, whose representation theory generalizes that of the Hecke group algebra, with the combinatorics of descents replaced by that of blocks and of the cutting poset.

Résumé. Pour tout groupe de Coxeter fini W, nous définissons deux nouveaux objets : son ordre de coupures et son monoïde de Hecke double. L'ordre de coupures, construit au moyen d'une généralisation de la notion de bloc dans les matrices de permutations, est presque un treillis sur W. La construction du monoïde de Hecke double s'appuie sur le modèle combinatoire usuel de la 0-algèbre de Hecke $H_0(W)$ *i.e.*, pour le groupe symétrique, l'algèbre (ou le monoïde) engendré par les opérateurs de tri par bulles élémentaires. Les auteurs ont introduit précédemment l'algèbre de Hecke-groupe, construite comme l'algèbre engendrée conjointement par les opérateurs de tri et d'anti-tri, et décrit sa théorie des représentations. Dans cet article, nous considérons le monoïde engendré par ces opérateurs. Nous montrons qu'il admet $|W|$ modules simples et projectifs. Afin de construire ses modules simples, nous introduisons pour tout $w \in W$ un module combinatoire T_w dont le support est l'intervalle $[1, w]_R$ pour l'ordre faible droit. Ce module détermine une algèbre dont la théorie des représentations généralise celle de l'algèbre de Hecke groupe, en remplaçant la combinatoire des descentes par celle des blocs et de l'ordre de coupures.

Keywords: Coxeter groups, Hecke algebras, representation theory, blocks of permutation matrices

1 Introduction

The usual combinatorial model for the 0-Hecke algebra $H_0(\mathfrak{S}_n)$ of the symmetric group is the algebra (or monoid) generated by the (anti) bubble sort operators $\pi_1, \ldots, \pi_{n-1}$, where π_i acts on words of length n and sorts the letters in positions i and $i+1$ decreasingly. By symmetry, one can also construct the bubble sort operators $\overline{\pi}_1, \ldots, \overline{\pi}_{n-1}$, where $\overline{\pi}_i$ acts by sorting increasingly, and this gives an isomorphic construction $\overline{H}_0$ of the 0-Hecke algebra. This construction generalizes naturally to any finite Coxeter group W. Furthermore, when W is a Weyl group, and hence can be affinized, there is an additional operator π_0 projecting along the highest root.

In [HT09] the first and last author constructed the *Hecke group algebra* $\mathcal{H}W$ by gluing together the 0-Hecke algebra and the group algebra of W along their right regular representation. Alternatively, $\mathcal{H}W$ can be constructed as the *biHecke algebra* of W, by gluing together the two realizations $H_0(W)$ and $\overline{H}_0(W)$ of the 0-Hecke algebra. $\mathcal{H}W$ admits a more conceptual description as the algebra of all operators on $\mathbb{K}.W$ preserving left antisymmetries; the representation theory of $\mathcal{H}W$ follows, governed by the combinatorics of descents. In [HST09], the authors further proved that, when W is a Weyl group, $\mathcal{H}W$ is a natural quotient of the affine Hecke algebra.

In this paper, following a suggestion of Alain Lascoux, we study the *biHecke monoid* $M(W)$, obtained by gluing together the two 0-Hecke *monoids*. This involves the combinatorics of the usual poset structures on W (left, right, left-right, Bruhat order), as well as a new one, the cutting poset, which in type A is related to blocks in permutation matrices. The guiding principle is the use of representation theory to derive a (so far elusive) summation formula for the size of this monoid, using that the simple and projective modules of M are indexed by the elements of W.

In type A, the tower of algebras $(\mathbb{K}[M(\mathfrak{S}_n)])_{n\in\mathbb{N}}$ possesses long sought-after properties. Indeed, it is well-known that several combinatorial Hopf algebras arise as Grothendieck rings of towers of algebras. The prototypical example is the tower of algebras of the symmetric groups which gives rise to the Hopf algebra Sym of symmetric functions, on the Schur basis. Another example, due to Krob and Thibon [KT97], is the tower of the 0-Hecke algebras of the symmetric groups which gives rise to the Hopf algebra QSym of quasi-symmetric functions of [Ges84], on the F_I basis. The product rule on the F_I's is naturally lifted through the descent map to a product on permutations, leading to the Hopf algebra FQSym of free quasi-symmetric functions. This calls for the existence of a tower of algebras $(A_n)_{n\in\mathbb{N}}$, such that each A_n contains $H_0(\mathfrak{S}_n)$ and has its simple modules indexed by the elements of $\mathfrak{S}_n$. The biHecke monoids $M(\mathfrak{S}_n)$ and their Borel submonoids $M_1(\mathfrak{S}_n)$ satisfy these properties, and are therefore expected to yield new representation theoretical interpretations of the bases of FQSym.

In the remainder of this introduction, we briefly review Coxeter groups and their 0-Hecke monoids, and introduce our main objects of study: the biHecke monoids.

1.1 Coxeter groups

Let (W, S) be a Coxeter group, that is, a group W with a presentation

$$W = \langle\, S \;\mid\; (ss')^{m(s,s')},\ \forall s, s' \in S \,\rangle\,, \tag{1}$$

with $m(s, s') \in \{1, 2, \ldots, \infty\}$ and $m(s, s) = 1$. The elements $s \in S$ are called *simple reflections*, and the relations can be rewritten as $s^2 = 1$, where 1 is the identity in W and:

$$\underbrace{ss'ss's\cdots}_{m(s,s')} = \underbrace{s'ss'ss'\cdots}_{m(s,s')} \qquad \text{for all } s, s' \in S\,. \tag{2}$$

Most of the time, we just write W for (W, S). In general, we follow the notation of [BB05], and we refer to this monograph and to [Hum90] for details on Coxeter groups and their Hecke algebras. Unless stated otherwise, we always assume that W is finite, and denote its generators by $S = (s_i)_{i\in I}$, where $I = \{1, 2, \ldots, n\}$ is the *index set* of W.

The prototypical example is the Coxeter group of type A_{n-1} which is the n-th symmetric group $(W, S) := (\mathfrak{S}_n, \{s_1, \dots, s_{n-1}\})$, where s_i denotes the *elementary transposition* which exchanges i and $i+1$. The relations are $s_i^2 = 1$ for $1 \le i \le n-1$ and the *braid relations*:

$$\begin{aligned} s_i s_j &= s_j s_i\,, && \text{for } |i-j| \ge 2\,,\\ s_i s_{i+1} s_i &= s_{i+1} s_i s_{i+1}\,, && \text{for } 1 \le i \le n-2\,. \end{aligned} \tag{3}$$

When writing a permutation $\mu \in \mathfrak{S}_n$ explicitly, we use the *one-line notation*, that is the sequence $\mu_1\mu_2 \cdots \mu_n$, where $\mu_i := \mu(i)$.

A *reduced word* $i_1 \dots i_k$ for an element $\mu \in W$ corresponds to a decomposition $\mu = s_{i_1} \cdots s_{i_k}$ of μ into a product of generators in S of minimal length $k = \ell(\mu)$. A *(right) descent* of μ is an element $i \in I$ such that $\ell(\mu s_i) < \ell(\mu)$. If μ is a permutation, this translates into $\mu_i > \mu_{i+1}$. *Left descents* are defined analogously. The sets of left and right descents of μ are denoted by $\mathrm{D}_L(\mu)$ and $\mathrm{D}_R(\mu)$, respectively.

A Coxeter group W comes equipped with four natural graded poset structures. Namely $\mu < \nu$ in *Bruhat order* (resp. *left (weak)*, *right (weak)*, *left-right (weak) order*) if some reduced word for μ is a subword (resp. right factor, left factor, factor) of some reduced word for ν. In type A, the left (resp. right) order give the usual left (resp. right) permutahedron.

For $J \subseteq I$, we denote by $W_J = \langle s_j \mid j \in J \rangle$ the subgroup of W generated by s_j with $j \in J$. Furthermore, the longest element in W_J (resp. W) is denoted by s_J (resp. w_0).

1.2 The 0*-Hecke monoid*

The 0-*Hecke monoid* $H_0(W) = \langle \pi_i \mid i \in I \rangle$ of a Coxeter group W is generated by the *simple projections* π_i with relations $\pi_i^2 = \pi_i$ for $i \in I$ and

$$\underbrace{\pi_i \pi_j \pi_i \pi_j \cdots}_{m(s_i, s_j)} = \underbrace{\pi_j \pi_i \pi_j \pi_i \cdots}_{m(s_i, s_j)} \qquad \text{for all } i, j \in I\,. \tag{4}$$

Thanks to these relations, the elements of $H_0(W)$ are canonically indexed by the elements of W by setting $\pi_w := \pi_{i_1} \cdots \pi_{i_k}$ for any reduced word $i_1 \dots i_k$ of w. We further denote by π_J the longest element of the *parabolic submonoid* $H_0(W_J) := \langle \pi_i \mid i \in J \rangle$.

The right regular representation of $H_0(W)$ induces a concrete realization of $H_0(W)$ as a monoid of operators acting on W, with generators $\pi_1, \dots, \pi_n$ defined by:

$$w.\pi_i := \begin{cases} w & \text{if } i \in \mathrm{D}_R(w),\\ ws_i & \text{otherwise.} \end{cases} \tag{5}$$

In type A, π_i sorts the letters at positions i and $i+1$ decreasingly, and for any permutation w, $w.\pi_{w_0} = n \cdots 21$. This justifies naming π_i an *elementary bubble antisorting operator*.

Another concrete realization of $H_0(W)$ can be obtained by considering instead the *elementary bubble sorting operators* $\overline{\pi}_1, \dots, \overline{\pi}_n$, whose action on W are defined by:

$$w.\overline{\pi}_i := \begin{cases} ws_i & \text{if } i \in \mathrm{D}_R(w),\\ w & \text{otherwise.} \end{cases} \tag{6}$$

In type A, and for any permutation w, one has $w.\overline{\pi}_{w_0} = 12 \cdots n$.

1.3 The biHecke monoid

Definition 1.1 *Let W be a finite Coxeter group. The* biHecke monoid *is the submonoid of functions from W to W generated simultaneously by the elementary bubble sorting and antisorting operators of* (5) *and* (6)*:*

$$M := M(W) := \langle \pi_1, \pi_2, \dots, \pi_n, \overline{\pi}_1, \overline{\pi}_2, \dots, \overline{\pi}_n \rangle .$$

1.4 Outline

The remainder of this paper is organized as follows. In Section 2, we generalize the notion of blocks of permutation matrices to any Coxeter group, and use it to define a new poset structure on W, which we call the *cutting poset*; we prove that it is (almost) a lattice, and derive that its Möbius function is essentially that of the hypercube.

In Section 3, we study the combinatorial properties of $M(W)$. In particular, we prove that it preserves left and Bruhat order, derive consequences on the fibers and image set of its elements, prove that it is aperiodic, and study its conjugacy classes of idempotents.

In Section 4, our strategy is to consider a "Borel" triangular submonoid of $M(W)$ whose representation theory is simpler, but with the same number of simple modules, in the hope to later induce back information about the representation theory of $M(W)$. Namely, we study the submonoid $M_1(W)$ of the elements fixing 1 in $M(W)$. This monoid not only preserves Bruhat order, but furthermore is contracting. It follows that it is $\mathcal{J}$-trivial which is the desired triangularity property. It is for example easily derived that $M_1(W)$ has $|W|$ simple modules, all of dimension 1. In fact most of our results about M_1 generalize to any $\mathcal{J}$-trivial monoid, which is the topic of a separate paper on the representation theory of $\mathcal{J}$-trivial monoids [DHST10].

In Section 5, we construct, for each $w \in W$, the *translation module* T_w by induction of the corresponding simple M_1-module. It is a quotient of the indecomposable projective module P_w of $M(W)$, and therefore admits the simple module S_w of $M(W)$ as top. It further admits a simple combinatorial model with the interval $[1, w]_R$ as support, and which passes down to S_w. We derive a formula for the dimension of S_w, using an inclusion-exclusion on the sizes of intervals in $(W, \leq_R)$, along the cutting poset. On the way, we study the algebra $\mathcal{H}W^{(w)}$ induced by the action of $M(W)$ on T_w. It turns out to be a natural w-analogue of the Hecke group algebra, acting not anymore on the full Coxeter group, but on the interval $[1, w]_R$ in right order. All the properties of the Hecke group algebra pass through this generalization, with the combinatorics of descents being replaced by that of blocks and of the cutting poset. In particular, $\mathcal{H}W^{(w)}$ is Morita equivalent to the incidence algebra of a lattice.

In Section 6, we derive (parts of) the representation theory of $M(W)$ from Sections 3, 4, and 5.

A long version of this paper with all proofs included will appear separately.

Acknowledgments

We would like to thank Jason Bandlow, Mahir Can, Brant Jones, Alain Lascoux, Jean-Christophe Novelli, Jean-Éric Pin, Vic Reiner, Franco Saliola, Benjamin Steinberg, and Jean-Yves Thibon for enlightening discussions. This research was driven by computer exploration, using the open-source mathematical software `Sage` [S$^+$09] and its algebraic combinatorics features developed by the `Sage-Combinat` community [SCc08]. AS and NT were in part supported by NSF grants DMS–0652641 and DMS–0652652. AS was in part supported by the Kastler foundation. FH was partly supported by ANR grant 06-BLAN-0380.

2 Blocks of Coxeter group elements and the cutting poset

In this section, we develop the combinatorics underlying the representation theory of the translation modules studied in Section 5. The key question is: for which subsets $J \subseteq I$ does the canonical bijection between a Coxeter group W and the Cartesian product $W_J \times {}^J W$ of a parabolic subgroup W_J by its set of coset representatives ${}^J W$ in W restrict properly to an interval $[1, w]_R$ in right order (see Figure 1)? In type A, the answer is given by the so-called *blocks* in the permutation matrix of w, and we generalize this notion to any Coxeter group. After reviewing parabolic subgroups and cosets representatives in Section 2.1, we define blocks of Coxeter group elements in Section 2.2, and show in Section 2.3 how this notion specializes to type A. Finally, in Section 2.4, we introduce and study the *cutting poset*.

2.1 Parabolic subgroups and cosets representatives

For a subset $J \subseteq I$, the *parabolic subgroup* W_J of W is the Coxeter subgroup of W generated by $s_j, j \in J$. A complete system of minimal length representatives of the right cosets $W_J w$ (resp. of the left cosets wW_J) are given respectively by:

$$ {}^J W := \{x \in W \mid \mathrm{D}_L(x) \cap J = \emptyset\} \qquad \text{and} \qquad W^J := \{x \in W \mid \mathrm{D}_R(x) \cap J = \emptyset\} \, . $$

Every $w \in W$ has a unique decomposition $w = w_J \, {}^J w$ with $w_J \in W_J$ and ${}^J w \in {}^J W$. Similarly, there is a unique decomposition $w = w^K \, {}_K w$ with ${}_K w \in {}_K W = W_K$ and $w^K \in W^K$. A subset J is *left reduced* w.r.t. w if $J' \subset J$ implies ${}^{J'} w >_L {}^J w$. *Right reduced* K's are defined analogously.

2.2 Blocks of Coxeter group elements

We now come to the definition of blocks of Coxeter group elements, and associated cutting points.

Definition 2.1 (Blocks and cutting points) *Let $w \in W$. We say $K \subseteq I$ is a* right block *(resp. $J \subseteq I$ is a* left block*) of w, if there exists $J \subseteq I$ (resp. $K \subseteq I$) such that $wW_K = W_J w$.*

In that case, $v := w^K$ is called a cutting point *of w, which we denote by $v \sqsubseteq w$. Furthermore, K is* proper *if $K \neq \emptyset$ and $K \neq I$; it is* nontrivial *if $w^K \neq w$ (or equivalently ${}_K w \neq 1$); analogous definitions are made for left blocks.*

We denote by $\mathcal{B}_{\mathcal{R}}(w)$ (resp. $\mathcal{B}_{\mathcal{L}}(w)$) the set of all right (resp. left) blocks for w, and by $\mathcal{RB}_{\mathcal{R}}(w)$ (resp. $\mathcal{RB}_{\mathcal{L}}(w)$) the set of all reduced right (resp. left) blocks for w.

Proposition 2.2 *Assuming that J is reduced, J is a left block of w if and only if the bijection*

$$ \begin{cases} W_J \times {}^J W & \to W \\ (u, v) & \mapsto uv \end{cases} $$

restricts to a bijection $[1, w_J]_R \times [1, {}^J w]_R \to [1, w]_R$ (see Figure 1).

Due to Proposition 2.2, we also say that $[1, v]_R$ *tiles* $[1, w]_R$ if $v = {}^J w$ for some left block J.

Example 2.3 *For $w = w_0$, any $K \subseteq I$ is a reduced right block; of course, $w_0^K \leq_L w_0$ and ${}_K w_0$ is the maximal element of the parabolic subgroup $W_K = {}_K W$. The cutting point $w^K \sqsubseteq w$ is the maximal element of the right descent class for the complement of K.*

Proposition 2.4 *The set $\mathcal{B}_{\mathcal{L}}(w)$ (resp. $\mathcal{B}_{\mathcal{R}}(w)$) of left (resp. right) blocks is stable under union and intersection. So, they form a sublattice of the Boolean lattice.*

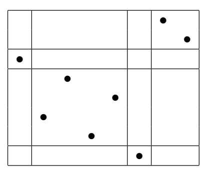
The sets $\mathcal{RB}_{\mathcal{L}}(w)$ and $\mathcal{RB}_{\mathcal{R}}(w)$ are (dual) Moore families and therefore lattices.

Definition 2.5 (**w-codescent sets**) *For $u \in [1, w]_R$ define $K^{(w)}(u)$ to be the maximal reduced right block K of w such that u is below the corresponding cutting point, that is $u \leq_R w^K$. Let $J^{(w)}(u)$ be the left block corresponding to $K^{(w)}(u)$.*

Example 2.6 *When $w = w_0$, any $J \subseteq I$ is a reduced left block. Furthermore, for $u \in W$, $J^{(w_0)}(u)$ is the complement of its left-descent set: $J^{(w_0)}(u) = I \setminus \mathrm{D}_L(u)$. Idem on the right.*

2.3 Blocks of permutations

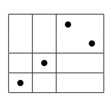
We now specialize to type A_{n-1}. For a permutation $w \in \mathfrak{S}_n$, the blocks of Definition 2.1 correspond to the usual notion of blocks of the matrix of w (or unions thereof), and the cutting points w^K for right blocks K correspond to putting the identity in the matrix-blocks of w. Namely, a *matrix-block* of a permutation w is an interval $[k', k'+1, \dots, k]$ which is mapped to another interval. Pictorially, this corresponds to a square submatrix of the matrix of w which is again a permutation matrix (that of the *associated permutation*). For example, the interval $[2, 3, 4, 5]$ is mapped to the interval $[4, 5, 6, 7]$ by the permutation $w = 36475812 \in \mathfrak{S}_8$, and is therefore a matrix-block of w with associated permutation 3142. Similarly, $[7, 8]$ is a matrix-block with associated permutation 12:



The singletons $[i]$ and the full set $[1, 2, \dots, n]$ are always matrix-blocks; the other matrix-blocks are called *proper*. A permutation with no proper matrix-block, such as 58317462, is called *simple*. See [AA05] for a review of simple permutations or, equivalently, dimension 2 posets.

Proposition 2.7 *Let $w \in \mathfrak{S}_n$. The right blocks of w are in bijection with disjoint unions of (non singleton) matrix-blocks for w; each matrix-block with column set $[i, i+1, \dots, k]$ contributes $\{i, i+1, \dots, k-1\}$ to the right block; each matrix-block with row set $[i, i+1, \dots, k]$ contributes $\{i, i+1, \dots, k-1\}$ to the left block. In addition, trivial right blocks correspond to unions of identity matrix-blocks. Also, reduced right blocks correspond to unions of connected matrix-blocks.*

Example 2.8 *As in Figure 1, consider the permutation* 4312*, whose permutation matrix is:*



The reduced (right)-blocks are $K = \{\}$, $\{1\}$, $\{2, 3\}$, and $\{1, 2, 3\}$. The cutting points are 4312*,* 3412*,* 4123*, and* 1234*, respectively. The corresponding left blocks are $J = \{\}$, $\{3\}$, $\{1, 2\}$ and $\{1, 2, 3\}$, respectively. The non-reduced blocks are $\{3\}$ and $\{1, 3\}$, as they are respectively equivalent to the blocks $\{\}$ and $\{1\}$. Finally, the trivial blocks are $\{\}$ and $\{3\}$.*

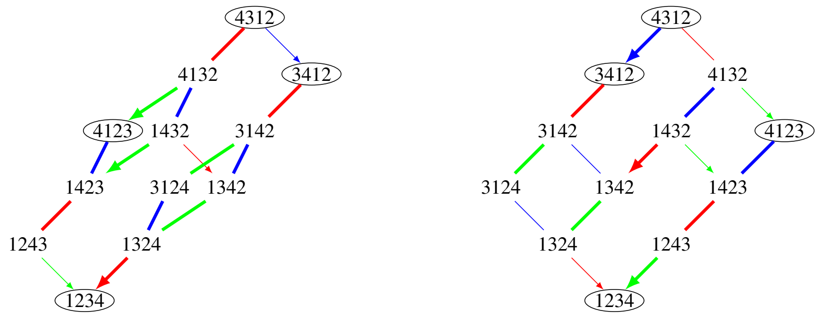


Fig. 1: Two pictures of the interval $[1234, 4312]_R$ in right order illustrating its proper tilings, for $J := \{3\}$ and $J := \{1, 2\}$, respectively. The thick edges highlight the tiling. The circled permutations are the cutting points, which are at the top of the tiling intervals. Blue, red, green lines correspond to s_1, s_2, s_3, respectively. See Section 5.4 for the definition of the orientation of the edges (this is $G^{(4312)}$); edges with no arrow tips point in both directions.

2.4 The cutting poset

Theorem 2.9 *$(W, \sqsubseteq)$ is a poset with 1 as minimal element; it is further a subposet of both left and right order. Every interval of $(W, \sqsubseteq)$ is a sublattice of both left and right order.*

The $\sqsubseteq$-lower covers of an element w correspond to the nontrivial blocks of w which are minimal for inclusion. The meet-semilattice L_w they generate is free, and is in correspondence with the lattice of unions of these minimal nontrivial blocks, or alternatively of the intersections of the intervals $[1, u]_R$ for u $\sqsubseteq$-lower covers of w.

The Möbius function is given by $\mu(u, w) = \pm 1$ if u is in L_w (with alternating sign according to the usual rule for the Boolean lattice), and 0 otherwise.

This Möbius function is used in Section 5.4 to compute the size of the simple modules of M.

Conjecture 2.10 *$(W, \sqsubseteq)$ is a meet-semilattice whose intervals are all distributive lattices.*

3 The combinatorics of $M(W)$

In this section we study the combinatorics of the biHecke monoid $M(W)$ of a finite Coxeter group W. In particular, we prove in Section 3.1 that its elements preserve left order and Bruhat order, and derive in Section 3.2 properties of their image sets and fibers. Finally, in Section 3.3, we prove the key combinatorial ingredients for the enumeration of the simple modules of $M(W)$ in Section 6: $M(W)$ is aperiodic and admits $|W|$ conjugacy classes of idempotents.

3.1 Preservation of left and Bruhat order

Lemma 3.1 *Take $f \in M(W)$, $w \in W$, and $j \in I$. Then, $(s_j w).f$ is either $w.f$ or $s_j(w.f)$.*

Proposition 3.2 *The elements f of M preserve left order: $u \leq_L v \Rightarrow u.f \leq_L v.f$.*

Proposition 3.3 *The elements f of M preserve Bruhat order: $u \leq_B v \Rightarrow u.f \leq_B v.f$.*

Proposition 3.4 *Any $f \in M$ such that $1.f = 1$ is contracting for Bruhat order: $w.f \leq_B w$.*

3.2 Fibers and image sets

Proposition 3.5 *The image set of an idempotent in $M(W)$ is an interval in left order.*

Proposition 3.6 *Take $f \in M(W)$, and consider the Hasse diagram of left order contracted with respect to the fibers of f. Then, this graph is isomorphic to left order restricted on the image set.*

Proposition 3.7 *Any element $f \in M(W)$ is characterized by its set of fibers and its image set.*

A monoid M is called *aperiodic* if for any $f \in M$, there exists $k > 0$ such that $f^{k+1} = f^k$. Note that, in that case, $f^\infty := f^k = f^{k+1} = \dots$ is an idempotent.

Proposition 3.8 *The biHecke monoid $M(W)$ is* aperiodic.

3.3 Conjugacy classes of idempotents

Proposition 3.9 *For $w \in W$, $e_w := \pi_{w^{-1}w_0}\overline{\pi}_{w_0 w}$ is the unique idempotent with image set $[1, w]_L$. For $u \in W$, it satisfies $e_w(u) = \max_{\leq_B}\big([1, u]_B \cap [1, w]_L\big)$.*

Corollary 3.10 *For $u, w \in W$, the intersection $[1, u]_B \cap [1, w]_L$ is a lower $\leq_L$ ideal with a unique maximal element v in Bruhat order. The maximum is given by $v = e_w(u)$.*

We are now in the position to describe the conjugacy relations between the idempotents of M.

Lemma 3.11 *Let e and f be idempotents with respective image sets $[a, b]_L$ and $[c, d]_L$. Then, $f \in MeM$ if and only if $dc^{-1} \leq_{LR} ba^{-1}$. In particular, e and f are conjugates if and only if the intervals $[a, b]_L$ and $[c, d]_L$ are of the same type: $dc^{-1} = ba^{-1}$.*

Corollary 3.12 *The idempotents $(e_w)_{w \in W}$ form a complete set of representatives of the conjugacy classes of idempotents in M.*

4 The Borel submonoid $M_1(W)$ and its representation theory

In the previous section, we outlined the importance of the idempotents $(e_w)_{w \in W}$. A crucial feature is that they live in a "Borel" submonoid $M_1 := \{f \in M \mid 1.f = 1\}$. In fact:

Theorem 4.1 *M_1 has a unique minimal generating set which consists of the ($2^n - n$ in type A) idempotents e_w where $w_0 w^{-1}$ is Grassmanian.*

Furthermore, the elements of M_1 are both order-preserving and contracting for Bruhat order.

Corollary 4.2 *For $f, g \in M_1$, define the relation $f \leq g$ if $w.f \leq w.g$ for all $w \in W$. Then, $\leq$ defines a partial order on M_1 such that $fg \leq f$ and $fg \leq g$ for any $f, g \in M_1$.*

M_1 is therefore *$\mathcal{J}$-trivial* (see e.g. [Pin09]). The generalization of most of the representation theoretical results summarized below to any *$\mathcal{J}$-trivial* monoid is the topic of [DHST10].

For each $w \in W$ define S_w to be the one-dimensional vector space with basis $\{\epsilon_w\}$ together with the right operation of any $f \in M_1$ given by $\epsilon_w.f := \epsilon_w$ if $w.f = w$ and $\epsilon_w.f := 0$ otherwise. The basic features of the representation theory of M_1 can be stated as follows:

Theorem 4.3 *The radical of $\mathbb{K}[M_1]$ is the ideal with basis $(f^\infty - f)_f$, for f non-idempotent. The quotient of $\mathbb{K}[M_1]$ by its radical is commutative. Therefore, all simple M_1-module are one dimensional. In fact, the family $\{S_w\}_{w \in W}$ forms a complete system of representatives of the simple M_1-modules.*

To describe the indecomposable projective modules, we note that the restriction of the conjugacy relation ($\mathcal{J}$-order) to idempotents has a very simple description:

Proposition 4.4 *For $u, v \in W$, the following are equivalent:*

- $e_u e_v = e_u$;
- $e_v e_u = e_u$;
- *$v \leq_L u$ for left order;*
- *there exists $x, y \in M_1$ such that $e_u = x e_v y$.*

Moreover $(e_u e_v)^\infty = e_{u \vee_L v}$, where $u \vee_L v$ is the join of u and v in left order.

Definition 4.5 *For any element $x \in M$, define*

$$\operatorname{lfix}(x) := \min_{\leq_L}\{u \in W \mid e_u x = x\} \qquad \textit{and} \qquad \operatorname{rfix}(x) := \min_{\leq_L}\{u \in W \mid x e_u = x\}\,. \tag{7}$$

Then, the projective modules and Cartan invariants can be described as follows:

Theorem 4.6 *There is an explicit basis $(b_f)_{f \in M_1}$ of $\mathbb{K}[M_1]$ such that, for all $w \in W$,*

- *the family $\{b_x \mid \operatorname{lfix}(x) = w\}$ is a basis for the right projective module associated to S_w;*
- *the family $\{b_x \mid \operatorname{rfix}(x) = w\}$ is a basis for the left projective module associated to S_w.*

Moreover, the Cartan invariant of $\mathbb{K}[M_1]$ defined by $c_{u,v} := \dim(e_u \mathbb{K}[M_1] e_v)$ for $u, v \in W$ is given by $c_{u,v} = |C_{u,v}|$, where $C_{u,v} := \{f \in M_1 \mid u = \operatorname{lfix}(f)$ and $v = \operatorname{rfix}(f)\}$.

5 Translation modules and w-biHecke algebras

The main purpose of this section is to pave the ground for the construction of the simple modules S_w of the biHecke monoid $M = M(W)$ in Section 6. To this end, in Section 5.1, we endow the interval $[1, w]_R$ with a natural structure of a combinatorial M-module T_w, called *translation module*. This module is closely related to the projective module P_w of M (Corollary 6.2), which explains its important role. By taking the quotient of $\mathbb{K}[M]$ through its representation on T_w, we obtain a w-analogue $\mathcal{H}W^{(w)}$ of the biHecke algebra $\mathcal{H}W$. This algebra turns out to be interesting in its own right, and we proceed by generalizing most of the results of [HT09] on the representation theory of $\mathcal{H}W$.

As a first step, we introduce in Section 5.2 a collection of submodules $P_J^{(w)}$ of T_w, which are analogues of the projective modules of $\mathcal{H}W$. Unlike for $\mathcal{H}W$, not any subset J of I yields such a submodule, and this is where the combinatorics of the blocks of w as introduced in Section 2 enters the game. In a second step, we derive in Section 5.3 a lower bound on the dimension of $\mathcal{H}W^{(w)}$; this requires a (fairly involved) combinatorial construction of a family of functions on $[1, w]_R$ which is triangular with respect to Bruhat order. In Section 5.4 we combine these results to derive the dimension and representation theory of $\mathcal{H}W^{(w)}$: projective and simple modules, Cartan matrix, quiver, etc.

5.1 Translation modules and w-biHecke algebras

For $f \in M$, define the *type* of f by $\operatorname{type}(f) := (w_0.f)(1.f)^{-1}$. By Proposition 3.2, we know that for $f, g \in M$ either $\operatorname{type}(fg) = \operatorname{type}(f)$, or $\ell(w_0.(fg)) - \ell(1.(fg)) < \ell(w_0.f) - \ell(1.f)$ and hence $\operatorname{type}(fg) \neq \operatorname{type}(f)$. The second case occurs precisely when $\operatorname{fiber}(f)$ is strictly finer than $\operatorname{fiber}(fg)$ or equivalently $\operatorname{rank}(fg) < \operatorname{rank}(f)$, where the *rank* is the cardinality of the image set.

Definition 5.1 *Fix $f \in M$. The right M-module*

$$\mathrm{trans}(f) := \mathbb{K}.fM/\mathbb{K}.\{h \in fM \mid \mathrm{rank}(h) < \mathrm{rank}(f)\}$$

is called the translation module *associated with f.*

Proposition 5.2 *Fix $f \in M$. Then:*

$$\{h \in fM \mid \mathrm{rank}(h) = \mathrm{rank}(f)\} = \{f_u \mid u \in [1, \mathrm{type}(f)^{-1}w_0]_R\}\,, \tag{8}$$

where f_u is the unique element of M such that $\mathrm{fiber}(f_u) = \mathrm{fiber}(f)$ *and* $1.f_u = u$.

Proposition 5.3 *Set $w = \mathrm{type}(f)^{-1}w_0$. Then, $(f_u)_{u\in[1,w]_R}$ forms a basis of* $\mathrm{trans}(f)$ *such that:*

$$f_u.\pi_i = \begin{cases} f_u & \text{if } i \in \mathrm{D}_R(u) \\ f_{us_i} & \text{if } i \notin \mathrm{D}_R(u) \text{ and } us_i \in [1,w]_R \\ 0 & \text{otherwise;} \end{cases} \quad f_u.\overline{\pi}_i = \begin{cases} f_{us_i} & \text{if } i \in \mathrm{D}_R(u) \text{ and } us_i \in [1,w]_R \\ f_u & \text{if } i \notin \mathrm{D}_R(u) \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

This proposition gives a combinatorial model for translation modules. It is clear that two functions with the same type yield isomorphic translation modules. The converse also holds:

Proposition 5.4 *For any $f, f' \in M$, the translation modules* $\mathrm{trans}(f)$ *and* $\mathrm{trans}(f')$ *are isomorphic if and only if* $\mathrm{type}(f) = \mathrm{type}(f')$.

By the previous proposition, we may choose a canonical representative for translation modules. We choose $T_w := \mathrm{trans}(e_{w,w_0})$, and identify its basis with $[1,w]_R$ via $u \mapsto f_u$.

Definition 5.5 *The w-biHecke algebra $\mathcal{H}W^{(w)}$ is the natural quotient of $\mathbb{K}[M(W)]$ through its representation on T_w. In other words, it is the subalgebra of $\mathrm{End}(T_w)$ generated by the operators π_i and $\overline{\pi}_i$ of Proposition 5.3.*

5.2 Left antisymmetric submodules

By analogy with the simple reflections in the Hecke group algebra, we define for each $i \in I$ the operator $s_i := \pi_i + \overline{\pi}_i - 1$. For $u \in [1,w]_R$, it satisfies $u.s_i = us_i$ if $us_i \in [1,w]_R$ and $u.s_i = -u$ otherwise. These operators are still involutions, but do not quite satisfy the braid relations. One can further define operators $\overleftarrow{s}_i$ acting similarly on the left.

Definition 5.6 *For $J \subseteq I$, set $P_J^{(w)} := \{v \in T_w \mid \overleftarrow{s}_i.v = -v, \quad \forall i \in J\}$.*

For $w = w_0$, these are the projective modules P_J of the biHecke algebra.

Proposition 5.7 *Take $w \in W$ and $J \subseteq I$ left reduced. Then, J is a reduced left block of w if and only if $P_J^{(w)}$ is a submodule of T_w.*

It is clear from the definition that for $J_1, J_2 \subseteq I$, $P_{J_1\cup J_2}^{(w)} = P_{J_1}^{(w)} \cap P_{J_2}^{(w)}$. Since the set $\mathcal{RB_L}(w)$ of reduced left blocks of w is stable under union, the set of modules $(P_J^{(w)})_{J\in\mathcal{RB_L}(w)}$ is stable under intersection. On the other hand, unless J_1 and J_2 are comparable, $P_{J_1\cup J_2}^{(w)}$ is a strict subspace of $P_{J_1}^{(w)} + P_{J_2}^{(w)}$. Hence, for $J \in \mathcal{RB_L}(w)$, we set $S_J^{(w)} := P_J^{(w)} / \sum_{J' \supsetneq J, J' \in \mathcal{RB_L}(w)} P_{J'}^{(w)}$.

5.3 A (maximal) Bruhat-triangular family of $\mathcal{H}W^{(w)}$

Consider the submonoid F in $\mathcal{H}W^{(w)}$ generated by the operators π_i, $\overline{\pi}_i$, and s_i, for $i \in I$. For $f \in F$ and $u \in [1, w]_R$, we have $u.f = \pm v$ for some $v \in [1, w]_R$. For our purposes, the signs can be ignored and f be considered as a function from $[1, w]_R$ to $[1, w]_R$.

Definition 5.8 *For $u, v \in [1, w]_R$, a function $f \in F$ is called* (u, v)-triangular *(for Bruhat order) if v is the unique minimal element of* $\mathrm{im}(f)$ *and u is the unique maximal element of $f^{-1}(v)$ (all minimal and maximal elements in this context are with respect to Bruhat order).*

Proposition 5.9 *Take $u, v \in [1, w]_R$ such $K^{(w)}(u) \subseteq K^{(w)}(v)$. Then, there exists a (u, v)-triangular function $f_{u,v}$ in F.*

For example, for $w = 4312$ in $\mathfrak{S}_4$, the condition on u and v is equivalent to the existence of a path from u to v in the digraph $G^{(4312)}$ (see Figure 1 and Section 5.4).

The construction of $f_{u,v}$ is explicit, and the triangularity derives from $f_{u,v}$ being either in M, or close enough to be bounded below by an element of M. It follows from the upcoming Theorem 5.10 that the condition on u and v is not only sufficient but also necessary.

5.4 Representation theory of w-biHecke algebras

Consider the digraph $G^{(w)}$ on $[1, w]_R$ with an edge $u \mapsto v$ if $u = vs_i$ for some i, and $J^{(w)}(u) \subseteq J^{(w)}(v)$. Up to orientation, this is the Hasse diagram of right order (see for example Figure 1). The following theorem is a generalization of [HT09, Section 3.3].

Theorem 5.10 *$\mathcal{H}W^{(w)}$ is the maximal algebra stabilizing all the modules $P_J^{(w)}$, for $J \in \mathcal{RB}_{\mathcal{L}}(w)$.*

The elements $f_{u,v}$ of Proposition 5.9 form a basis $\mathcal{H}W^{(w)}$; in particular,

$$\dim \mathcal{H}W^{(w)} = |\{(u, v) \mid J^{(w)}(u) \subseteq J^{(w)}(v)\}| \,. \tag{10}$$

$\mathcal{H}W^{(w)}$ is the digraph algebra of the graph $G^{(w)}$.

The family $(P_J)_{J\in\mathcal{RB}_{\mathcal{L}}(w)}$ forms a complete system of representatives of the indecomposable projective modules of $\mathcal{H}W^{(w)}$.

The family $(S_J)_{J\in\mathcal{RB}_{\mathcal{L}}(w)}$ forms a complete system of representatives of the simple modules of $\mathcal{H}W^{(w)}$. The dimension of S_J is the size of the corresponding w-descent class.

$\mathcal{H}W^{(w)}$ is Morita equivalent to the poset algebra of the lattice $[1, w]_{\sqsubseteq}$.

6 The representation theory of $M(W)$

Theorem 6.1 *The monoid $M = M(W)$ admits $|W|$ non-isomorphic simple modules $(S_w)_{w\in W}$ (resp. projective indecomposable modules $(P_w)_{w\in W}$).*

The simple module S_w is isomorphic to the top simple module $S_{\{\}}^{(w)}$ of the translation module T_w. In general, the simple quotient module $S_J^{(w)}$ of T_w is isomorphic to S_{J_w} of M.

For example, the simple module S_{4312} is of dimension 3, with basis $\{4312, 4132, 1432\}$ (see Figure 1). The other simple modules S_{3412}, S_{4123}, and S_{1234} are of dimension 5, 3, and 1.

Corollary 6.2 *The translation module T_w is an indecomposable M-module, quotient of the projective module P_w of M.*

M_1 is a submonoid of M. Therefore any M-module X is a M_1-module, and its M_1-character $[X]_{M_1}$ depends only on its M-character $[X]_M$. This defines a $\mathbb{Z}$-linear map $[X]_M \mapsto [X]_{M_1}$. Let $(S_w)_{w\in W}$ and $(S^1_w)_{w\in W}$ be complete families of simple modules representatives for M and M_1, respectively. The matrix of $[X]_M \mapsto [X]_{M_1}$ is called the *decomposition matrix* of M over M_1; its coefficient (u, v) is the multiplicity of S^1_u as a composition factor of S_v viewed as an M_1-module.

Theorem 6.3 *The decomposition matrix of M over M_1 is upper uni-triangular for right order, with* $0, 1$ *entries. Furthermore, T_w is isomorphic to the induction to M of the simple module S^1_w.*

References

[AA05] M. H. Albert and M. D. Atkinson. Simple permutations and pattern restricted permutations. *Discrete Math.*, 300(1-3):1–15, 2005.

[BB05] A. Björner and F. Brenti. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York, 2005.

[DHST10] T. Denton, F. Hivert, A. Schilling, and N. M. Thiéry. The representation theory of $\mathcal{J}$-trivial monoids. In preparation, 2010.

[Ges84] I. M. Gessel. Multipartite P-partitions and inner products of skew Schur functions. In *Combinatorics and algebra (Boulder, Colo., 1983)*, volume 34 of *Contemp. Math.*, pages 289–317. Amer. Math. Soc., Providence, RI, 1984.

[HST09] F. Hivert, A. Schilling, and N. M. Thiéry. Hecke group algebras as quotients of affine Hecke algebras at level 0. *J. Combin. Theory Ser. A*, 116(4):844–863, 2009.

[HT09] F. Hivert and N. M. Thiéry. The Hecke group algebra of a Coxeter group and its representation theory. *J. Algebra*, 321(8):2230–2258, 2009.

[Hum90] J. E. Humphreys. *Reflection groups and Coxeter groups*, volume 29 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[KT97] D. Krob and J.-Y. Thibon. Noncommutative symmetric functions. IV. Quantum linear groups and Hecke algebras at $q = 0$. *J. Algebraic Combin.*, 6(4):339–376, 1997.

[Pin09] J.-É. Pin. *Mathematical grounds of automata theory*. 2009.

[S$^+$09] W. A. Stein et al. *Sage Mathematics Software (Version 3.3)*. The Sage Development Team, 2009. `http://www.sagemath.org`.

[SCc08] The Sage-Combinat community. Sage-Combinat: enhancing sage as a toolbox for computer exploration in algebraic combinatorics, 2008. `http://combinat.sagemath.org`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 198–209

Criteria for rational smoothness of some symmetric orbit closures

Axel Hultman

Department of Mathematics, KTH-Royal Institute of Technology, SE-100 44, Stockholm, Sweden.

Abstract. Let G be a connected reductive linear algebraic group over $\mathbb{C}$ with an involution θ. Denote by K the subgroup of fixed points. In certain cases, the K-orbits in the flag variety G/B are indexed by the twisted identities $\iota(\theta) = \{\theta(w^{-1})w \mid w \in W\}$ in the Weyl group W. Under this assumption, we establish a criterion for rational smoothness of orbit closures which generalises classical results of Carrell and Peterson for Schubert varieties. That is, whether an orbit closure is rationally smooth at a given point can be determined by examining the degrees in a "Bruhat graph" whose vertices form a subset of $\iota(\theta)$. Moreover, an orbit closure is rationally smooth everywhere if and only if its corresponding interval in the Bruhat order on $\iota(\theta)$ is rank symmetric.

In the special case $K = \mathrm{Sp}_{2n}(\mathbb{C})$, $G = \mathrm{SL}_{2n}(\mathbb{C})$, we strengthen our criterion by showing that only the degree of a single vertex, the "bottom one", needs to be examined. This generalises a result of Deodhar for type A Schubert varieties.

Résumé. Soit G un groupe algébrique connexe réductif sur $\mathbb{C}$, équipé d'une involution θ. Soit K le sous–groupe de ses points fixes. Dans certains cas, les orbites des points de la variété de drapeaux G/B sous l'action de K sont indexées par les identités tordues, $\iota(\theta) = \{\theta(w^{-1})w \mid w \in W\}$, du groupe de Weyl W. Sous cette hypothèse, on établit un critère pour la lissité rationnelle des adhérences des orbites, qui généralise des résultats classiques de Carrell et Peterson pour les variétés de Schubert. Plus précisément, on peut déterminer si l'adhérence d'une orbite est rationnellement lisse en examinant les degrés dans un "graphe de Bruhat" dont les sommets forment un sous–ensemble de $\iota(\theta)$. En outre, l'adhérence d'une orbite est partout rationnellement lisse si et seulement si l'intervalle correspondant dans l'ordre de Bruhat de $\iota(\theta)$ est symétrique respectivement au rang.

Dans le cas particulier $K = \mathrm{Sp}_{2n}(\mathbb{C})$, $G = \mathrm{SL}_{2n}(\mathbb{C})$, nous améliorons notre critère en montrant qu'il suffit d'examiner le degré d'un seul sommet, celui "du bas". Ceci généralise un résultat de Deodhar pour les variétés de Schubert de type A.

Keywords: Rational smoothness, symmetric orbit, Bruhat graph

1 Introduction

Let G be a connected reductive complex linear algebraic group equipped with an automorphism θ of order 2. There is a θ-stable Borel subgroup B which contains a θ-stable maximal torus T [Ste68, §7] with normaliser N. Let $K = G^\theta$ be the fixed point subgroup. We may always assume θ to be the complexification of the Cartan involution of some real form $G_{\mathbb{R}}$ of G.

The flag variety $X = G/B$ decomposes into finitely many orbits under the action of the symmetric subgroup K by left translations. A natural "Bruhat-like" partial order on the set of orbits $K\backslash X$ is defined by inclusion of their closures. Let V denote this poset. Richardson and Springer [RS90, RS94] defined a poset map $\varphi : V \to \mathrm{Br}(W)$, where $\mathrm{Br}(W)$ is the Bruhat order on the Weyl group $W = N/T$. The image of φ is contained in the set of *twisted involutions* $\mathcal{I}(\theta) = \{w \in W \mid \theta(w) = w^{-1}\}$. In general, φ is neither injective nor surjective. For certain choices of G and θ, however, φ produces a poset isomorphism $V \cong \mathrm{Br}(\iota(\theta))$, where $\iota(\theta) = \{\theta(w^{-1})w \mid w \in W\} \subseteq \mathcal{I}(\theta)$ is the set of *twisted identities* and $\mathrm{Br}(\cdot)$ denotes induced subposet of $\mathrm{Br}(W)$. In Section 3, we shall make explicit under what circumstances this fairly restrictive assumption holds. Now suppose that φ is such an isomorphism and let $\overline{\mathcal{O}_w}$, $w \in \iota(\theta)$, denote the closure of the orbit $\mathcal{O}_w = \varphi^{-1}(w)$. In this article we express the rationally singular locus of $\overline{\mathcal{O}_w}$ in terms of the combinatorics of $\iota(\theta)$.

With each $w \in \iota(\theta)$, we associate a *Bruhat graph* $\mathrm{BG}(w)$ with vertex set $I_w = \{u \in \iota(\theta) \mid u \leq w\}$. Our first main result, Theorem 5.7, states that $\overline{\mathcal{O}_w}$ is rationally smooth at $\mathcal{O}_u$ if and only if v is contained in $\rho(w)$ edges for all $u \leq v \leq w$, where $\rho(w)$ is the rank of w in $\mathrm{Br}(\iota(\theta))$. In particular, $\overline{\mathcal{O}_w}$ is rationally smooth if and only if $\mathrm{BG}(w)$ is $\rho(w)$-regular. This latter statement also turns out to be equivalent to the principal order ideal $\mathrm{Br}(I_w)$ being rank-symmetric; see Theorem 5.8 below.

The assertions just stated generalise celebrated criteria due to Carrell and Peterson [Car94] for rational smoothness of Schubert varieties. We recover their results in the special case where $G = G' \times G'$ and $\theta(x, y) = (y, x)$.

The main brushstrokes of our proofs are completely similar to those of Carrell and Peterson. Below the surface, however, their results rely on delicate connections between Kazhdan-Lusztig polynomials and the combinatorics of (ordinary) Bruhat graphs. Our chief contribution is to extend these properties to a more general setting. Very roughly, here is what we do:

First, properties of $\iota(\theta)$ are established that combined with results of Brion [Bri99] imply a bound on the degrees in $\mathrm{BG}(w)$ that generalises "Deodhar's inequality" for degrees in ordinary Bruhat graphs of Weyl groups.

Second, an explicit procedure, in terms of combinatorial properties of $\iota(\theta)$, for computing the "R-polynomials" of [LV83, Vog83] is extracted from the correspondence $V \leftrightarrow \iota(\theta)$. Using this procedure we establish several properties of these polynomials (and therefore of Kazhdan-Lusztig-Vogan polynomials) and relate them to degrees in the graphs $\mathrm{BG}(w)$. This generalises well known properties of ordinary Kazhdan-Lusztig polynomials and R-polynomials and how they are related to ordinary Bruhat graphs.

The most prominent example where our results say something which is not contained in [Car94] is $G = \mathrm{SL}_{2n}(\mathbb{C})$, $K = \mathrm{Sp}_{2n}(\mathbb{C})$. For this setting, we prove the stronger statement (Corollary 6.5) that the degree of the bottom vertex alone suffices to decide rational smoothness. That is, $\overline{\mathcal{O}_w}$ is rationally smooth at $\mathcal{O}_u$ if and only if the degree of u in $\mathrm{BG}(w)$ is $\rho(w)$. This is analogous to a corresponding result for type A Schubert varieties which is due to Deodhar [Deo85]. Again, that result is contained in ours as a special case.

Remark 1.1 After a preliminary version of [Hul09] was circulated, McGovern [McG09] has applied our results in order to deduce a criterion for (rational) smoothness in the case $G = \mathrm{SL}_{2n}(\mathbb{C})$, $K = \mathrm{Sp}_{2n}(\mathbb{C})$ in terms of pattern avoidance among fixed point free involutions. Moreover, he proved that in this case the rationally singular loci in fact coincide with the singular loci.

In Section 3, we make precise the assumptions on θ for which our results are valid. Thereafter, the Bruhat graphs $\mathrm{BG}(w)$ are introduced in Section 4. Our Carrell-Peterson type criteria for rational smoothness are deduced in Section 5. Finally, in Section 6, we prove that the bottom vertex alone suffices to decide rational smoothness when $G = \mathrm{SL}_{2n}(\mathbb{C})$, $K = \mathrm{Sp}_{2n}(\mathbb{C})$.

Details left out in the present extended abstract can be found in [Hul09].

2 Kazhdan-Lusztig-Vogan polynomials

In the present paper, the principal method for detecting rational singularities of symmetric orbit closures is via Kazhdan-Lusztig-Vogan polynomials. Here, we briefly review some of their properties and establish notation. For more information we refer the reader to [LV83] or [Vog83]. Our terminology chiefly follows the latter reference.

Let $\mathcal{D}$ denote the set of pairs $(\mathcal{O}, \gamma)$, where $\mathcal{O} \in K\backslash X$ and γ is a K-equivariant local system on $\mathcal{O}$. The choice of γ is equivalent to the choice of a character of the component group of the stabiliser K_x of a point $x \in \mathcal{O}$. In particular, γ is unique if K_x is connected. Since $\mathcal{O}$ is determined by γ, we may abuse notation and write γ for $(\mathcal{O}, \gamma)$. With each pair $\gamma, \delta \in \mathcal{D}$, we associate polynomials $R_{\gamma,\delta}, P_{\gamma,\delta} \in \mathbb{Z}[q]$. The R-polynomials can be computed using a recursive procedure which we refrain from stating in full generality here; see [Vog83, Lemma 6.8] for details. A special case sufficient for our purposes is formulated in Proposition 5.2 below.

Let $\mathcal{M}$ denote the free $\mathbb{Z}[q, q^{-1}]$ module with basis $\mathcal{D}$. For fixed $\delta \in \mathcal{D}$, we have in $\mathcal{M}$ the identity

$$q^{-l(\delta)} \sum_{\gamma \leq \delta} P_{\gamma,\delta}(q)\gamma = \sum_{\beta \leq \gamma \leq \delta} (-1)^{l(\beta)-l(\gamma)} q^{-l(\gamma)} P_{\gamma,\delta}(q^{-1}) R_{\beta,\gamma}(q)\beta$$

which subject to the restrictions $P_{\gamma,\gamma} = 1$ and $\deg(P_{\gamma,\delta}) \leq (l(\delta) - l(\gamma) - 1)/2$ uniquely determines the *Kazhdan-Lusztig-Vogan (KLV)* polynomials $P_{\gamma,\delta}$ [Vog83, Corollary 6.12].(i) Here, $l(\cdot)$ indicates the dimension of the corresponding orbit, and the order on $\mathcal{D}$ is the *Bruhat $\mathcal{G}$-order* [Vog83, Definition 5.8].

KLV polynomials serve as measures of the singularities of symmetric orbit closures; *cf.* [Vog83, Theorem 1.12]. In particular, their coefficients are nonnegative. Another consequence is the following:

Proposition 2.1 *Let $\leq$ denote the order relation in V,* i.e. *containment among orbit closures. Given orbits $\mathcal{P}, \mathcal{O} \in K\backslash X$ with $\mathcal{P} \leq \mathcal{O}$, let $\delta = (\mathcal{O}, \mathbb{C}_{\mathcal{O}})$, where $\mathbb{C}_{\mathcal{O}}$ is the trivial local system. Then, $\overline{\mathcal{O}}$ is rationally smooth at some (equivalently, every) point in $\mathcal{P}$ if and only if*

$$P_{\gamma,\delta} = \begin{cases} 1 & \text{if } L = \mathbb{C}_{\mathcal{Q}}, \\ 0 & \text{if } L \neq \mathbb{C}_{\mathcal{Q}}, \end{cases}$$

for all $\gamma = (\mathcal{Q}, L) \in \mathcal{D}$ with $\mathcal{P} \leq \mathcal{Q} \leq \mathcal{O}$.

The gadgets just described are fundamental ingredients in the representation theory of $G_{\mathbb{R}}$. More precisely, the KLV polynomials govern the transition between important families of $(\mathfrak{g}, K_{\mathbb{R}})$-modules. See [LV83, Vog83] for more details.

(i) Note that there is a typo which has an impact on the cited result. We are grateful to D. A. Vogan for pointing out that the displayed formula in the statement of [Vog83, Lemma 6.8] should read

$$D(\delta) = u^{-l(\delta)} \sum_{\gamma} (-1)^{l(\gamma)-l(\delta)} R_{\gamma,\delta}(u)\gamma.$$

3 Restricting the involution

Consider the set $\mathcal{V} = \{g \in G \mid \theta(g^{-1})g \in N\}$. The set of orbits $K\backslash\mathcal{V}/T$ parametrises $K\backslash X$. In this way, the map $\mathcal{V} \to W$ given by $g \mapsto \theta(g^{-1})gT$ induces the map $\varphi : V \to W$ which was mentioned in the introduction. Observe that the image of φ is contained in $\mathcal{I}(\theta)$.

Throughout this paper we shall only allow certain choices of θ. More precisely, *we from now on assume that θ obeys the following condition:*

Hypothesis 3.1 *The fixed point subgroup K is connected. Moreover, $\varphi : V \to W$ satisfies $\varphi(v_0) \in \iota(\theta)$, where $v_0 \in V$ is the maximum element,* i.e. *the dense orbit.*

Remark 3.2 If G is semisimple and simply connected, then K is necessarily connected. This result is due to Steinberg [Ste68, Theorem 8.1]. In some sense, the general situation can be reduced to the study of semisimple simply connected G; see [RS90].

Several consequences are collected in the next proposition. For the proof, see [Hul09]. Let Φ denote the root system of G, T and write $R \subset W$ for the corresponding set of reflections.

Proposition 3.3 *Hypothesis 3.1 implies the following:*

(i) The map φ yields a poset isomorphism $V \to \mathrm{Br}(\iota(\theta))$.

(ii) There is a unique K-equivariant local system, namely $\mathbb{C}_{\mathcal{O}}$, on each orbit $\mathcal{O} \in K\backslash X$. In particular, the sets $\mathcal{D}$, $K\backslash X$ and $\iota(\theta)$ may be identified, and the Bruhat $\mathcal{G}$-order on $\mathcal{D}$ coincides with V and $\mathrm{Br}(\iota(\theta))$.

(iii) Let $\alpha \in \Phi$ and denote by $G_\alpha \subseteq G$ the corresponding rank one semisimple group. Then, we are in one of the following two situations:

  (a) The root α is compact imaginary. *That is, $G_\alpha \subseteq K$.*

  (b) The root α is complex *(meaning $\theta(\alpha) \neq \alpha$) and $\theta(\alpha) + \alpha \notin \Phi$.*

(iv) If $r \in R$, then $\theta(r)r = r\theta(r)$.

(v) The poset $\mathrm{Br}(\iota(\theta))$ is graded with rank function ρ being half the ordinary Coxeter length. Moreover, $\rho(w) = l(\mathcal{O}_w) - l(\mathcal{O}_{\mathrm{id}})$.

The following example allows us to consider many of our results as generalisations of statements about Schubert varieties.

Example 3.4 If G' is a connected reductive complex linear algebraic group and $G = G' \times G'$, the involution θ which interchanges the two factors makes K the diagonal subgroup. In this case, $\iota(\theta) = \mathcal{I}(\theta)$, so Hypothesis 3.1 is satisfied. The poset $\mathrm{Br}(\iota(\theta))$ coincides with $\mathrm{Br}(W')$, where W' is the Weyl group of G'. There is a one-to-one correspondence between K-orbits in X and Schubert cells in the Bruhat decomposition of the flag variety of G' which preserves a lot of structure including the property of having rationally smooth closure at a given orbit.

In addition to the setting in Example 3.4 there are a few more cases that satisfy Hypothesis 3.1. They are denoted AII, DII and EIV in the classification of symmetric spaces $G_{\mathbb{R}}/K_{\mathbb{R}}$ given *e.g.* in Helgason [Hel78].[(ii)] The corresponding Weyl groups are A_{2n+1}, D_n and E_6, respectively, with θ in each case restricting to the Weyl group as the unique nontrivial Dynkin diagram involution. Types D and E could in principle be handled separately. In the former case, $\iota(\theta)$ has a very simple structure (*cf.* [Hul08, proof of Theorem 5.2]), whereas the latter admits a brute force computation. Thus, the main substance lies in the A_{2n+1} case where $\mathrm{Br}(\iota(\theta))$ is an incarnation of the containments among closures of $\mathrm{Sp}_{2n}(\mathbb{C})$ orbits in the flag variety $\mathrm{SL}_{2n}(\mathbb{C})/B$; see [RS90, Example 10.4] for a discussion of this case. Nevertheless, we have opted to keep our arguments type independent regarding all assertions that are valid in the full generality of Hypothesis 3.1. There are two reasons. First, the natural habitat for Theorems 5.7 and 5.8 is the general setting; no simplicity would be gained by formulating the arguments in type A specific terminology. Second, we hope that the less specialised viewpoint shall prove suitable as point of departure for generalisations beyond Hypothesis 3.1.

4 "Bruhat graphs"

Let $*$ denote the θ-twisted right conjugation action of W on itself, *i.e.* $u * w = \theta(w^{-1})uw$ for $u, w \in W$. Then $\iota(\theta)$ is the orbit of the identity element $\mathrm{id} \in W$.

Recall that $I_w = \{u \in \iota(\theta) \mid u \leq w\}$.

Definition 4.1 *Given $w \in \iota(\theta)$, let $\mathrm{BG}(w)$ be the graph with vertex set I_w and an edge $\{u, v\}$ whenever $u = v * t \neq v$ for some reflection $t \in R$.*

Notice that $\mathrm{BG}(u)$ is an induced subgraph of $\mathrm{BG}(w)$ if $u \leq w$. See Figure 1 for an illustration.

We shall refer to graphs of the form $\mathrm{BG}(w)$ as *Bruhat graphs*, because in the setting of Example 3.4, they coincide with (undirected versions of) the ordinary Bruhat graphs in W' introduced by Dyer [Dye91].

Next, we list some useful properties of Bruhat graphs. The proofs rely on combinatorial considerations and results from [Dye91]; see [Hul09] for details.

Lemma 4.2 *Let $w \in \iota(\theta)$ and $u, v \in I_w$, $u \neq v$. Write $u = \theta(x^{-1})x$ for $x \in W$. The following are equivalent:*

(i) *$\{u, v\}$ is an edge in $\mathrm{BG}(w)$.*

(ii) *There are exactly two distinct reflections $t \in R$ such that $u * t = v$.*

(iii) *There are exactly two distinct reflections $t \in R$ such that $\theta(x^{-1})\theta(t)tx = v$. If t is one of these reflections, then $\theta(t)$ is the other.*

Lemma 4.3 *If $\{u, v\}$ is an edge in $\mathrm{BG}(w)$, then either $u < v$ or $v < u$. Furthermore, v has exactly $\rho(v)$ neighbours u such that $u < v$.*

Combining the first part of Brion's [Bri99, Theorem 2.5] with part (iii) of Proposition 3.3 shows that the rank of a vertex $v = \theta(x^{-1})x$ in $\mathrm{BG}(w)$ is at most half the number of complex reflections (*i.e.* reflections that correspond to complex roots) $t \in R$ such that $\theta(x^{-1})\theta(t)tx \leq w$. By Lemma 4.2, this is precisely the degree of v in $\mathrm{BG}(w)$. We thus have the following fact:

(ii) The "usual" construction of DII would yield $G = \mathrm{SO}_{2n}(\mathbb{C})$, $K = \mathrm{S}(\mathrm{O}_{2n-1}(\mathbb{C}) \times \mathrm{O}_1(\mathbb{C})) \cong \mathrm{O}_{2n-1}(\mathbb{C})$ so that K is disconnected. However, passing to the fundamental cover, we have $G = \mathrm{Spin}_{2n}(\mathbb{C})$, $K = \mathrm{Spin}_{2n-1}(\mathbb{C})$ in agreement with Hypothesis 3.1.

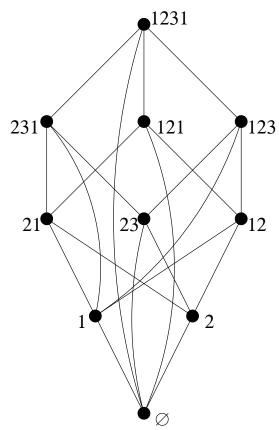


Fig. 1: A picture of the Bruhat graph $\mathrm{BG}(w)$ where $w = s_5s_3s_4s_5s_1s_2s_3s_1 \in \iota(\theta) \subset A_5$. Here, s_i denotes the simple reflection $(i, i+1)$ in the usual manifestation of A_5 as the symmetric group S_6. The involution θ sends s_{6-i} to s_i. A vertex $u \in I_w$ is labelled by the indices of a sequence of simple reflections whose product x satisfies $u = \theta(x^{-1})x$. The straight edges indicate the covering relation of $\mathrm{Br}(\iota(\theta))$.

Theorem 4.4 *For $w \in \iota(\theta)$, the degree of each vertex in $\mathrm{BG}(w)$ is at least $\rho(w)$.*

Remark 4.5 In the setting of Example 3.4, Theorem 4.4 specialises to "Deodhar's inequality" in W'; see [BL00, §6] and the references cited there.

5 A criterion for rational smoothness

In general, the recursion for the R-polynomials mentioned in Section 2 is technically rather involved. Since we are assuming Hypothesis 3.1, however, the situation is simpler. Proposition 3.3 allows us to identify the indexing set $\mathcal{D}$ with $\iota(\theta)$. Rather than working with the actual R-polynomials as defined in [Vog83], we shall find it more convenient to use the following simple variation:

Definition 5.1 *For $u, v \in \iota(\theta)$, let $Q_{u,v}(q) = (-q)^{\rho(v)-\rho(u)}R_{u,v}(q^{-1})$.*

With some labour, a combinatorially explicit recursion for the $Q_{u,v}$ can be extracted from the identity

$$\sum_{u\in\iota(\theta)} (-1)^{\rho(u)}R_{u,w}(q)u = -\sum_{u\in\iota(\theta)} (-1)^{\rho(u)}R_{u,w*s}(q)(T_s+1-q)u;$$

see [Vog83, proof of Lemma 6.8]. The key is that the definition of the maps T_s (see [Vog83, Definition 6.4]) simplifies a fair amount under Hypothesis 3.1. We refer the reader to [Hul09] for the details.

With $D_{\mathrm{R}}(v)$ denoting the *descent set* of $v \in \iota(\theta)$, *i.e.* the set of simple reflections s such that $vs < v$, or equivalently $v * s < v$, the recursion takes the following form:

Proposition 5.2 *For $s \in D_{\mathrm{R}}(v)$, we have*

$$Q_{u,v}(q) = \begin{cases} Q_{u*s,v*s}(q) & \text{if } u*s<u, \\ qQ_{u*s,v*s}(q)+(q-1)Q_{u,v*s}(q) & \text{if } u*s>u, \\ qQ_{u,v*s}(q) & \text{if } u*s=u. \end{cases}$$

Together with the "initial values" $Q_{u,u}(q) = 1$ and $Q_{u,v}(q) = 0$ if $u \not\leq v$, we may calculate any $Q_{u,v}$ using Proposition 5.2.

In the setting of Example 3.4, both the $R_{u,v}(q)$ and the $Q_{u,v}(q)$ coincide with the classical Kazhdan-Lusztig R-polynomials introduced in [KL79]. The three lemmata coming up next hint that the $Q_{u,v}(q)$ may provide the more useful generalisation.

Lemma 5.3 *For $u, v \in \iota(\theta)$, we have*

$$Q'_{u,v}(1) = \begin{cases} 1 & \textit{if } u < v \textit{ and } \{u,v\} \textit{ is an edge in } \mathrm{BG}(v), \\ 0 & \textit{otherwise.} \end{cases}$$

Proof: Suppose $s \in D_{\mathrm{R}}(v)$. Differentiating the equation in Proposition 5.2 with respect to q, and using that $Q_{u,v}(1) = R_{u,v}(1) = \delta_{u,v}$ (Kronecker's delta), it follows that

$$Q'_{u,v}(1) = Q'_{u*s,v*s}(1) + \delta_{u,v*s}.$$

It is clear that $\{u * s, v * s\}$ is an edge in $\mathrm{BG}(v)$ if and only if the same is true about $\{u, v\}$. Employing induction on $\rho(v)$, it thus suffices to show that $u * s < v * s$ if $v * s \neq u < v$ and $\{u, v\}$ is an edge. Lemma 4.3 shows that $u * s$ and $v * s$ are comparable in this situation. The assertion $u * s > v * s$ would contradict the Lifting Property [Hul08, Lemma 2.7], and we are done. □

Lemma 5.4 *Denote by μ the Möbius function of $\mathrm{Br}(\iota(\theta))$. Then, $\mu(u, v) = Q_{u,v}(0)$ for all $u, v \in \iota(\theta)$.*

Proof: Let us induct on $\rho(v)$. The assertion holds for $\rho(v) = 0$ because $Q_{\mathrm{id,id}}(q) = R_{\mathrm{id,id}}(q) = 1$. We shall demonstrate that $\mu(u, v)$ satisfies the recursion for $Q_{u,v}(0)$ derived from Proposition 5.2.

Borrowing terminology from [Hul08], call $[u, v]$ *full* if every twisted involution in the interval $[u, v]$ is in fact a twisted identity. Combining Philip Hall's theorem (see *e.g.* [Sta97, Proposition 3.8.5]) with the topological results in [Hul08, Theorem 4.12] shows that

$$\mu(u, v) = \begin{cases} (-1)^{\rho(v)-\rho(u)} & \text{if } [u, v] \text{ is full,} \\ 0 & \text{otherwise.} \end{cases}$$

Pick $s \in D_{\mathrm{R}}(v)$. In case $u * s = u$, $[u, v]$ is not full, and $\mu(u, v) = 0$ as desired. If $u * s > u$, it follows from [Hul08, Lemma 4.10] that $[u, v * s]$ is full if and only if $[u, v]$ is full. Thus, $\mu(u, v) = -\mu(u, v * s)$, and we are done. Finally, suppose $u * s < u$. If $[u * s, v * s]$ is full then $[u, v]$ is also full, again by [Hul08, Lemma 4.10]. On the other hand, [Hul08, Theorem 4.9] implies that $\mu(u * s, v) = -\mu(u, v)$, so if $[u * s, v * s]$ (and therefore $[u * s, v]$) is not full, then $[u, v]$ cannot be full either. Completing the proof, we conclude $\mu(u, v) = \mu(u * s, v * s)$. □

Lemma 5.5 *For all $v \in \iota(\theta)$,*

$$\sum_{u \leq v} Q_{u,v}(q) = q^{\rho(v)}.$$

Proof: We prove the lemma using induction on $\rho(v)$. Given $s \in D_{\mathrm{R}}(v)$, partition I_v into three sets:

$$A = \{u \leq v \mid u * s < u\},\ B = \{u \leq v \mid u * s > u\},\ C = \{u \leq v \mid u * s = u\}.$$

By the Lifting Property [Hul08, Lemma 2.7], the map $u \mapsto u * s$ is a bijection between A and B. The recursion in Proposition 5.2 therefore yields

$$\begin{aligned}\sum_{u \leq v} Q_{u,v}(q) &= \sum_{\substack{u \in A \\ u \leq v*s}} qQ_{u,v*s}(q) + \sum_{\substack{u \in B \\ u \leq v*s}} (1+q-1)Q_{u,v*s}(q) + \sum_{\substack{u \in C \\ u \leq v*s}} qQ_{u,v*s}(q) \\ &= q \sum_{u \leq v*s} Q_{u,v*s}(q),\end{aligned}$$

proving the claim. □

Lemma 5.6 *We have $P_{u,v}(0) = 1$ whenever $u \leq v$ in $\iota(\theta)$.*

Proof: The assertion is clear if $u = v$, and we employ induction on $\rho(v) - \rho(u)$.
Vogan's [Vog83, Corollary 6.12] translates to

$$q^{\rho(v)-\rho(u)} P_{u,v}(q^{-1}) = \sum_{u \leq w \leq v} Q_{u,w}(q) P_{w,v}(q).$$

The left hand side is a polynomial with zero constant term. Hence, Lemma 5.4 implies

$$P_{u,v}(0) = -\sum_{u < w \leq v} \mu(u,w) = \mu(u,u) = 1.$$

□

We are finally in position to prove the main results. Since all necessary technical prerequisites have been established, the corresponding arguments from [Car94] can now be transferred to our setting more or less verbatim.

Theorem 5.7 *Suppose $u, v \in \iota(\theta)$, $u \leq w$. The following conditions are equivalent:*

(i) The degree of v in $\mathrm{BG}(w)$ is $\rho(w)$ for all $u \leq v \leq w$.

(ii) The KLV polynomials satisfy $P_{v,w}(q) = 1$ for all $u \leq v \leq w$. That is, the orbit closure $\overline{\mathcal{O}_w}$ is rationally smooth at $\mathcal{O}_u$.

Proof: Define
$$f_{u,w}(q) = q^{\rho(w)-\rho(u)}(P_{u,w}(q^{-2}) - 1).$$
The P-polynomials have nonnegative coefficients. By Lemma 5.6, $f_{u,w}(q)$ too is a polynomial with nonnegative coefficients. Since it has vanishing constant term, $f'_{u,w}(1) = 0$ if and only if $f_{u,w}(q) = 0$ which, in turn, is equivalent to $P_{u,w}(q) = 1$.
Now,
$$f'_{u,w}(1) = (\rho(w) - \rho(u))(P_{u,w}(1) - 1) - 2P'_{u,w}(1).$$

Since $Q_{u,w}(1) = \delta_{u,w}$, we have

$$\begin{aligned}-2P'_{u,w}(1) &= \frac{d}{dq}P_{u,w}(q^{-2})|_{q=1}\\ &= 2(\rho(u) - \rho(w))P_{u,w}(1) + 2\sum_{u\le v\le w} Q'_{u,v}(1)P_{v,w}(1) + 2P'_{u,w}(1).\end{aligned}$$

Hence,

$$f'_{u,w}(1) = \rho(u) - \rho(w) + \sum_{u\le v\le w} Q'_{u,v}(1)P_{v,w}(1).$$

To begin with, assume (ii) holds. Then,

$$\rho(w) - \rho(v) = \sum_{v\le v'\le w} Q'_{v,v'}(1)$$

for all $u \le v \le w$. Condition (i) now follows from Lemma 5.3 together with Lemma 4.3.

Finally, let us prove (i) $\Rightarrow$ (ii) by induction on $\rho(w) - \rho(u)$. Suppose $u < v \le w$ in $\mathrm{Br}(\iota(\theta))$. By Lemma 5.3 and the induction assumption, $Q'_{u,v}(1)P_{v,w}(1)$ is one if $\{u, v\}$ is an edge in $\mathrm{BG}(w)$, zero otherwise. Since $\deg(u) = \rho(w)$, u has exactly $\rho(w) - \rho(u)$ neighbours v such that $u < v$. We conclude $f'_{u,w}(1) = 0$ as desired. □

Theorem 5.8 *For $w \in \iota(\theta)$, the following are equivalent:*

(i) For all i, $[\mathrm{id}, w] = \mathrm{Br}(I_w)$ has equally many elements of rank i as of rank $\rho(w) - i$.

(ii) The graph $\mathrm{BG}(w)$ is regular.

(iii) $P_{u,w}(q) = 1$ for all $u \le w$.

Proof: (i) $\Rightarrow$ (ii): Let $n(i)$ denote the number of elements of rank i in $[e, w]$. Now, using Lemma 4.3 and Theorem 4.4, we count the edges in $\mathrm{BG}(w)$ in two ways and obtain

$$\sum_{i=0}^{\rho(w)} n(i)i \ge \sum_{i=0}^{\rho(w)} n(i)(\rho(w) - i)$$

with equality if and only if $\mathrm{BG}(w)$ is $\rho(w)$-regular. However, if $n(i) = n(\rho(w)-i)$ for all i, then equality does hold.

(ii) $\Rightarrow$ (iii): This follows from Theorem 5.7.

(iii) $\Rightarrow$ (i): We claim that

$$F_w(q) = \sum_{u\le w} P_{u,w}(q)q^{\rho(u)}$$

is a symmetric polynomial, *i.e.* $F_w(q) = q^{\rho(w)}F_w(q^{-1})$. If the P-polynomials all are 1, this means

$$\sum_{u\le w} q^{\rho(u)} = \sum_{u\le w} q^{\rho(w)-\rho(u)}.$$

It therefore remains to verify the claim. Observe that

$$q^{\rho(w)}F_w(q^{-1}) = \sum_{u\le w} q^{\rho(w)-\rho(u)}P_{u,w}(q^{-1}) = \sum_{u\le w}\sum_{u\le v\le w} Q_{u,v}(q)P_{v,w}(q) = \sum_{v\le w} P_{v,w}(q)\sum_{u\le v} Q_{u,v}(q).$$

The claim now follows from Lemma 5.5. □

To illustrate these results, consider Figure 1. The interval $[\mathrm{id}, w]$ has three elements of rank three but only two of rank $\rho(w) - 3 = 1$. By Theorem 5.8, $\overline{\mathcal{O}_w}$ is rationally singular. A more careful inspection of the graph shows that s_5s_1 and e both have degree five whereas all other vertices have degree $\rho(w) = 4$. By Theorem 5.7, the rationally singular locus of $\overline{\mathcal{O}_w}$ therefore is $\mathcal{O}_{s_5s_1} \cup \mathcal{O}_e$. Also, observe that the degree never decreases as we move down in the graph. This phenomenon is explained in the next section.

6 Sufficiency of the bottom vertex

In this final section, the criterion given in Theorem 5.7 is significantly improved in the special case $G = \mathrm{SL}_{2n}(\mathbb{C})$, $K = \mathrm{Sp}_{2n}(\mathbb{C})$. In that case, as we shall see, whether or not an orbit closure $\overline{\mathcal{O}_w}$ is rationally smooth at $\mathcal{O}_u$ is determined by the degree of u alone (Corollary 6.5 below). The corresponding statement for Schubert varieties is known to be true in type A [Deo85] but false in general (see [BG03] for some elaboration on this). Necessarily, therefore, this section must be type specific since the results cannot possibly extend to the situation in Example 3.4 for arbitrary G'.

We work in the set F_{2n} of fixed point free involutions on $\{1, \dots, 2n\}$. Let $\star$ denote the conjugation action from the right by the symmetric group S_{2n} on itself, *i.e.* $\sigma \star \pi = \pi^{-1}\sigma\pi$. Then, $F_{2n} = w_0 \star S_{2n}$, where w_0 is the reverse permutation $i \mapsto 2n + 1 - i$.

Let $\preceq$ denote the dual of the subposet of the Bruhat order on S_{2n} induced by F_{2n}. The bottom element of this poset is w_0. Observe that if $u \neq u \star t$, then $u \star t \succ u$ iff t is an *inversion* of u (meaning $t = (a, b)$ with $a < b$ and $u(a) > u(b)$).

For $w \in F_{2n}$, define the *Bruhat graph* $\mathrm{BG}(w)$ as the graph with vertex set $I_w = \{u \in F_{2n} \mid u \preceq w\}$ and an edge $\{u, v\}$ whenever $u \neq v = u \star t$ for some transposition t. Thus, each edge has exactly two transpositions associated with it, and the graph is simple (no loops or multiple edges). If w is understood from the context and $u \preceq w$, let $\mathcal{E}(u)$ denote the set of edges incident to u in $\mathrm{BG}(w)$. Also, define $\deg(u) = |\mathcal{E}(u)|$.

Proposition 6.1 *Suppose $W = A_{2n-1} \cong S_{2n}$ with $\theta : W \to W$ given by the unique nontrivial involution of the Dynkin diagram. Then, $x \mapsto w_0x$ defines a bijection $F_{2n} \to \iota(\theta)$. Moreover, the bijection is an isomorphism of Bruhat graphs,* i.e. $u \preceq w \Leftrightarrow w_0u \le w_0w$ *and* $w_0(w \star t) = w_0w * t$.

Proof: This is immediate from the well known facts that $\theta(x) = w_0xw_0$ and that $x \mapsto w_0x$ is an antiautomorphism of $\mathrm{Br}(W)$. □

Suppose $w \succeq u \neq w_0$ and let $r = (i, j)$, $i < j$, be a transposition such that $u \star r \prec u$. Let $a = u(i)$ and $b = u(j)$. Thus, $a < b \neq i$.

For a transposition $t = (x, y)$, we use the notation $\mathrm{supp}(t) = \{x, y\}$.

Definition 6.2 *Call a transposition t* compatible *(with respect to u and r) if either* $\mathrm{supp}(t) \cap \{i, j\} \neq \emptyset$ *or* $\mathrm{supp}(t) \cap \{a, b, i, j\} = \emptyset$.

Given an edge $e \in \mathcal{E}(u)$ there are precisely two transpositions t and $t' \neq t$ such that $e = \{u, u \star t\} = \{u, u \star t'\}$. At least one is compatible; let t_e be such a one.

Definition 6.3 *Given* $e \in \mathcal{E}(u)$*, define* $\epsilon(e) = \{u \star r, u \star r\tau_e\}$*, where*

$$\tau_e = \begin{cases} rt_e r & \textit{if } u \star t_e r \preceq w, \\ t_e & \textit{otherwise.} \end{cases}$$

It is not *a priori* clear that $\epsilon(e)$ is independent of the choice of t_e, but this turns out to be the case. Here is the point of all this:

Theorem 6.4 *Definition 6.3 defines an injective map* $\epsilon : \mathcal{E}(u) \to \mathcal{E}(u \star r)$.

The proof of Theorem 6.4 hinges on combinatorial considerations revolving around the Standard Criterion characterising Bruhat order in symmetric groups; see *e.g.* [BB05, Theorem 2.1.5]. The details can be found in [Hul09].

By Theorem 6.4, the degree can never decrease as we go down along edges in a Bruhat graph. In particular, if a vertex has the minimum possible degree, then so does every vertex above it:

Corollary 6.5 *We have* $\deg(v) = \deg(w)$ *for all* $u \preceq v \preceq w$ *if and only if* $\deg(u) = \deg(w)$.

Thus, to determine whether Condition (i) of Theorem 5.7 is satisfied, it suffices to check the degree of u.

Remark 6.6 The set $\mathcal{S}_{2n} = \{w \in F_{2n} \mid i \leq n \Rightarrow w(i) \geq n+1\}$ is in natural bijective correspondence with S_{2n} in a way which identifies $\mathrm{Br}(S_{2n})$ with $\preceq$. Restricted to $w \in \mathcal{S}_{2n}$, Corollary 6.5 specialises to a result of Deodhar [Deo85] for type A Schubert varieties. In that setting, our arguments are closely related to work of Billey and Warrington [BW03, §6]

Remark 6.7 Observe that for $G = \mathrm{SL}_{2n}(\mathbb{C})$, $K = \mathrm{Sp}_{2n}(\mathbb{C})$, Theorem 4.4 follows directly from Theorem 6.4. Thus, we have rederived Brion's [Bri99, Theorem 2.5] in this case.

Acknowledgements

The author is grateful to W. M. McGovern for many helpful discussions and to P. McNamara and E. Briand for translation assistance. Pardon my French.

References

[BL00] S. C. Billey and V. Lakshmibai, *Singular loci of Schubert varieties*, Progress in Mathematics **182**, Birkhäuser Boston, Inc., Boston, MA, 2000.

[BW03] S. C. Billey and G. S. Warrington, Maximal singular loci of Schubert varieties in $\mathrm{S}L(n)/B$, *Trans. Amer. Math. Soc.* **355** (2003), 3915–3945.

[BB05] A. Björner and F. Brenti, *Combinatorics of Coxeter groups*, Graduate Texts in Mathematics, Vol. 231, Springer-Verlag, New York, 2005.

[BG03] B. D. Boe and W. Graham, A lookup conjecture for rational smoothness, *Amer. J. Math.* **125** (2003), 317–356.

[Bri99] M. Brion, Rational smoothness and fixed points of torus actions, *Transform. Groups* **4** (1999), 127–156.

[Car94] J. B. Carrell, The Bruhat graph of a Coxeter group, a conjecture of Deodhar, and rational smoothness of Schubert varieties, *Proc. Sympos. Pure Math.* **56** (1994), 53–61.

[Deo85] V. V. Deodhar, Local Poincaré duality and nonsingularity of Schubert varieties, *Comm. Algebra* **13** (1985), 1379–1388.

[Dye91] M. Dyer, On the "Bruhat graph" of a Coxeter system, *Comp. Math.* **78** (1991), 185–191.

[Hel78] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics **80**, Academic Press, New York, 1978.

[Hul08] A. Hultman, Twisted identities in Coxeter groups, *J. Algebraic Combin.* **28** (2008), 313–332.

[Hul09] A. Hultman, Criteria for rational smoothness of some symmetric orbit closures, preprint 2009.

[KL79] D. Kazhdan and G. Lusztig, Representations of Coxeter groups and Hecke algebras, *Invent. Math.* **53** (1979), 165–184.

[LV83] G. Lusztig and D. A. Vogan, Jr., Singularities of closures of K-orbits on flag manifolds, *Invent. Math.* **71** (1983), 365–379.

[McG09] W. M. McGovern, Closures of K-orbits in the flag variety for $SU^*(2n)$, preprint 2009.

[Ric82] R. W. Richardson, Orbits, invariants and representations associated to involutions of reductive groups, *Invent. Math.* **66** (1982), 287–312.

[RS90] R. W. Richardson and T. A. Springer, The Bruhat order on symmetric varieties, *Geom. Dedicata* **35** (1990), 389–436.

[RS94] R. W. Richardson and T. A. Springer, Complements to: The Bruhat order on symmetric varieties, *Geom. Dedicata* **49** (1994), 231–238.

[Spr85] T. A. Springer, Some results on algebraic groups with involutions, *Advanced Studies in Pure Math.* **6**, 525–543, Kinokuniya/North-Holland, 1985.

[Spr94] T. A. Springer, A combinatorial result on K-orbits on a flag manifold, *The Sophus Lie Memorial Conference (Oslo, 1992), Scand. Univ. Press, Oslo*, 1994, 363–370.

[Sta97] R. P. Stanley, *Enumerative combinatorics*, vol. 1, Cambridge Univ. Press, 1997.

[Ste68] R. Steinberg, Endomorphisms of linear algebraic groups, *Mem. Amer. Math. Soc.* **80** (1968), 1–108.

[Vog83] D. A. Vogan, Jr., Irreducible characters of semisimple Lie groups III. Proof of Kazhdan-Lusztig conjecture in the integral case, *Invent. Math.* **71** (1983), 381–417.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 210–221

Combinatorics of the PASEP partition function

Matthieu Josuat-Vergès[1 †]

[1] *LRI, Bât. 490, Université Paris-sud 11, 91405 Orsay CEDEX*

Abstract. We consider a three-parameter PASEP model on N sites. A closed formula for the partition function was obtained analytically by Blythe et al. We give a new formula which generalizes the one of Blythe et al, and is proved in two combinatorial ways. Moreover the first proof can be adapted to give the moments of Al-Salam-Chihara polynomials.

Résumé. Nous considérons un modèle de PASEP à trois paramètres sur N sites. Une formule close pour la fonction de partition a été obtenue analytiquement par Blythe et al. Nous donnons une formule qui généralise celle de Blythe et al, prouvée combinatoirement de deux manières diffèrentes. Par ailleurs la première preuve peut être adaptée de sorte à obtenir les moments des polynômes d'Al-Salam-Chihara.

Keywords: asymmetric exclusion process, lattice paths, orthogonal polynomials, enumeration

1 Introduction

The partially asymmetric simple exclusion process (also called PASEP) is a Markov chain describing the evolution of particles in N sites arranged in a line, each site being either empty or occupied by one particle. Particles may enter the leftmost site at a rate $\alpha \geq 0$, go out the rightmost site at a rate $\beta \geq 0$, hop left at a rate $q \geq 0$ and hop right at a rate $p \geq 0$ when possible. By rescaling time it is always possible to assume that the latter parameter is 1 without loss of generality. It is possible to define either a continuous-time model or a discrete-time model, but they are equivalent in the sense that their stationary distributions are the same. In this work we only study some combinatorial properties of the partition function. For precisions, background about the model, and much more, we refer to [2, 3, 4, 5, 7, 9, 18]. We refer particularly to the long survey of Blythe and Evans [2] and all references therein to give evidence that this is a widely studied model. Indeed, it is quite rich and some important features are the various phase transitions, and spontaneous symmetry breaking for example, so that it is considered to be a fundamental model of nonequilibrium statistical physics.

A method to obtain the stationary distribution and the partition function Z_N of the model is the Matrix Ansatz of Derrida, Evans, Hakim and Pasquier [9]. We suppose that D and E are linear operators, $\langle W|$ is a vector, $|V\rangle$ is a linear form, such that:

$$DE - qED = D + E, \qquad \langle W|\alpha E = \langle W|, \qquad \beta D|V\rangle = |V\rangle, \qquad \langle W|V\rangle = 1, \tag{1}$$

†Partially supported by the grant ANR08-JCJC-0011.

then the non-normalized probability of each state can be obtained by taking the product $\langle W|t_1 \dots t_N|V\rangle$ where t_i is D if the ith site is occupied and E if it is empty. It follows that the normalization, or partition function, is given by $\langle W|(D+E)^N|V\rangle$. It is possible to introduce another variable y, which is not a parameter of the probabilistic model, but is a formal parameter such that the coefficient of y^k in the partition function corresponds to the states with exactly k particles (physically it could be called a *fugacity*). The partition function is then:

$$Z_N = \langle W|(yD+E)^N|V\rangle, \tag{2}$$

which we may take as a definition on the combinatorial point of view. An interesting property is the symmetry:

$$Z_N(\alpha, \beta, y, q) = y^N Z_N(\beta, \alpha, y^{-1}, q), \tag{3}$$

which can be seen on the physical point of view by exchanging the empty sites with occupied sites. It can also be obtained from the Matrix Ansatz by using the transposed matrices D^* and E^* and the transposed vectors $\langle V|$ and $|W\rangle$, which satisfies a similar Matrix Ansatz with α and β exchanged.

In section 2, we will use the explicit solution of the Matrix Ansatz found by Derrida & al. [9], and it will permit to make use of weighted lattice paths as in [4].

An exact formula for Z_N was given by Blythe & al. [3, Equation (57)] in the case where $y = 1$. It was obtained from the eigenvalues and eigenvectors of the operator $D + E$ as defined in (10) and (11) below. This method gives an integral form for Z_N, which can be simplified so as to obtain a finite sum rather than an integral. Moreover this expression for Z_N was used to obtain various properties of the large system size limit, such as phases diagrams and currents. Here we generalize this result since we also have the variable y, and the proofs are combinatorial. This is an important result since most interesting properties of a model can be derived from the partition function. The interest of the result is also due to the plentiful combinatorial information of Z_N [7], in the full version of this work we will show that it is the generating function of permutations in $\mathfrak{S}_{N+1}$ with respect to right-to-left minima, right-to-left maxima, ascents, and occurrences of the pattern 31-2 (see [6] for a close result).

Theorem 1.1 *Let $\tilde{\alpha} = (1-q)\frac{1}{\alpha} - 1$ and $\tilde{\beta} = (1-q)\beta - 1$. We have:*

$$Z_N = \frac{1}{(1-q)^N} \sum_{n=0}^{N} R_{N,n}(y,q) B_n(\tilde{\alpha}, \tilde{\beta}, y, q), \tag{4}$$

where

$$R_{N,n}(y,q) = \sum_{i=0}^{\lfloor \frac{N-n}{2} \rfloor} (-y)^i q^{\binom{i+1}{2}} {\textstyle\begin{bmatrix} n+i \\ i \end{bmatrix}_q} \sum_{j=0}^{N-n-2i} y^j \left(\tbinom{N}{j}\tbinom{N}{n+2i+j} - \tbinom{N}{j-1}\tbinom{N}{n+2i+j+1} \right) \tag{5}$$

and

$$B_n(\tilde{\alpha}, \tilde{\beta}, y, q) = \sum_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix}_q \tilde{\alpha}^k (y\tilde{\beta})^{n-k}. \tag{6}$$

In the case where $y = 1$, one sum can be simplified by the Vandermonde identity $\sum_j \binom{N}{j}\binom{N}{m-j} = \binom{2N}{m}$, and we recover the expression given in [3, Equation (54)] by Blythe & al:

$$R_{N,n}(1,q) = \sum_{i=0}^{\lfloor \frac{N-n}{2} \rfloor} (-1)^i \left(\binom{2N}{N-n-2i} - \binom{2N}{N-n-2i-2} \right) q^{\binom{i+1}{2}} {\begin{bmatrix} n+i \\ i \end{bmatrix}}_q. \tag{7}$$

In the case where $\alpha = \beta = 1$, it is known [8, 13] that:

$$Z_N = \frac{1}{(1-q)^{N+1}} \sum_{k=0}^{N+1} (-1)^k \left(\sum_{j=0}^{N+1-k} y^j \left(\binom{N+1}{j}\binom{N+1}{j+k} - \binom{N+1}{j-1}\binom{N+1}{j+k+1} \right) \right) \left(\sum_{i=0}^{k} y^i q^{i(k+1-i)} \right) \tag{8}$$

(see Remarks 2.4 and 3.3 for a comparison between this previous result and the new one in Theorem 1.1). And in the case where $y = q = 1$, from a recursive construction of permutation tableaux [6] or lattice paths combinatorics [4] it is known that :

$$Z_N = \prod_{i=0}^{N-1} \left(\frac{1}{\alpha} + \frac{1}{\beta} + i \right). \tag{9}$$

Our first proof of (4) is a purely combinatorial enumeration of some weighted Motzkin paths defined below in (13), appearing from explicit representations of the operators D and E of the Matrix Ansatz. It partially relies on results of [8, 13] through Proposition 2.1 below. In contrast, our second proof does not use a particular representation of the operators D and E, but only on the combinatorics of the normal ordering process. It also relies on previous results of [13] (through Proposition 3.1 below), but we will sketch a self-contained proof. Additionally we will show that our first proof of Theorem 1.1 can be adapted to give a formula for Al-Salam-Chihara moments [1].

2 A first combinatorial derivation of Z_N using lattice paths

We use the solution of the Matrix Ansatz (1) given by Derrida & al. [9]. Let $\tilde{\alpha} = (1-q)\frac{1}{\alpha} - 1$ and $\tilde{\beta} = (1-q)\frac{1}{\beta} - 1$, their matrices are $D = (D_{i,j})_{i,j\in\mathbb{N}}$ and $E = (E_{i,j})_{i,j\in\mathbb{N}}$ with coefficients :

$$(1-q)D_{i,i} = 1 + \tilde{\beta}q^i, \qquad (1-q)D_{i,i+1} = 1 - \tilde{\alpha}\tilde{\beta}q^i, \tag{10}$$

$$(1-q)E_{i,i} = 1 + \tilde{\alpha}q^i, \qquad (1-q)E_{i+1,i} = 1 - q^{i+1}, \tag{11}$$

all other coefficients being 0, and vectors:

$$\langle W| = (1,0,0,\dots), \qquad |V\rangle = (1,0,0,\dots)^*, \tag{12}$$

(*i.e.* $|V\rangle$ is the transpose of $\langle W|$). Even if infinite-dimensional, they have the nice property of being tridiagonal and this lead to a combinatorial interpretation of Z_N in terms of lattice paths [4]. Indeed, we can see $yD + E$ as a transfer matrix for walks in the non-negative integers, and from (2) we obtain that $(1-q)^N Z_N$ is the sum of weights of Motzkin paths of length N with weights:

- $1 - q^{h+1}$ for a step $\nearrow$ starting at height h,
- $(1+y) + (\tilde{\alpha} + y\tilde{\beta})q^h$ for a step $\rightarrow$ starting at height h, (13)
- $y(1 - \tilde{\alpha}\tilde{\beta}q^{h-1})$ for a step $\searrow$ starting at height h.

To give a bijective proof of Theorem 1.1, we need to consider the set $\mathfrak{P}_N$ of weighted Motzkin paths of length N such that:

- the weight of a step $\nearrow$ starting at height h is $q^i - q^{i+1}$ for some $i \in \{0, \dots, h\}$,
- the weight of a step $\rightarrow$ starting at height h is either $1 + y$ or $(\tilde{\alpha} + y\tilde{\beta})q^h$,
- the weight of a step $\searrow$ starting at height h is either y or $-y\tilde{\alpha}\tilde{\beta}q^{h-1}$.

The sum of weights of elements in $\mathfrak{P}_N$ is $(1-q)^N Z_N$ because the weights sum to the ones in (13). We stress that on the combinatorial point of view, it will be important to distinguish $(h+1)$ kinds of step $\nearrow$ starting at height h, instead of one kind of step $\nearrow$ with weight $1 - q^{h+1}$.

We will show that each element of $\mathfrak{P}_N$ bijectively corresponds to a pair of weighted Motzkin paths. The first path (respectively, second path) belongs to a set whose generating function is $R_{N,n}(y,q)$ (respectively, $B_n(\tilde{\alpha}, \tilde{\beta}, y, q)$) for some $n \in \{0, \dots, N\}$. Following this scheme, our first combinatorial proof of (4) is a consequence of Propositions 2.1, 2.2, and 2.3 below.

Let $\mathfrak{R}_{N,n}$ be the set of weighted Motzkin paths of length N such that:

- the weight of a step $\nearrow$ starting at height h is $q^i - q^{i+1}$ for some $i \in \{0, \dots, h\}$,
- the weight of a step $\rightarrow$ starting at height h is either $1 + y$ or q^h, and there are exactly n steps $\rightarrow$ weighted by a power of q,
- the weight of a step $\searrow$ is y,

Proposition 2.1 *The sum of weights of elements in $\mathfrak{R}_{N,n}$ is $R_{N,n}(y,q)$.*

This can be obtained with the methods used in [8, 13]. Some precisions are in order. In [8] and [13], we obtained the formula (8) which is the special case $\alpha = \beta = 1$ in Z_N, and is the Nth moment of q-Laguerre polynomials [15] which are a rescaling of Al-Salam-Chihara polynomials. Since Z_N is also very closely related with these polynomials (see Section 4) it is not surprising that some steps are in common between these previous results and the present ones. See also Remark 2.4 below.

Let $\mathfrak{B}_n$ be the set of weighted Motzkin paths of length n such that:

- the weight of a step $\nearrow$ starting at height h is $q^i - q^{i+1}$ for some $i \in \{0, \dots, h\}$,
- the weight of a step $\rightarrow$ starting at height h is $(\tilde{\alpha} + y\tilde{\beta})q^h$,
- the weight of a step $\searrow$ starting at height h is $-y\tilde{\alpha}\tilde{\beta}q^{h-1}$.

Proposition 2.2 *The sum of weights of elements in* $\mathfrak{B}_n$ *is* $B_n(\tilde{\alpha}, \tilde{\beta}, y, q)$.

It is a consequence of properties of the Al-Salam-Carlitz orthogonal polynomials. Indeed, a standard argument [10, 20] shows that $B_n(\tilde{\alpha}, \tilde{\beta}, y, q)$ is the nth moment of an orthogonal sequence whose three-term recurrence relation is derived from the weights in the Motzkin paths. These polynomials are a rescaled version of Al-Salam-Carlitz polynomials, whose moments are known [16]. The result follows.

Proposition 2.3 *There exists a weight-preserving bijection* Φ *between the disjoint union of* $\mathfrak{R}_{N,n} \times \mathfrak{B}_n$ *over* $n \in \{0, \dots, N\}$*, and* $\mathfrak{P}_n$ *(we understand that the weight of a pair is the product of the weights of each element).*

To define the bijection, we start from a pair $(H_1, H_2) \in \mathfrak{R}_{N,n} \times \mathfrak{B}_n$ for some $n \in \{0, \dots, N\}$ and build a path $\Phi(H_1, H_2) \in \mathfrak{P}_N$. Let $i \in \{1, \dots, N\}$.

- If the ith step of H_1 is a step $\rightarrow$ weighted by a power of q, say the jth one among the n such steps, then:
  - the ith step $\Phi(H_1, H_2)$ has the same direction as the jth step of H_2,
  - its weight is the product of weights of the ith step of H_1 and the jth step of H_2.
- Otherwise the ith step of $\Phi(H_1, H_2)$ has the same direction and same weight as the ith step of H_1.

See Figure 1 for an example, where the thick steps correspond to the ones in the first of the two cases considered above. It is immediate that the total weight of $\Phi(H_1, H_2)$ is the product of the total weights of H_1 and H_2.

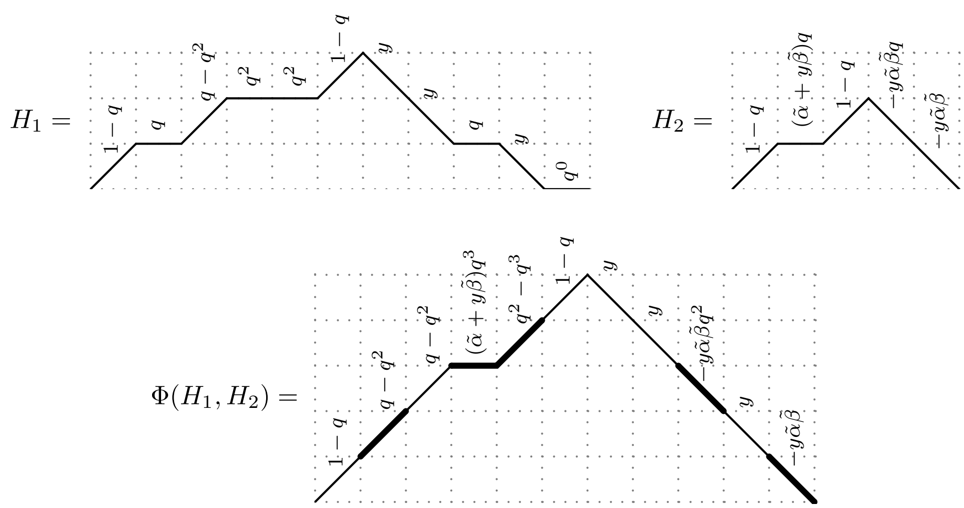


Fig. 1: Example of paths H_1, H_2 and their image $\Phi(H_1, H_2)$.

The inverse bijection is not as simple. It can be checked that H_1 and H_2 can be recovered by reading $\Phi(H_1, H_2)$ step by step from right to left.

Remark 2.4 The decomposition Φ is the key step in our first proof of Theorem 1.1. This makes the proof quite different from the one in the case $\alpha = \beta = 1$ [8], even though we have used results from [8] to prove an intermediate step (namely Proposition 2.1). Actually, it might be possible to have a direct adaptation of the case $\alpha = \beta = 1$ [8] to prove Theorem 1.1, but it should give rise to many computational steps. In contrast our decomposition Φ explains the formula for Z_N as a sum of products.

3 A second derivation of Z_N using the Matrix Ansatz

In this section we build on our previous work [13] to give a second proof of (4). In this reference we define the operators

$$\hat{D} = \frac{q-1}{q} D + \frac{1}{q} I \qquad \text{and} \qquad \hat{E} = \frac{q-1}{q} E + \frac{1}{q} I, \tag{14}$$

where I is the identity. The new relations for these operators are:

$$\hat{D}\hat{E} - q\hat{E}\hat{D} = \frac{1-q}{q^2}, \qquad \langle W|\hat{E} = -\frac{\tilde{\alpha}}{q}\langle W|, \qquad \text{and} \quad \hat{D}|V\rangle = -\frac{\tilde{\beta}}{q}|V\rangle, \tag{15}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ are defined as in the previous section. While the normal ordering problem for D and E leads to *permutation tableaux* [7], for $\hat{D}$ and $\hat{E}$ it leads to *rook placements* as was shown for example in [21]. The combinatorics of rook placements lead to the following proposition.

Proposition 3.1 *We have:*

$$\langle W|(qy\hat{D} + q\hat{E})^k|V\rangle = \sum_{\substack{i+j\le k \\ i+j\equiv k \text{ mod } 2}} \begin{bmatrix} i+j \\ i \end{bmatrix}_q (-\tilde{\alpha})^i(-y\tilde{\beta})^j M_{\frac{k-i-j}{2},k} \tag{16}$$

where

$$M_{\ell,k} = y^\ell \sum_{u=0}^{\ell} (-1)^u q^{\binom{u+1}{2}} \begin{bmatrix} k-2\ell+u \\ u \end{bmatrix}_q \left(\binom{k}{\ell-u} - \binom{k}{\ell-u-1} \right). \tag{17}$$

Proof: This is a consequence of results in [13] (see Section 2, Corollary 1, Proposition 12). We also give here a self-contained recursive proof. By means of the commutation relation in (15), we can write $(yq\hat{D} + q\hat{E})^k$ as a *normal form*:

$$(yq\hat{D} + q\hat{E})^k = \sum_{i,j\ge 0} d_{i,j}^{(k)} (q\hat{E})^i (qy\hat{D})^j, \tag{18}$$

where $d_{i,j}^{(k)}$ are polynomials in y and q, and only finitely many of them are non-zero. From the commutation relation we also obtain:

$$(qy\hat{D})^j(q\hat{E}) = q^j(q\hat{E})(qy\hat{D})^j + y(1-q^j)(qy\hat{D})^{j-1}. \tag{19}$$

If we multiply (18) by $yq\hat{D} + q\hat{E}$ to the right, using (19) we can get a recurrence relation for the coefficients $d_{i,j}^{(k)}$, which reads:

$$d_{i,j}^{(k+1)} = d_{i,j-1}^{(k)} + q^j d_{i-1,j}^{(k)} + y(1-q^{j+1})d_{i,j+1}^{(k)}. \tag{20}$$

The initial case is that $d^{(0)}_{i,j}$ is 1 if $(i,j) = (0,0)$ and 0 otherwise. It can be directly checked that the recurrence is solved by:

$$d^{(k)}_{i,j} = \begin{bmatrix} i+j \\ i \end{bmatrix}_q M_{\frac{k-i-j}{2},k} \tag{21}$$

where we understand that $M_{\frac{k-i-j}{2},k}$ is 0 when $k-i-j$ is not even. More precisely, if we let $e^{(k)}_{i,j} = \left[\begin{smallmatrix} i+j \\ i \end{smallmatrix}\right]_q M_{\frac{k-i-j}{2},k}$ then we have:

$$e^{(k)}_{i,j-1} + q^j e^{(k)}_{i-1,j} = \begin{bmatrix} i+j \\ i \end{bmatrix}_q M_{\frac{k-i-j+1}{2},k}, \tag{22}$$

and also

$$y(1-q^{j+1})e^{(k)}_{i,j+1} = y(1-q^{i+j+1}) \begin{bmatrix} i+j \\ i \end{bmatrix}_q M_{\frac{k-i-j-1}{2},k}. \tag{23}$$

So to prove $d^{(k)}_{i,j} = e^{(k)}_{i,j}$ it remains only to check that

$$M_{\frac{k-i-j+1}{2},k} + y(1-q^{i+j+1})M_{\frac{k-i-j-1}{2},k} = M_{\frac{k-i-j}{2},k}. \tag{24}$$

See for example [13, Proposition 12] (actually this recurrence already appeared more than fifty years ago in the work of Touchard, see *loc. cit.* for precisions). □

Now we can give our second proof of Theorem 1.1.

Proof: From (2) and (14) we have:

$$(1-q)^N Z_N = \langle W|((1+y)I - qy\hat{D} - q\hat{E})^N|V\rangle = \sum_{k=0}^{N} \binom{N}{k} (1+y)^{N-k}(-1)^k \langle W|(qy\hat{D} + q\hat{E})^k|V\rangle.$$

So, from Proposition 3.1 we have:

$$(1-q)^N Z_N = \sum_{k=0}^{N} \sum_{\substack{i+j\le k \\ i+j\equiv k \bmod 2}} \begin{bmatrix} i+j \\ i \end{bmatrix}_q \tilde{\alpha}^i (y\tilde{\beta})^j \binom{N}{k} (1+y)^{N-k} M_{\frac{k-i-j}{2},k}$$

(the $(-1)^k$ cancels with a $(-1)^{i+j}$). Setting $n = i+j$, we have:

$$(1-q)^N Z_N \quad = \quad \sum_{n=0}^{N} B_n(\tilde{\alpha},\tilde{\beta},y,q) \sum_{\substack{n\le k\le N \\ k\equiv n \bmod 2}} \binom{N}{k} (1+y)^{N-k} M_{\frac{k-n}{2},k}.$$

So it remains only to show that the latter sum is $R_{N,n}(y,q)$. If we change the indices so that k becomes $n+2k$, this sum is:

$$\sum_{k=0}^{\lfloor \frac{N-n}{2} \rfloor} \tbinom{N}{n+2k} (1+y)^{N-n-2k} y^k \sum_{i=0}^{k} (-1)^i q^{\binom{i+1}{2}} \begin{bmatrix} n+i \\ i \end{bmatrix}_q \left(\tbinom{n+2k}{k-i} - \tbinom{n+2k}{k-i-1} \right)$$

$$= \sum_{i=0}^{\lfloor \frac{N-n}{2} \rfloor} (-y)^i q^{\binom{i+1}{2}} \begin{bmatrix} n+i \\ i \end{bmatrix}_q \sum_{k=i}^{\lfloor \frac{N-n}{2} \rfloor} y^{k-i} \binom{N}{n+2k} (1+y)^{N-n-2k} \left(\binom{n+2k}{k-i} - \binom{n+2k}{k-i-1} \right).$$

We can simplify the latter sum by Lemma 3.2 below and obtain $R_{N,n}(y,q)$. This completes the proof. □

Lemma 3.2 *For any $N, n, i \geq 0$ we have:*

$$\sum_{k=i}^{\lfloor \frac{N-n}{2} \rfloor} y^{k-i} \binom{N}{n+2k} (1+y)^{N-n-2k} \left(\binom{n+2k}{k} - \binom{n+2k}{k-1} \right)$$
$$= \sum_{j=0}^{N-n-2i} y^j \left(\binom{N}{j} \binom{N}{n+2i+j} - \binom{N}{j-1} \binom{N}{n+2i+j+1} \right). \quad (25)$$

Proof: It can be shown that the right-hand side of (25) is the number of Motzkin prefixes of length N, final height $n+2i$, and a weight $1+y$ on each step $\rightarrow$ and y on each step $\searrow$. Similarly, $y^{k-i}(\binom{n+2k}{k-i} - \binom{n+2k}{k-i-1})$ is the number of Dyck prefixes of length $n+2k$ and final height $n+2i$, with a weight y on each step $\searrow$. From these two combinatorial interpretations it is straightforward to obtain a bijective proof of (25). Each Motzkin prefix is built from a shorter Dyck prefix with the same final height, by choosing where are the $N-n-2k$ steps $\rightarrow$. □

Remark 3.3 All the ideas in this proof were present in [13] where we obtained the case $\alpha = \beta = 1$. The particular case was actually more difficult to prove because several q-binomial and binomial simplifications were needed. In particular, it is natural to ask if the formula in (8) for $Z_N|_{\alpha=\beta=1}$ can be recovered from the general expression in Theorem 1.1, and the (affirmative) answer is essentially given in [13] (see also Subsection 4.2 below for a very similar simplification).

4 Moments of Al-Salam-Chihara polynomials

The link between the PASEP and the Al-Salam-Chihara orthogonal polynomials $Q_n(x;a,b\,|\,q)$ was described in [18]. These polynomials, denoted by $Q_n(x)$ when we do not need to precise the other parameters, are defined by the recurrence [17]:

$$2xQ_n(x) = Q_{n+1}(x) + (a+b)q^n Q_n(x) + (1-q^n)(1-abq^{n-1})Q_{n-1}(x) \quad (26)$$

together with $Q_{-1}(x) = 0$ and $Q_0(x) = 1$. They were introduced as the most general orthogonal sequence that is a convolution of two orthogonal sequences [1]. In terms of the Askey-Wilson polynomials $p_n(x;a,b,c,d\,|\,q)$, Al-Salam-Chihara polynomials are an important particular case since $Q_n(x;a,b\,|\,q) = p_n(x;a,b,0,0\,|\,q)$ [17].

4.1 Closed formulas for the moments

It can be checked that the specialization $(1-q)^N Z_N|_{y=1}$ is the Nth moment of the sequence $\{Q_n(\frac{x}{2}-1;\tilde{\alpha},\tilde{\beta}\,|\,q)\}_{n\in\mathbb{N}}$, where $\tilde{\alpha}=(1-q)\frac{1}{\alpha}-1$ and $\tilde{\beta}=(1-q)\frac{1}{\beta}-1$ as before. There is a simple relation between the moments of an orthogonal sequence and the ones of a rescaled sequence, so that assuming $a=\tilde{\alpha}$ and $b=\tilde{\beta}$ the Nth moment μ_N of the Al-Salam-Chihara polynomials can be obtained via the relation:

$$\mu_N=\sum_{k=0}^{N}\binom{N}{k}(-1)^{N-k}2^{-k}(1-q)^k Z_k|_{y=1}. \tag{27}$$

Actually, the methods of Section 2 also give a direct proof of the following.

Theorem 4.1 *The Nth moment of the Al-Salam-Chihara polynomials is:*

$$\mu_N=\frac{1}{2^N}\sum_{\substack{0\le n\le N\\ n\equiv N \text{ mod } 2}}\left(\sum_{j=0}^{\frac{N-n}{2}}(-1)^j q^{\binom{j+1}{2}}\begin{bmatrix}n+j\\ j\end{bmatrix}_q\left(\binom{N}{\frac{N-n}{2}-j}-\binom{N}{\frac{N-n}{2}-j-1}\right)\right)\left(\sum_{k=0}^{n}\begin{bmatrix}n\\ k\end{bmatrix}_q a^k b^{n-k}\right). \tag{28}$$

Proof: The general idea is to adapt the proof of Theorem 1.1 in Section 2. Let $\mathfrak{P}'_N\subset\mathfrak{P}_N$ be the subset of paths which contain no step $\to$ with weight $1+y$. The sum of weights of elements in $\mathfrak{P}'_N$ specialized at $y=1$, gives the Nth moment of the sequence $\{Q_n(\frac{x}{2})\}_{n\ge 0}$. This can be seen by comparing the weights in the Motzkin paths and the recurrence (26). But the Nth moment of this sequence is also $2^N\mu_N$.

From the definition of the bijection Φ in Section 2, we see that $\Phi(H_1,H_2)$ has no step $\to$ with weight $1+y$ if and only if H_1 has the same property. So from Proposition 2.3 the bijection Φ^{-1} gives a weight-preserving bijection between $\mathfrak{P}'_N$ and the disjoint union of $\mathfrak{R}'_{N,n}\times\mathfrak{B}_n$ over $n\in\{0,\dots,N\}$, where $\mathfrak{R}'_{N,n}\subset\mathfrak{R}_{N,n}$ is the subset of paths which contain no horizontal step with weight $1+y$. Note that $\mathfrak{R}'_{N,n}$ is empty when n and N do not have the same parity, because now n has to be the number of steps $\to$ in a Motzkin path of length N. In particular we can restrict the sum over n to the case $n\equiv N$ mod 2.

At this point it remains only to adapt the proof of Proposition 2.1 to compute the sum of weights of elements in $\mathfrak{R}'_{N,n}$, and obtain the sum over j in (28). As in the previous case we can adapt the methods from [8]. $\square$

We have to mention that there are analytical methods to obtain the moments μ_N of these polynomials. A nice formula for the Askey-Wilson moments was given by D. Stanton [19], as a consequence of joint results with M. Ismail [12, equation (1.16)]. As a particular case they have the Al-Salam-Chihara moments:

$$\mu_N=\frac{1}{2^N}\sum_{k=0}^{N}(ab;q)_k q^k\sum_{j=0}^{k}\frac{q^{-j^2}a^{-2j}(q^j a+q^{-j}a^{-1})^N}{(q,a^{-2}q^{-2j+1};q)_j(q,a^2q^{1+2j};q)_{k-j}}, \tag{29}$$

where we use the q-Pochhammer symbol. The latter formula has no apparent symmetry in a and b and has denominators, but D. Stanton gave evidence [19] that (29) can be simplified down to (28) using binomial, q-binomial, and q-Vandermonde summation theorems. Moreover (29) is equivalent to a formula for rescaled polynomials given in [15] (Section 4, Theorem 1 and equation (29)).

4.2 Some particular cases of Al-Salam-Chihara moments

When $a = b = 0$ in (28) we immediately recover the known result for the continuous q-Hermite moments. This is 0 if N is odd, and the Touchard-Riordan formula if N is even. Other interesting cases are the q-secant numbers $E_{2n}(q)$ and q-tangent numbers $E_{2n+1}(q)$, defined in [11] by continued fraction expansions of the generating functions:

$$\sum_{n\geq 0} E_{2n}(q)t^n = \cfrac{1}{1-\cfrac{[1]_q^2 t}{1-\cfrac{[2]_q^2 t}{1-\cfrac{[3]_q^2 t}{\ddots}}}} \quad \text{and} \quad \sum_{n\geq 0} E_{2n+1}(q)t^n = \cfrac{1}{1-\cfrac{[1]_q[2]_q t}{1-\cfrac{[2]_q[3]_q t}{1-\cfrac{[3]_q[4]_q t}{\ddots}}}}. \tag{30}$$

The exponential generating function of the numbers $E_n(1)$ is the function $\tan(x)+\sec(x)$. They have the following combinatorial interpretation [11, 14]: $E_n(q)$ is the generating function counting the occurrences of the generalized pattern 31-2 in alternating permutations of size n. We say that $\sigma \in \mathfrak{S}_n$ is alternating when $\sigma(1) > \sigma(2) < \sigma(3) > \dots$. From (30) these numbers are particular case of Al-Salam-Chihara moments:

$$E_{2n}(q) = (\tfrac{2}{1-q})^{2n} \mu_{2n}|_{a=-b=i\sqrt{q}}, \qquad \text{and} \quad E_{2n+1}(q) = (\tfrac{2}{1-q})^{2n} \mu_{2n}|_{a=-b=iq} \tag{31}$$

(where $i^2 = -1$). From (28) and q-binomial identities it is possible to obtain the closed formulas for $E_{2n}(q)$ and $E_{2n+1}(q)$ that were given in [14], in a similar manner that (4) can be simplified into (8) when $\alpha = \beta = 1$. Indeed, from (28) we can rewrite:

$$2^{2n}\mu_{2n} = \sum_{m=0}^{n} \left(\tbinom{2n}{n-m} - \tbinom{2n}{n-m-1}\right) \sum_{j,k\geq 0} (-1)^j q^{\binom{j+1}{2}} \left[\begin{smallmatrix}2m-j\\ j\end{smallmatrix}\right]_q \left[\begin{smallmatrix}2m-2j\\ k\end{smallmatrix}\right]_q \left(\tfrac{b}{a}\right)^k a^{2m-2j}. \tag{32}$$

This latter sum over j and k is also

$$\sum_{j,k\geq 0} (-1)^j q^{\binom{j+1}{2}} \left[\begin{smallmatrix}2m-j\\ j+k\end{smallmatrix}\right]_q \left[\begin{smallmatrix}j+k\\ j\end{smallmatrix}\right]_q \left(\tfrac{b}{a}\right)^k a^{2m-2j} = \sum_{\ell\geq j\geq 0} (-1)^j q^{\binom{j+1}{2}} \left[\begin{smallmatrix}2m-j\\ \ell\end{smallmatrix}\right]_q \left[\begin{smallmatrix}\ell\\ j\end{smallmatrix}\right]_q \left(\tfrac{b}{a}\right)^{\ell-j} a^{2m-2j}. \tag{33}$$

The sum over j can be simplified in the case $a = -b = i\sqrt{q}$, or $a = -b = iq$, using the q-binomial identities already used in [13] (see Lemma 2):

$$\sum_{j\geq 0} (-1)^j q^{\binom{j}{2}} \begin{bmatrix}2m-j\\ \ell\end{bmatrix}_q \begin{bmatrix}\ell\\ j\end{bmatrix}_q = q^{\ell(2m-\ell)}, \tag{34}$$

and

$$\sum_{j\geq 0} (-1)^j q^{\binom{j-1}{2}} \begin{bmatrix}2m-j\\ \ell\end{bmatrix}_q \begin{bmatrix}\ell\\ j\end{bmatrix}_q = \frac{q^{(\ell+1)(2m-\ell)} - q^{\ell(2m-\ell)} + q^{\ell(2m-\ell+1)} - q^{(\ell+1)(2m-\ell+1)}}{q^{2m-1}(1-q)}. \tag{35}$$

Omitting details, this gives a new proof of the Touchard-Riordan-like formulas for q-secant and q-tangent numbers [14]:

$$E_{2n}(q) = \frac{1}{(1-q)^{2n}} \sum_{m=0}^{n} \left(\tbinom{2n}{n-m} - \tbinom{2n}{n-m-1} \right) \sum_{\ell=0}^{2m} (-1)^{\ell+m} q^{\ell(2m-\ell)+m}, \tag{36}$$

and

$$E_{2n+1}(q) = \frac{1}{(1-q)^{2n+1}} \sum_{m=0}^{n} \left(\tbinom{2n+1}{n-m} - \tbinom{2n+1}{n-m-1} \right) \sum_{\ell=0}^{2m+1} (-1)^{\ell+m} q^{\ell(2m+2-\ell)}. \tag{37}$$

Acknowledgements

I thank my advisor Sylvie Corteel for her help, advice, support and kindness. I thank Einar Steingrímsson, Jiang Zeng and Lauren Williams for their help.

References

[1] W. A. Al-Salam and T. S. Chihara, Convolutions of orthonormal polynomials, SIAM J. Math. Anal. 7 (1976), 16–28.

[2] R. A. Blythe and M. R. Evans, Nonequilibrium steady states of matrix product form: A solver's guide, J. Phys. A: Math. gen. 40, 333–441.

[3] R. A. Blythe, M. R. Evans, F. Colaiori and F. H. L. Essler, Exact solution of a partially asymmetric exclusion model using a deformed oscillator algebra, J. Phys. A: Math. Gen. 33 (2000), 2313–2332.

[4] R. Brak, S. Corteel, J. Essam, R. Parviainen and A. Rechnitzer, A combinatorial derivation of the PASEP stationary state, Electron. J. Combin. 13(1) (2006), R108.

[5] S. Corteel, Crossings and alignments of permutations, Adv. in Appl. Math. 38(2) (2007), 149–163.

[6] S. Corteel, P. Nadeau, Bijections for permutation tableaux, European. J. Combin. 30(1) (2009), 295–310.

[7] S. Corteel and L. K. Williams, Tableaux combinatorics for the asymmetric exclusion process, Adv. in Appl. Math. 39(3) (2007), 293–310.

[8] S. Corteel, M. Josuat-Vergès, T. Prellberg and M. Rubey, Matrix Ansatz, lattice paths and rook placements, Proc. FPSAC 2009.

[9] B. Derrida, M. Evans, V. Hakim and V. Pasquier, Exact solution of a 1D asymmetric exclusion model using a matrix formulation, J. Phys. A: Math. Gen. 26 (1993), 1493–1517.

[10] P. Flajolet, Combinatorial aspects of continued fractions, Discrete Math. 41 (1982), 145–153.

[11] G.-N. Han, A. Randrianarivony, J. Zeng, Un autre q-analogue des nombres d'Euler, Séminaire Lotharingien de Combinatoire, B42e (1999).

[12] M. Ismail and D. Stanton, q-Taylor theorems, polynomial expansions, and interpolation of entire functions, J. Approx. Th. 123 (2003), 125–146.

[13] M. Josuat-Vergès, Rook placements in Young diagrams and permutation enumeration, preprint 2008, arXiv:0811.0524v2 [math.CO].

[14] M. Josuat-Vergès, A q-enumeration of alternating permutations, to appear in European J. Combin.

[15] A. Kasraoui, D. Stanton and J. Zeng, The Combinatorics of Al-Salam-Chihara q-Laguerre polynomials, preprint 2008, arXiv:0810.3232v1 [math.CO].

[16] D. Kim, On combinatorics of Al-Salam Carlitz polynomials, European. J. Combin. 18(3) (1997), 295–302.

[17] R. Koekoek and R. F. Swarttouw, The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue, Delft University of Technology, Report no. 98–17 (1998).

[18] T. Sasamoto, One-dimensional partially asymmetric simple exclusion process with open boundaries: orthogonal polynomials approach, J. Phys. A: Math. Gen. 32 (1999), 7109–7131.

[19] D. Stanton, personal communication.

[20] X. G. Viennot, Une théorie combinatoire des polynômes orthogonaux, Notes de cours, UQAM, Montréal 1988.

[21] A. Varvak, Rook numbers and the normal ordering problem, J. Combin. Theory Ser. A 112(2) (2005), 292–307.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 222–233

Generalized Energy Statistics and Kostka–Macdonald Polynomials

Anatol N. Kirillov[1] and Reiho Sakamoto[2†]

[1] *Research Institute for Mathematical Sciences, Kyoto University, Sakyo-ku, Kyoto, 606-8502, Japan,*
e-mail: kirillov@kurims.kyoto-u.ac.jp
[2] *Department of Physics, Tokyo University of Science, Kagurazaka, Shinjuku-ku, Tokyo, 162-8601, Japan*
e-mail: reiho@rs.kagu.tus.ac.jp

Abstract. We give an interpretation of the $t = 1$ specialization of the modified Macdonald polynomial as a generating function of the energy statistics defined on the set of paths arising in the context of Box-Ball Systems (BBS-paths for short). We also introduce one parameter generalizations of the energy statistics on the set of BBS-paths which all, conjecturally, have the same distribution.

Résumé. Nous donnons une intérprétation de la spécialisation à $t = 1$ du polynôme de Macdonald modifié comme fonction génératrice des statistiques d'énergie définies sur l'ensemble des chemins qui apparaissent dans la théorie des Systèmes BBS (BBS-chemins). Nous présentons également des généralisations à un paramètre de la statistique d'énergie sur les chemins BBS qui toutes, conjecturalement, ont la même distribution.

Keywords: modified Macdonald polynomials, box-ball systems

1 Introduction

The purpose of the present paper is two-fold. First of all we would like to draw attention to a rich combinatorics hidden behind the dynamics of Box-Ball Systems, and secondly, to connect the former with the theory of modified Macdonald polynomials. More specifically, our final goal is to give an interpretation of the Kostka–Macdonald polynomials $K_{\lambda,\mu}(q,t)$ as a *refined partition function* of a certain box-ball systems depending on initial data λ and μ.

Box-Ball Systems (BBS for short) were invented by Takahashi–Satsuma [29, 28] as a wide class of discrete integrable soliton systems. In the simplest case, BBS are described by simple combinatorial procedures using boxes and balls. One can see the simplest but still very interesting examples of the BBS by the free software available at [26]. Despite its simple outlook, it is known that the BBS have various remarkably deep properties:

- Local time evolution rule of the BBS coincides with the isomorphism of the crystal bases [7, 2]. Thus the BBS possesses quantum integrability.

† Supported by Grant-in-Aid for Scientific Research (No.21740114), JSPS.

- BBS are ultradiscrete (or tropical) limit of the usual soliton systems [30, 20]. Thus the BBS possesses classical integrability at the same time.
- Inverse scattering formalism of the BBS [19] coincides with the rigged configuration bijection originating in completeness problem of the Bethe states [14, 16], see also [25].

Let us say a few words about the main results of this note.

- We will identify the space of states of a BBS with the corresponding weight subspace in the tensor product of fundamental (or rectangular) representations of the Lie algebra $\mathfrak{gl}(n)$.
- In the case of statistics *tau*, our main result can be formulated as a computation of the corresponding partition function for the BBS in terms of the values of the Kostka–Macdonald polynomials at $t = 1$.
- In the case of the statistics *energy*, our result can be formulated as an interpretation of the corresponding partition function for the BBS as the q-weight multiplicity of a certain irreducible representation of the Lie algebra $\mathfrak{gl}(n)$ in the tensor product of the fundamental representations. We *expect* that the same statement is valid for the BBS corresponding to the tensor product of rectangular representations.

  Let us remind that a *q-analogue of the multiplicity* of a highest weight λ in the tensor product $\bigotimes_{a=1}^{L} V_{s_a \omega_{r_a}}$ of the highest weight $s_a\ \omega_{r_a}$, $a = 1, \ldots, L$, irreducible representations $V_{s_a\omega_{r_a}}$ of the Lie algebra $\mathfrak{gl}(n)$ is defined as

  $$q\text{-Mult}\,[V_\lambda\ :\ \bigotimes_{a=1}^{L} V_{s_a\omega_{r_a}}] = \sum_{\eta} K_{\eta,R}\, K_{\eta,\lambda}(q),$$

  where $K_{\eta,R}$ stands for the parabolic Kostka number corresponding to the sequence of rectangles $R := \{(s_a^{r_a})\}_{a=1,\ldots,L}$, see e.g. [15], [18].

A combinatorial description of the modified Macdonald polynomials has been obtained by Haglund–Haiman–Loehr [5]. In Section 5 we give an interpretation of two Haglund's statistics in the context of the box-ball systems, i.e., in terms of the BBS-paths. Namely, we identify the set of BBS paths of weight α with the set $\mathcal{P}(\alpha)$ which is the weight α component in the tensor product of crystals corresponding to vector representations. We have observed that from the proof given in [5] one can prove the following identity

$$\sum_{p\in\mathcal{P}(\alpha)} q^{\mathrm{inv}_\mu(p)} t^{\mathrm{maj}_\mu(p)} = \sum_{\eta\vdash|\mu|} K_{\eta,\alpha}\tilde{K}_{\eta,\mu}(q,t), \tag{1}$$

see Proposition 6.2 and Corollary 6.3. One of the main problems we are interested in is to generalize the identity Eq.(1) on more wider set of the BBS-paths.

Our result about connections of the energy partition functions for BBS and q-weight multiplicities suggests a deep hidden connections between partition functions for the BBS and characters of the Demazure modules, solutions to the q-difference Toda equations, cf.[3], ...

As an interesting open problem we want to give raise a question about an interpretation of the sums $\sum_\eta K_{\eta,R}\, K_{\eta,\lambda}(q,t)$, where $K_{\eta,\lambda}(q,t)$ denotes the Kostka–Macdonald polynomials [21], as *refined partition functions* for the BBS corresponding to the tensor product of rectangular representations $R =$

$\{(s_a^{r_a})\}_{1\le a\le n}$. In other words, one can ask: what is a meaning of the second statistics (see [5]) in the Kashiwara theory [11] of crystal bases (of type A) ?

This paper is abbreviated and updated version of our paper [17]. The main novelty of the present paper is the definition of a one parameter family of statistics on the set of BBS-paths which generalizes those introduced in [17], see Conjecture 7.2. It conjecturally gives a new family of MacMahonian statistics on the set of transportation matrices, see [15].

Organization of the present paper is as follows. In Section 2 we remind algorithms of the combinatorial R-matrix and the energy functions. In Section 3, we introduce the energy statistics and the set of the BBS. In Section 4 we remind definition of box-ball systems and state some of their simplest properties. In Section 5 we remind definition of the Haglund's statistics and give their interpretation in terms of the BBS-paths. Sections 6 and 7 contain our main results and conjectures. In particular it is not difficult to see that Haglund's statistics maj_μ and inv_μ do not compatible with the Kostka–Macdonald polynomials for general partitions λ and μ. In Section 6 we state a conjecture which describes the all pairs of partitions (λ,μ) for those the restriction of the Haglund–Haiman–Loehr formula on the set of highest weight paths of shape μ coincide with the Kostka–Macdonald polynomial $\tilde{K}_{\lambda,\mu}(q,t)$.

2 Combinatorial R and energy function

Let $B^{r,s}$ be the Kirillov–Reshetikhin crystals of type $A_n^{(1)}$ (see [11, 12, 10], see also section 2 of [17]). Here $r\in\{1,2,\cdots,n\}$ and $s\in\mathbb{Z}_{>0}$. As the set, $B^{r,s}$ is consisting of all semistandard tableaux of height r and width s. In this section, we recall an explicit description of the combinatorial R-matrix (combinatorial R for short) and energy function on $B^{r,s}\otimes B^{r',s'}$. To begin with we define few terminologies about Young tableaux. Denote rows of a Young tableaux Y by $y_1,y_2,\ldots y_r$ from top to bottom. Then row word $row(Y)$ is defined by concatenating rows as $row(Y)=y_r y_{r-1}\ldots y_1$. Let $x=(x_1,x_2,\ldots)$ and $y=(y_1,y_2,\ldots)$ be two partitions. We define concatenation of x and y by the partition $(x_1+y_1,x_2+y_2,\ldots)$.

Proposition 2.1 ([27]) *$b\otimes b'\in B^{r,s}\otimes B^{r',s'}$ is mapped to $\tilde{b}'\otimes\tilde{b}\in B^{r',s'}\otimes B^{r,s}$ under the combinatorial R, i.e.,*

$$b\otimes b'\overset{R}{\simeq}\tilde{b}'\otimes\tilde{b},\tag{2}$$

if and only if

$$(b'\leftarrow row(b))=(\tilde{b}\leftarrow row(\tilde{b}')).\tag{3}$$

Moreover, the energy function $H(b\otimes b')$ is given by the number of nodes of $(b'\leftarrow row(b))$ outside the concatenation of partitions (s^r) and $(s'^{r'})$.

For special cases of $B^{1,s}\otimes B^{1,s'}$, the function H is called unwinding number in [22]. Explicit values for the case $b\otimes b'\in B^{1,1}\otimes B^{1,1}$ are given by $H(b\otimes b')=\chi(b<b')$ where $\chi(\mathrm{True})=1$ and $\chi(\mathrm{False})=0$.

In order to describe the algorithm for finding $\tilde{b}$ and $\tilde{b}'$ from the data $(b'\leftarrow row(b))$, we introduce a terminology. Let Y be a tableau, and Y' be a subset of Y such that Y' is also a tableau. Consider the set theoretic subtraction $\theta=Y\setminus Y'$. If the number of nodes contained in θ is r and if the number of nodes of θ contained in each row is always 0 or 1, then θ is called vertical r-strip.

Given a tableau $Y=(b'\leftarrow row(b))$, let Y' be the upper left part of Y whose shape is (s^r). We assign numbers from 1 to $r's'$ for each node contained in $\theta=Y\setminus Y'$ by the following procedure. Let θ_1 be the vertical r'-strip of θ as upper as possible. For each node in θ_1, we assign numbers 1 through r' from the bottom to top. Next we consider $\theta \setminus \theta_1$, and find the vertical r' strip θ_2 by the same way. Continue this procedure until all nodes of θ are assigned numbers up to $r's'$. Then we apply inverse bumping procedure according to the labeling of nodes in θ. Denote by u_1 the integer which is ejected when we apply inverse bumping procedure starting from the node with label 1. Denote by Y_1 the tableau such that $(Y_1 \leftarrow u_1) = Y$. Next we apply inverse bumping procedure starting from the node of Y_1 labeled by 2, and obtain the integer u_2 and tableau Y_2. We do this procedure until we obtain $u_{r's'}$ and $Y_{r's'}$. Finally, we have

$$\tilde{b}' = (\emptyset \leftarrow u_{r's'} u_{r's'-1} \cdots u_1), \qquad \tilde{b} = Y_{r's'}. \tag{4}$$

3 Energy statistics and its generalizations on the set of paths

For a path $b_1 \otimes b_2 \otimes \cdots \otimes b_L \in B^{r_1,s_1} \otimes B^{r_2,s_2} \otimes \cdots \otimes B^{r_L,s_L}$, let us define elements $b_j^{(i)} \in B^{r_j,s_j}$ for $i < j$ by the following isomorphisms of the combinatorial R;

$$\begin{aligned}
& b_1 \otimes b_2 \otimes \cdots \otimes b_{i-1} \otimes b_i \otimes \cdots \otimes b_{j-1} \otimes b_j \otimes \cdots \\
\simeq \quad & b_1 \otimes b_2 \otimes \cdots \otimes b_{i-1} \otimes b_i \otimes \cdots \otimes b_j^{(j-1)} \otimes b'_{j-1} \otimes \cdots \\
\simeq \quad & \cdots \\
\simeq \quad & b_1 \otimes b_2 \otimes \cdots \otimes b_{i-1} \otimes b_j^{(i)} \otimes \cdots \otimes b'_{j-2} \otimes b'_{j-1} \otimes \cdots ,
\end{aligned} \tag{5}$$

where we have written $b_k \otimes b_j^{(k+1)} \simeq b_j^{(k)} \otimes b'_k$ assuming that $b_j^{(j)} = b_j$.

Define the statistics $\mathrm{maj}(p)$ by

$$\mathrm{maj}(p) = \sum_{i<j} H(b_i \otimes b_j^{(i+1)}). \tag{6}$$

For example, consider a path $a = a_1 \otimes a_2 \otimes \cdots \otimes a_L \in (B^{1,1})^{\otimes L}$. In this case, we have $a_j^{(i)} = a_i$, since the combinatorial R act on $B^{1,1} \otimes B^{1,1}$ as identity. Therefore, we have

$$\mathrm{maj}(a) = \sum_{i=1}^{L-1} (L-i)\chi(a_i < a_{i+1}). \tag{7}$$

Define another statistics *tau* as follows.

Definition 3.1 *For the path $p \in B^{r_1,s_1} \otimes B^{r_2,s_2} \otimes \cdots \otimes B^{r_L,s_L}$, define $\tau^{r,s}$ by*

$$\tau^{r,s}(p) = \mathrm{maj}(u_s^{(r)} \otimes p), \tag{8}$$

where $u_s^{(r)}$ is the highest element of $B^{r,s}$.

Here the highest element $u_s^{(r)} \in B^{r,s}$ is the tableau whose i-th row is occupied by integers i. For example, $u_4^{(3)} = \begin{array}{|c|c|c|c|}\hline 1&1&1&1\\\hline 2&2&2&2\\\hline 3&3&3&3\\\hline\end{array}$. In particular, the statistics $\tau^{r,1}$ on $B^{1,1}$ type paths $a \in (B^{1,1})^{\otimes L}$ has the following form;

$$\tau^{r,1}(a) = L \cdot \chi(r < a_1) + \sum_{i=1}^{L-1} (L-i)\chi(a_i < a_{i+1}), \tag{9}$$

where a_1 denotes the first letter of the path a. Note that $\tau^{1,1}$ is a special case of the tau functions for the box-ball systems [20, 24] which originates as an ultradiscrete limit of the tau functions for the KP hierarchy [9].

Definition 3.2 *For composition $\mu = (\mu_1, \mu_2, \cdots, \mu_n)$, write $\mu_{[i]} = \sum_{j=1}^{i} \mu_j$ with convention $\mu_{[0]} = 0$. Then we define a generalization of $\tau^{r,1}$ by*

$$\tau_{\mu}^{r,1}(a) = \sum_{i=1}^{n} \tau^{r,1}(a_{[i]}), \tag{10}$$

where

$$a_{[i]} = a_{\mu_{[i-1]}+1} \otimes a_{\mu_{[i-1]}+2} \otimes \cdots \otimes a_{\mu_{[i]}} \in (B^{1,1})^{\otimes \mu_i}. \tag{11}$$

Note that we have $a = a_{[1]} \otimes a_{[2]} \otimes \cdots \otimes a_{[n]}$, i.e., the path a is partitioned according to μ.

4 Box-ball system

In this section, we summarize basic facts about the box-ball system in order to explain physical origin of $\tau^{1,1}$. For our purpose, it is convenient to express the isomorphism of the combinatorial R: $a \otimes b \simeq b' \otimes a'$ by the following vertex diagram:

$$\begin{array}{ccc} & b & \\ & \vert & \\ a & \text{---}\!\!+\!\!\text{---} & a' \\ & \vert & \\ & b' & \end{array}.$$

Successive applications of the combinatorial R is depicted by concatenating these vertices.

Following [7, 2], we define time evolution of the box-ball system $T_l^{(a)}$. Let $u_{l,0}^{(a)} = u_l^{(a)} \in B^{a,l}$ be the highest element and $b_i \in B^{r_i,s_i}$. Define $u_{l,j}^{(a)}$ and $b_i' \in B^{r_i,s_i}$ by the following diagram.

$$\begin{array}{ccccccccccc} & b_1 & & b_2 & & & & b_L & \\ & \vert & & \vert & & & & \vert & \\ u_{l,0}^{(a)} & \text{---}\!\!+\!\!\text{---} & u_{l,1}^{(a)} & \text{---}\!\!+\!\!\text{---} & u_{l,2}^{(a)} & \cdots\cdots\cdots & u_{l,L-1}^{(a)} & \text{---}\!\!+\!\!\text{---} & u_{l,L}^{(a)} \\ & \vert & & \vert & & & & \vert & \\ & b_1' & & b_2' & & & & b_L' & \end{array} \tag{12}$$

$u_{l,j}^{(a)}$ are usually called *carrier* and we set $u_{l,0}^{(a)} := u_l^{(a)}$. Then we define operator $T_l^{(a)}$ by

$$T_l^{(a)}(b) = b' = b_1' \otimes b_2' \otimes \cdots \otimes b_L'. \tag{13}$$

Recently [25], operators $T_l^{(a)}$ have used to derive crystal theoretical meaning of the rigged configuration bijection.

It is known ([19] Theorem 2.7) that there exists some $l \in \mathbb{Z}_{>0}$ such that

$$T_l^{(a)} = T_{l+1}^{(a)} = T_{l+2}^{(a)} = \cdots (=: T_\infty^{(a)}). \tag{14}$$

If the corresponding path is $b \in (B^{1,1})^{\otimes L}$, we have the following combinatorial description of the box-ball system [29, 28]. We regard $\boxed{1} \in B^{1,1}$ as an empty box of capacity 1, and $\boxed{i} \in B^{1,1}$ as a ball of label (or internal degree of freedom) i contained in the box. Then we have:

Proposition 4.1 ([7]) *For a path $b \in (B^{1,1})^{\otimes L}$ of type $A^{(1)}_n$, $T^{(1)}_\infty(b)$ is given by the following procedure.*

1. *Move every ball only once.*
2. *Move the leftmost ball with label $n+1$ to the nearest right empty box.*
3. *Move the leftmost ball with label $n+1$ among the rest to its nearest right empty box.*
4. *Repeat this procedure until all of the balls with label $n+1$ are moved.*
5. *Do the same procedure 2–4 for the balls with label n.*
6. *Repeat this procedure successively until all of the balls with label 2 are moved.*

There are extensions of this box and ball algorithm corresponding to generalizations of the box-ball systems with respect to each affine Lie algebra, see e.g., [8]. Using this box and ball interpretation, our statistics $\tau^{1,1}(b)$ admits the following interpretation.

Theorem 4.2 ([20] Theorem 7.4) *For a path $b \in (B^{1,1})^{\otimes L}$ of type $A^{(1)}_n$, $\tau^{1,1}(b)$ coincides with number of all balls $2, \cdots, n+1$ contained in paths b, $T^{(1)}_\infty(b)$, $\cdots$, $(T^{(1)}_\infty)^{L-1}(b)$.*

Example 4.3 Consider the path $p = a \otimes b$ where $a = 4311211111$, $b = 4321111111$. Note that we omit all frames of tableaux of $B^{1,1}$ and symbols for tensor product. We compute $\tau_{(10,10)}(p)$ by using Theorem 4.2. According to Proposition 4.1, the time evolutions of the paths a and b are as follows:

4	3	1	1	2	1	1	1	1	1
1	1	4	3	1	2	1	1	1	1
1	1	1	1	4	1	3	2	1	1
1	1	1	1	1	4	1	1	3	2
1	1	1	1	1	1	4	1	1	1
1	1	1	1	1	1	1	4	1	1
1	1	1	1	1	1	1	1	4	1
1	1	1	1	1	1	1	1	1	4

4	3	2	1	1	1	1	1	1	1
1	1	1	4	3	2	1	1	1	1
1	1	1	1	1	1	4	3	2	1
1	1	1	1	1	1	1	1	1	4
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	1	1

Here the left and right tables correspond to a and b, respectively. Rows of left (resp. right) table represent a, $T^{(1)}_\infty(a)$, $\cdots$, $(T^{(1)}_\infty)^L(a)$ (resp., those for b) from top to bottom. Counting letters 2, 3 and 4 in each table, we have $\tau^{1,1}(a) = 16$, $\tau^{1,1}(b) = 10$ and we get $\tau^{1,1}_{(10,10)}(p) = 16 + 10 = 26$, which coincides with the computation by Eq.(9). Meanings of the above two dynamics corresponding to paths a and b are summarized as follows:

(a) Dynamics of the path a. In the first two rows, there are two solitons (length two soliton 43 and length one soliton 2), and in the lower rows, there are also two solitons (length one soliton 4 and length two soliton 32). This is scattering of two solitons. After the scattering, soliton 4 propagates at velocity one and soliton 32 propagates at velocity two without scattering.

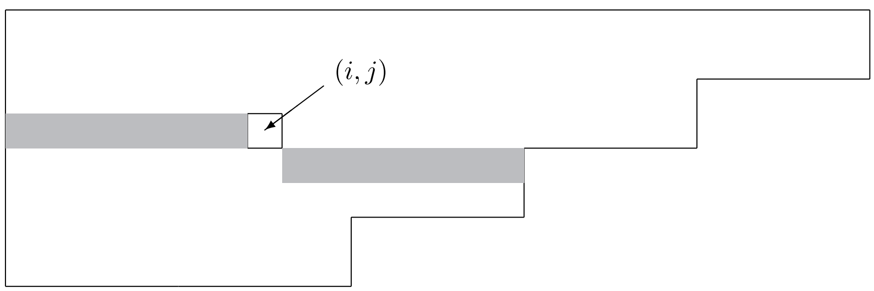
(b) Dynamics of the path b. This shows free propagation of one soliton of length three 432 at velocity three.

5 Haglund's statistics

Tableaux language description For a given path $a = a_1 \otimes a_2 \otimes \cdots \otimes a_L \in (B^{1,1})^{\otimes L}$, associate tabloid t of shape μ whose reading word coincides with a. For example, to path $p = abcdefgh$ and the composition $\mu = (3, 2, 3)$ one associates the tabloid

$$\begin{array}{|c|c|c|} \hline c & b & a \\ \hline e & d & \\ \hline h & g & f \\ \hline \end{array} \quad . \tag{15}$$

Denote the cell at the i-th row, j-th column (we denote the coordinate by (i, j)) of the tabloid t by t_{ij}. Attacking region of the cell at (i, j) is all cells (i, k) with $k < j$ or $(i + 1, k)$ with $k > j$. In the following diagram, gray zonal regions are the attacking regions of the cell (i, j).



Follow [5], define $|\mathrm{Inv}_{ij}|$ by

$$|\mathrm{Inv}_{ij}| = \#\{(k,l) \in \text{ attacking region for } (i,j) \,|\, t_{kl} > t_{ij}\}. \tag{16}$$

Then we define

$$|\mathrm{Inv}_\mu(a)| = \sum_{(i,j)\in\mu} |\mathrm{Inv}_{ij}|. \tag{17}$$

If we have $t_{(i-1)j} < t_{ij}$, then the cell (i, j) is called by *descent*. Then define

$$\mathrm{Des}_\mu(a) = \sum_{\text{all descent } (i,j)} (\mu_i - j). \tag{18}$$

Note that $(\mu_i - j)$ is the arm length of the cell (i, j).

Path language description Consider two paths $a^{(1)}, a^{(2)} \in (B^{1,1})^{\otimes\mu}$. We denote by $a^{(1)} \otimes a^{(2)} = a_1 \otimes a_2 \otimes \cdots \otimes a_{2\mu}$. Then we define

$$\mathrm{Inv}_{(\mu,\mu)}(a^{(1)}, a^{(2)}) = \sum_{k=1}^{\mu} \sum_{i=k+1}^{k+\mu-1} \chi(a_k < a_i). \tag{19}$$

For more general cases $a^{(1)} \in (B^{1,1})^{\otimes \mu_1}$ and $a^{(2)} \in (B^{1,1})^{\otimes \mu_2}$ satisfying $\mu_1 > \mu_2$, we define

$$\mathrm{Inv}_{(\mu_1,\mu_2)}(a^{(1)}, a^{(2)}) := \mathrm{Inv}_{(\mu_1,\mu_1)}(a^{(1)}, 1^{\otimes(\mu_1-\mu_2)} \otimes a^{(2)}). \tag{20}$$

Then the above definition of $|\mathrm{Inv}_\mu(a)|$ is equivalent to

$$|\mathrm{Inv}_\mu(a)| = \sum_{i=1}^{n-1} \mathrm{Inv}_{(\mu_i,\mu_{i+1})}. \tag{21}$$

Consider two paths $a^{(1)} \in (B^{1,1})^{\otimes \mu_1}$ and $a^{(2)} \in (B^{1,1})^{\otimes \mu_2}$ satisfying $\mu_1 \geq \mu_2$. Denote $a = a^{(1)} \otimes a^{(2)}$. Then define

$$\mathrm{Des}_{(\mu_1,\mu_2)}(a) = \sum_{k=\mu_1-\mu_2+1}^{\mu_1} (k - (\mu_1 - \mu_2) - 1)\chi(a_k < a_{k+\mu_2}). \tag{22}$$

For the tableau T of shape μ corresponding to the path a, we define

$$\mathrm{Des}_\mu(T) = \sum_{i=1}^{n} \mathrm{Des}_{(\mu_i,\mu_{i+1})}(a_{[i]} \otimes a_{[i+1]}). \tag{23}$$

Definition 5.1 ([4]) *For a path a, statistics maj_μ is defined by*

$$\mathrm{maj}_\mu(a) = \sum_{i=1}^{\mu_1} \mathrm{maj}(t_{1,i} \otimes t_{2,i} \otimes \cdots \otimes t_{\mu'_i,i}). \tag{24}$$

and $\mathrm{inv}_\mu(a)$ is defined by

$$\mathrm{inv}_\mu(a) = |\mathrm{Inv}_\mu(a)| - \mathrm{Des}_\mu(a). \tag{25}$$

If we associate to a given path $p \in \mathcal{P}(\lambda)$ with the shape μ tabloid T, we sometimes write $\mathrm{maj}_\mu(p) = \mathrm{maj}(T)$ and $\mathrm{inv}_\mu(p) = \mathrm{inv}(T)$.

6 Haglund–Haiman–Loehr formula

Let $\tilde{H}_\mu(x; q, t)$ be the (integral form) modified Macdonald polynomials where x stands for infinitely many variables $x_1, x_2, \cdots$. Here $\tilde{H}_\mu(x; q, t)$ is obtained by simple plethystic substitution (see, e.g., section 2 of [6]) from the original definition of the Macdonald polynomials [21]. Schur function expansion of $\tilde{H}_\mu(x; q, t)$ is given by

$$\tilde{H}_\mu(x; q, t) = \sum_\lambda \tilde{K}_{\lambda,\mu}(q, t) s_\lambda(x), \tag{26}$$

where $\tilde{K}_{\lambda,\mu}(q, t)$ stands for the following transformation of the Kostka–Macdonald polynomials:

$$\tilde{K}_{\lambda,\mu}(q, t) = t^{n(\mu)} K_{\lambda,\mu}(q, t^{-1}). \tag{27}$$

Here we have used notation $n(\mu) = \sum_i (i - 1)\mu_i$. Then the celebrated Haglund–Haiman–Loehr (HHL) formula is as follows.

Theorem 6.1 ([5]) *Let $\sigma : \mu \to \mathbb{Z}_{>0}$ be the filling of the Young diagram μ by positive integers $\mathbb{Z}_{>0}$, and define $x^\sigma = \prod_{u\in\mu} x_{\sigma(u)}$. Then the Macdonald polynomial $\tilde{H}_\mu(x;q,t)$ have the following explicit formula:*

$$\tilde{H}_\mu(x;q,t) = \sum_{\sigma:\mu\to\mathbb{Z}_{>0}} q^{\mathrm{inv}(\sigma)} t^{\mathrm{maj}(\sigma)} x^\sigma. \tag{28}$$

From the HHL formula, we can show the following formula.

Proposition 6.2 *For any partition μ and composition α of the same size, one has*

$$\sum_{p\in\mathcal{P}(\alpha)} q^{\mathrm{inv}_\mu(p)} t^{\mathrm{maj}_\mu(p)} = \sum_{\eta\vdash|\mu|} K_{\eta,\alpha} \tilde{K}_{\eta,\mu}(q,t), \tag{29}$$

where $\mathcal{P}(\alpha)$ stands for the set of type $B^{1,1}$ paths of weight $\alpha = (\alpha_1,\alpha_2,\ldots,\alpha_{n+1})$ and η runs over all partitions of size $|\mu|$.

Corollary 6.3 *The (modified) Macdonald polynomial $\tilde{H}_\mu(x;q,t)$ have the following expansion in terms of the monomial symmetric functions $m_\lambda(x)$:*

$$\tilde{H}_\mu(x;q,t) = \sum_{\lambda\vdash|\mu|} \left(\sum_{p\in\mathcal{P}(\lambda)} q^{\mathrm{inv}_\mu(p)} t^{\mathrm{maj}_\mu(p)} \right) m_\lambda(x), \tag{30}$$

where λ runs over all partitions of size $|\mu|$.

To find combinatorial interpretation of the Kostka–Macdonald polynomials $\tilde{K}_{\lambda,\mu}(q,t)$ remains significant open problem. Among many important partial results about this problem, we would like to mention the following theorem also due to Haglund–Haiman–Loehr:

Theorem 6.4 ([5] Proposition 9.2) *If $\mu_1 \leq 2$, we have*

$$\tilde{K}_{\lambda,\mu}(q,t) = \sum_{p\in\mathcal{P}_+(\lambda)} q^{\mathrm{inv}_\mu(p)} t^{\mathrm{maj}_\mu(p)}, \tag{31}$$

where $\mathcal{P}_+(\lambda)$ is the set of all highest weight elements of $\mathcal{P}(\lambda)$ according to the reading order explained in Eq.(15).

It is interesting to compare this formula with the formula obtained by S. Fishel [1], see also [14], [18].

Concerning validity of the formula Eq.(31), we state the following conjecture.

Conjecture 6.5 *Explicit formula for the Kostka–Macdonald polynomials*

$$\tilde{K}_{\lambda,\mu}(q,t) = \sum_{p\in\mathcal{P}_+(\lambda)} q^{\mathrm{inv}_\mu(p)} t^{\mathrm{maj}_\mu(p)}. \tag{32}$$

is valid if and only if at least one of the following two conditions is satisfied.

(i) $\mu_1 \leq 3$ *and* $\mu_2 \leq 2$.

(ii) λ *is a hook shape.*

7 Generating function of tau functions

In [17], we give an elementary proof for special case $t = 1$ of the formula Eq.(29) in the following form.

Theorem 7.1 *Let α be a composition and μ be a partition of the same size. Then,*

$$\sum_{p\in\mathcal{P}(\alpha)} q^{\mathrm{maj}_{\mu'}(p)} = \sum_{\eta\vdash|\mu|} K_{\eta,\alpha}\, K_{\eta,\mu}(q,1). \tag{33}$$

Conjecture 7.2 *Let α be a composition and μ be a partition of the same size. Then,*

$$q^{-\sum_{i>r}\alpha_i} \sum_{p\in\mathcal{P}(\alpha)} q^{\tau_\mu^{r,1}(p)} = \sum_{\eta\vdash|\mu|} K_{\eta,\alpha}\tilde{K}_{\eta,\mu}(q,1). \tag{34}$$

This conjecture contains Conjecture 5.8 of [17] and Theorem 7.1 above as special cases $r = 1$ and $r = \infty$, respectively. Also, extensions for paths of more general representations without partition μ are discussed in Section 5.3 of [17].

Example 7.3 Let us consider case $\alpha = (4, 1, 1)$ and $\mu = (4, 2)$. The following is a list of paths p and the corresponding value of tau function $\tau^{2,1}_{(4,2)}(p)$. For example, the top left corner $\boxed{111123 \quad 1}$ means $p = \boxed{1} \otimes \boxed{1} \otimes \boxed{1} \otimes \boxed{1} \otimes \boxed{2} \otimes \boxed{3}$ and $\tau^{2,1}_{(4,2)}(p) = 1$.

111123	1	111132	2	111213	2	111231	3	111312	2	111321	1
112113	3	112131	4	112311	3	113112	3	113121	2	113211	2
121113	4	121131	5	121311	4	123111	5	131112	4	131121	3
131211	4	132111	3	211113	1	211131	2	211311	1	213111	2
231111	3	311112	5	311121	4	311211	5	312111	6	321111	4

Summing up, LHS of Eq.(34) is

$$q^{-1} \sum_{p\in\mathcal{P}((4,1,1))} q^{\tau^{2,1}_{(4,2)}(p)} = q^5 + 4q^4 + 7q^3 + 7q^2 + 7q + 4$$

which coincides with the RHS of Eq.(34). Compare this with $\tau^{1,1}_{(4,2)}$ data for the same set of paths at Example 5.9 of [17].

References

[1] S. Fishel, Statistics for special q, t-Kostka polynomials, Proc. Amer. Math. Soc. **123** (1995) 2961–2969.

[2] K. Fukuda, M. Okado and Y. Yamada, Energy functions in box-ball systems, Int. J. Mod. Phys. **A15** (2000) 1379–1392, arXiv:math/9908116.

[3] A. Gerasimov, D. Lebedev and S. Oblezin, On q-deformed $gl(l + 1)$-Whittaker function I, Comm. Math. Phys. **294** (2010) 97–119, arXiv:0803.0145; II, Comm. Math. Phys. **294** (2010) 121–143, arXiv:0803.0970; III, arXiv:0805.3754.

[4] J. Haglund, A combinatorial model for the Macdonald polynomials, Proc. Nat. Acad. Sci. USA **101** (2004) 16127-16131.

[5] J. Haglund, M. Haiman and N. Loehr, A combinatorial formula for Macdonald polynomials, J. Amer. Math. Soc. **18** (2005) 735–761, arXiv:math/0409538.

[6] M. Haiman, Macdonald polynomials and geometry, *New perspectives in geometric combinatorics* (Billera, Björner, Greene, Simion and Stanley, eds.), MSRI publications, **38** Cambridge Univ. Press (1999) 207–254.

[7] G. Hatayama, K. Hikami, R. Inoue, A. Kuniba, T. Takagi and T. Tokihiro, The $A^{(1)}_M$ automata related to crystals of symmetric tensors, J. Math. Phys. **42** (2001) 274–308, arXiv:math/9912209.

[8] G. Hatayama, A. Kuniba and T. Takagi, Simple algorithm for factorized dynamics of $\mathfrak{g}_n$-automaton, J. Phys. A: Math. Gen. **34** (2001) 10697–10705, arXiv:nlin/0103022.

[9] M. Jimbo and T. Miwa, Solitons and infinite dimensional Lie algebras, Publ. RIMS. Kyoto Univ. **19** (1983) 943–1001.

[10] S.-J. Kang, M. Kashiwara, K. C. Misra, T. Miwa, T. Nakashima and A. Nakayashiki, Perfect crystals of quantum affine Lie algebras, Duke Math. J. **68** (1992) 499–607.

[11] M. Kashiwara, On crystal bases of the q-analogue of universal enveloping algebras, Duke Math. J. **63** (1991) 465–516.

[12] M. Kashiwara and T. Nakashima, Crystal graphs for representations of the q-analogue of classical Lie algebras, J. Algebra **165** (1994) 295–345.

[13] S. V. Kerov, A. N. Kirillov and N. Yu. Reshetikhin, Combinatorics, the Bethe ansatz and representations of the symmetric group, J. Soviet Math. **41** (1988) 916–924.

[14] A. N. Kirillov, Combinatorics of Young tableaux and rigged configurations (Russian) Proceedings of the St. Petersburg Math. Soc., **7** (1999), 23-115;
translation in Proceedings of the St. Petersburg Math. Soc. volume VII, Amer. Math. Soc. Transl. Ser.2, **203**, 17-98, AMS, Providence, RI, 2001

[15] A. N. Kirillov, New combinatorial formula for modified Hall–Littlewood polynomials, Contemp. Math. **254** (2000) 283–333, arXiv:math/9803006.

[16] A. N. Kirillov and N. Yu. Reshetikhin, The Bethe ansatz and the combinatorics of Young tableaux. J. Soviet Math. **41** (1988) 925–955.

[17] A. N. Kirillov and R. Sakamoto, Paths and Kostka–Macdonald polynomials, Moscow Math. J. **9** (2009) 823–854, arXiv:0811.1085.

[18] A. N. Kirillov and M. Shimozono, A generalization of the Kostka–Foulkas polynomials, J. Algebraic Combin. **15** (2002) 27–69, arXiv:math/9803062.

[19] A. Kuniba, M. Okado, R. Sakamoto, T. Takagi and Y. Yamada, Crystal interpretation of Kerov–Kirillov–Reshetikhin bijection, Nucl. Phys. **B740** (2006) 299–327, arXiv:math/0601630.

[20] A. Kuniba, R. Sakamoto and Y. Yamada, Tau functions in combinatorial Bethe ansatz, Nucl. Phys. **B786** (2007) 207–266, arXiv:math/0610505.

[21] I. G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd edition, Oxford Univ. Press, New York (1995) x+475 pp.

[22] A. Nakayashiki and Y. Yamada, Kostka polynomials and energy functions in solvable lattice models, Selecta Math. (N.S.) **3** (1997) 547–599, arXiv:q-alg/9512027.

[23] M. Okado, $X = M$ conjecture, MSJ Memoirs **17** (2007) 43–73.

[24] R. Sakamoto, Crystal interpretation of Kerov–Kirillov–Reshetikhin bijection II. Proof for $\mathfrak{sl}_n$ case, J. Algebraic Combin. **27** (2008) 55–98, arXiv:math/0601697.

[25] R. Sakamoto, Kirillov–Schilling–Shimozono bijection as energy functions of crystals, Int. Math. Res. Notices **2009** (2009) 579–614, arXiv:0711.4185.

[26] R. Sakamoto, "Periodic Box-Ball System" from The Wolfram Demonstrations Project, http://demonstrations.wolfram.com/PeriodicBoxBallSystem/

[27] M. Shimozono, Affine type A crystal structure on tensor products of rectangles, Demazure characters, and nilpotent varieties, J. Algebraic Combin. **15** (2002) 151–187, arXiv:math/9804039.

[28] D. Takahashi, On some soliton systems defined by using boxes and balls, *Proceedings of the International Symposium on Nonlinear Theory and Its Applications (NOLTA '93),* (1993) 555–558.

[29] D. Takahashi and J. Satsuma, A soliton cellular automaton, J. Phys. Soc. Japan, **59** (1990) 3514–3519.

[30] T. Tokihiro, D. Takahashi, J. Matsukidaira and J. Satsuma, From soliton equations to integrable cellular automata through a limiting procedure, Phys. Rev. Lett. **76** (1996) 3247–3250.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 234–244

Skew Littlewood–Richardson rules from Hopf algebras

Thomas Lam[1,†], Aaron Lauve[2], and Frank Sottile[2,‡]

[1]*Department of Mathematics*
University of Michigan
Ann Arbor, MI 48109
[2]*Department of Mathematics*
Texas A&M University
College Station, TX 77843

Abstract. We use Hopf algebras to prove a version of the Littlewood–Richardson rule for skew Schur functions, which implies a conjecture of Assaf and McNamara. We also establish skew Littlewood–Richardson rules for Schur P- and Q-functions and noncommutative ribbon Schur functions, as well as skew Pieri rules for k-Schur functions, dual k-Schur functions, and for the homology of the affine Grassmannian of the symplectic group.

Résumé. Nous utilisons des algèbres de Hopf pour prouver une version de la règle de Littlewood–Richardson pour les fonctions de Schur gauches, qui implique une conjecture d'Assaf et McNamara. Nous établissons également des règles de Littlewood–Richardson gauches pour les P- et Q-fonctions de Schur et les fonctions de Schur rubbans non commutatives, ainsi que des règles de Pieri gauches pour les k-fonctions de Schur, les k-fonctions de Schur duales, et pour l'homologie de la Grassmannienne affine du groupe symplectique.

Keywords: symmetric functions, Littlewood–Richardson rule, Pieri rule, Hopf algebras, antipode

Assaf and McNamara [AM] recently used combinatorics to give an elegant and surprising formula for the product of a skew Schur function and a complete homogeneous symmetric function. Their paper included a conjectural skew version of the Littlewood–Richardson rule, and also an appendix by one of us (Lam) with a simple algebraic proof of their formula. We show how these formulas and much more are special cases of a simple formula that holds for any pair of dual Hopf algebras. We first establish this Hopf-algebraic formula, and then apply it to obtain formulas in some well-known Hopf algebras in combinatorics.

1 A Hopf algebraic formula

We assume basic familiarity with Hopf algebras, as found in the opening chapters of the book [Mon93]. Let H, H^* be a pair of dual Hopf algebras over a field $\Bbbk$. This means that there is a nondegenerate pairing

† Supported by a Sloan Fellowship, and by NSF grants DMS-0901111 and DMS-0652641.
‡ Supported by the NSF grant DMS-0701050.

$\langle\cdot,\cdot\rangle\colon H\otimes H^*\to\Bbbk$ for which the structure of H^* is dual to that of H and vice-versa. For example, H could be finite-dimensional and H^* its linear dual, or H could be graded with each component finite-dimensional and H^* its graded dual. These algebras naturally act on each other [Mon93, 1.6.5]: suppose that $h\in H$ and $a\in H^*$ and set

$$h\rightharpoonup a \;:=\; \sum\langle h,a_2\rangle a_1 \qquad\text{and}\qquad a\rightharpoonup h \;:=\; \sum\langle h_2,a\rangle h_1\,. \tag{1}$$

(We use Sweedler notation for the coproduct, $\Delta h \;=\; \sum h_1\otimes h_2$.) These left actions are the adjoints of right multiplication: for $g,h\in H$ and $a,b\in H^*$,

$$\langle g, h\rightharpoonup a\rangle \;=\; \langle g\cdot h, a\rangle \qquad\text{and}\qquad \langle a\rightharpoonup h, b\rangle \;=\; \langle h, b\cdot a\rangle\,.$$

This shows that H^* is a left H-module under the action in (1). In fact, H^* is a left H–module algebra, meaning that for $a,b\in H^*$ and $h\in H$,

$$h\rightharpoonup(a\cdot b) \;=\; \sum(h_1\rightharpoonup a)\cdot(h_2\rightharpoonup b)\,. \tag{2}$$

Recall that the *counit* $\varepsilon\colon H\to\Bbbk$ and *antipode* $S\colon H\to H$ satisfy $\sum h_1\cdot\varepsilon(h_2)=h$ and $\sum h_1\cdot S(h_2)=\varepsilon(h)\cdot 1_H$ for all $h\in H$.

Lemma 1 *For $g,h\in H$ and $a\in H^*$, we have*

$$(a\rightharpoonup g)\cdot h \;=\; \sum(S(h_2)\rightharpoonup a)\rightharpoonup(g\cdot h_1)\,. \tag{3}$$

Proof: Let $b\in H^*$. We prove first the formula

$$(h\rightharpoonup b)\cdot a \;=\; \sum h_1\rightharpoonup(b\cdot(S(h_2)\rightharpoonup a))\,. \tag{4}$$

(This is essentially $(*)$ in the proof of Lemma 2.1.4 in [Mon93].) Expanding the sum using (2) and coassociativity, $(\Delta\otimes 1)\circ\Delta(h)=(1\otimes\Delta)\circ\Delta(h)=\sum h_1\otimes h_2\otimes h_3$, gives

$$\begin{aligned}\sum h_1\rightharpoonup(b\cdot(S(h_2)\rightharpoonup a)) \;&=\; \sum(h_1\rightharpoonup b)\cdot(h_2\rightharpoonup(S(h_3)\rightharpoonup a))\\ &=\; \sum(h_1\rightharpoonup b)\cdot((h_2\cdot S(h_3))\rightharpoonup a) &(5)\\ &=\; (h\rightharpoonup b)\cdot a\,. &(6)\end{aligned}$$

Here, (5) follows as H^* is an H-module and (6) from the antipode and counit conditions.

Note that $\langle(a\rightharpoonup g)\cdot h,b\rangle=\langle a\rightharpoonup g, h\rightharpoonup b\rangle=\langle g,(h\rightharpoonup b)\cdot a\rangle$. Using (4) this becomes

$$\begin{aligned}\Big\langle g,\sum h_1\rightharpoonup(b\cdot(S(h_2)\rightharpoonup a))\Big\rangle \;&=\; \sum\Big\langle g\cdot h_1, b\cdot(S(h_2)\rightharpoonup a)\Big\rangle\\ &=\; \Big\langle\,\sum(S(h_2)\rightharpoonup a)\rightharpoonup(g\cdot h_1), b\Big\rangle\,,\end{aligned}$$

which proves the lemma, as this holds for all $b\in H^*$. $\square$

Remark 2 This proof is identical to the argument in the appendix to [AM], where h was a complete homogeneous symmetric function in the Hopf algebra H of symmetric functions.

2 Application to distinguished bases

We apply Lemma 1 to produce skew Littlewood–Richardson rules for several Hopf algebras in algebraic combinatorics. We isolate the common features of those arguments.

In the notation of Section 1, let $\{L_\lambda\} \subset H$ and $\{R_\lambda\} \subset H^*$ be dual bases indexed by some set $\mathcal{P}$, so $\langle L_\lambda, R_\mu\rangle = \delta_{\lambda,\mu}$ for $\lambda, \mu \in \mathcal{P}$. Define structure constants for H and H^* via

$$L_\lambda \cdot L_\mu \;=\; \sum_\nu b^\nu_{\lambda,\mu} L_\nu \qquad \Delta(L_\nu) \;=\; \sum_{\lambda,\mu} c^\nu_{\lambda,\mu} L_\lambda \otimes L_\mu \;=\; \sum_\mu L_{\nu/\mu} \otimes L_\mu \tag{7}$$

$$R_\lambda \cdot R_\mu \;=\; \sum_\nu c^\nu_{\lambda,\mu} R_\nu \qquad \Delta(R_\nu) \;=\; \sum_{\lambda,\mu} b^\nu_{\lambda,\mu} R_\lambda \otimes R_\mu \;=\; \sum_\mu R_{\nu/\mu} \otimes R_\mu\,. \tag{8}$$

The *skew elements* $L_{\nu/\mu}$ and $R_{\nu/\mu}$ defined above co-multiply according to

$$\Delta(L_{\tau/\sigma}) \;=\; \sum_{\pi,\rho} c^\tau_{\pi,\rho,\sigma}\, L_\pi \otimes L_\rho \qquad \Delta(R_{\tau/\sigma}) \;=\; \sum_{\pi,\rho} b^\tau_{\pi,\rho,\sigma}\, R_\pi \otimes R_\rho\,. \tag{9}$$

(Note that the structure of H^* can be recovered from the structure of H. Thus, we may suppress the analogs of (8) and the second formula in (9) in the coming sections.)

Finally, suppose that the antipode acts on H in the L-basis according to the formula

$$S(L_\rho) \;=\; (-1)^{\mathsf{e}(\rho)} L_{\rho^{\mathsf{T}}} \tag{10}$$

for some functions $\mathsf{e}\colon \mathcal{P} \to \mathbb{N}$ and $(\cdot)^{\mathsf{T}}\colon \mathcal{P} \to \mathcal{P}$. Then Lemma 1 takes the following form.

Theorem 3 (Algebraic Littlewood–Richardson formula) *For any $\lambda, \mu, \sigma, \tau \in \mathcal{P}$, we have*

$$L_{\mu/\lambda} \cdot L_{\tau/\sigma} \;=\; \sum_{\pi,\rho,\lambda^-,\mu^+} (-1)^{\mathsf{e}(\rho)}\, c^\tau_{\pi,\rho,\sigma}\, b^\lambda_{\lambda^-,\rho^{\mathsf{T}}}\, b^{\mu^+}_{\mu,\pi}\; L_{\mu^+/\lambda^-}\,. \tag{11}$$

Swapping $L \leftrightarrow R$ and $b \leftrightarrow c$ in (11) *yields the analog for the skew elements $R_{\mu/\lambda}$ in H^*.*

Proof: The actions in (1) together with the second formulas for the coproducts in (7) and (8) show that $R_\lambda \rightharpoonup L_\mu = L_{\mu/\lambda}$ and $L_\lambda \rightharpoonup R_\mu = R_{\mu/\lambda}$. Now use (3) and (7)–(10) to obtain

$$\begin{aligned}
L_{\mu/\lambda} \cdot L_{\tau/\sigma} \;=\; (R_\lambda \rightharpoonup L_\mu) \cdot L_{\tau/\sigma} \;&=\; \sum_{\pi,\rho} (-1)^{\mathsf{e}(\rho)}\, c^\tau_{\pi,\rho,\sigma}\, \big((L_{\rho^{\mathsf{T}}} \rightharpoonup R_\lambda) \rightharpoonup (L_\mu \cdot L_\pi)\big)\\
&=\; \sum_{\pi,\rho,\mu^+} (-1)^{\mathsf{e}(\rho)}\, c^\tau_{\pi,\rho,\sigma}\, b^{\mu^+}_{\mu,\pi}\, \big(R_{\lambda/\rho^{\mathsf{T}}} \rightharpoonup L_{\mu^+}\big)\\
&=\; \sum_{\pi,\rho,\lambda^-,\mu^+} (-1)^{\mathsf{e}(\rho)}\, c^\tau_{\pi,\rho,\sigma}\, b^\lambda_{\lambda^-,\rho^{\mathsf{T}}}\, b^{\mu^+}_{\mu,\pi}\, (R_{\lambda^-} \rightharpoonup L_{\mu^+})\,.
\end{aligned}$$

This equals the right hand side of (11), since $R_{\lambda^-} \rightharpoonup L_{\mu^+} = L_{\mu^+/\lambda^-}$. □

Remark 4 The condition (10) is highly restrictive. It implies that the antipode S, as a linear map, is conjugate to a signed permutation matrix. Nevertheless, it holds for the Hopf algebras we consider. More generally, it holds if either H or H^* is commutative, for then S is an involution [Mon93, Cor. 1.5.12].

3 Skew Littlewood–Richardson rule for Schur functions

The commutative Hopf algebra Λ of symmetric functions is graded and self-dual under the Hall inner product $\langle\cdot,\cdot\rangle\colon \Lambda\otimes\Lambda\to\mathbb{Q}$. A systematic study of Λ from a Hopf algebra perspective appears in [Zel81]. We follow the definitions and notation in Chapter I of [Mac95]. The Schur basis of Λ (indexed by partitions) is self-dual, so (7) and (9) become

$$s_\lambda\cdot s_\mu \;=\; \sum_\nu c^\nu_{\lambda,\mu}s_\nu \qquad\qquad \Delta(s_\nu) \;=\; \sum_{\lambda,\mu} c^\nu_{\lambda,\mu}s_\lambda\otimes s_\mu \;=\; \sum_\mu s_{\nu/\mu}\otimes s_\mu \tag{12}$$

$$\Delta(s_{\tau/\sigma}) \;=\; \sum_{\pi,\rho} c^\tau_{\pi,\rho,\sigma}\, s_\pi\otimes s_\rho\,, \tag{13}$$

where the $c^\nu_{\lambda,\mu}$ are the *Littlewood–Richardson coefficients* and the $s_{\nu/\mu}$ are the *skew Schur functions* [Mac95, I.5]. Combinatorial expressions for the $c^\nu_{\lambda,\mu}$ and inner products $\langle s_{\mu/\lambda}, s_{\tau/\sigma}\rangle$ are derived using the Hopf algebraic structure of Λ in [Zel81]. The coefficients $c^\tau_{\pi,\rho,\sigma}$ occur in the triple product $s_\pi\cdot s_\rho\cdot s_\sigma$,

$$c^\tau_{\pi,\rho,\sigma} \;=\; \langle s_\pi\cdot s_\rho\cdot s_\sigma,\, s_\tau\rangle \;=\; \langle s_\pi\cdot s_\rho,\, s_{\tau/\sigma}\rangle \;=\; \langle s_\pi\otimes s_\rho,\, \Delta(s_{\tau/\sigma})\rangle\,.$$

Write ρ' for the conjugate (matrix-transpose) of ρ. Then the action of the antipode is

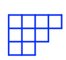
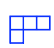
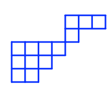
$$S(s_\rho) \;=\; (-1)^{|\rho|}s_{\rho'}\,, \tag{14}$$

which is just a twisted form of the fundamental involution ω that sends s_ρ to $s_{\rho'}$. Indeed, the formula $\sum_{i+j=n}(-1)^i e_i h_j \;=\; \delta_{0,n}$ shows that (14) holds on the generators $\{h_n \mid n\geq 1\}$ of Λ. The validity of (14) follows as both S and ω are algebra maps.

Since $c^\lambda_{\lambda^-,\rho'}=0$ unless $|\rho|=|\lambda/\lambda^-|$, we may write (11) as

$$s_{\mu/\lambda}\cdot s_{\tau/\sigma} \;=\; \sum_{\pi,\rho,\lambda^-,\mu^+} (-1)^{|\lambda/\lambda^-|}\; c^\tau_{\pi,\rho,\sigma}\; c^\lambda_{\lambda^-,\rho'}\; c^{\mu^+}_{\mu,\pi}\;\; s_{\mu^+/\lambda^-}\;. \tag{15}$$

We next formulate a combinatorial version of (15). Given partitions ρ and σ, form the skew shape $\rho*\sigma$ by placing ρ southwest of σ. Thus,

if ρ =  and σ =  then $\rho*\sigma$ =  .

Similarly, if R is a tableau of shape ρ and S a tableau of shape σ, then $R*S$ is the skew tableau of shape $\rho*\sigma$ obtained by placing R southwest of S. Fix a tableau T of shape τ. The Littlewood–Richardson coefficient $c^\tau_{\rho,\sigma}$ is the number of pairs (R,S) of tableaux of respective shapes ρ and σ with $R*S$ Knuth-equivalent to T. See [Ful97, Ch. 5, Cor. 2(v)]. Similarly, $c^\tau_{\pi,\rho,\sigma}$ is the number of triples (P,R,S) of tableaux of respective shapes π, ρ, and σ with $P*R*S$ Knuth-equivalent to T.

Write $\mathrm{sh}(S)$ for the shape of a tableau S and $S\equiv_K T$ if S is Knuth-equivalent to T.

Lemma 5 *Let σ,τ be partitions and fix a tableau T of shape τ. Then*

$$\Delta(s_{\tau/\sigma}) \;=\; \sum s_{\mathrm{sh}(R^-)}\otimes s_{\mathrm{sh}(R^+)}\,,$$

the sum taken over triples (R^-,R^+,S) of tableaux with $\mathrm{sh}(S)=\sigma$ and $R^-*R^+*S\equiv_K T$. $\square$

Note that $(\mu/\lambda)' = \mu'/\lambda'$ and the operation $*$ makes sense for skew tableaux. If S is a tableau of skew shape μ/λ, put $|S| = |\mu/\lambda| = |\mu| - |\lambda|$.

Theorem 6 (Skew Littlewood–Richardson rule) *Let $\lambda, \mu, \sigma, \tau$ be partitions and fix a tableau T of shape τ. Then*

$$s_{\mu/\lambda} \cdot s_{\tau/\sigma} \;=\; \sum (-1)^{|S^-|}\, s_{\mu^+/\lambda^-}\ , \tag{16}$$

the sum taken over triples (S^-, S^+, S) of skew tableaux of respective shapes $(\lambda/\lambda^-)'$, μ^+/μ, and σ such that $S^- * S^+ * S \equiv_K T$.

Remark 7 If T is the unique *Yamanouchi tableau* of shape τ whose ith row contains only the letter i, then this is *almost* Conjecture 6.1 in [AM]. Indeed, in this case S is Yamanouchi of shape σ, so the sum is really over pairs of tableaux, and this explains the σ-Yamanouchi condition in [AM]. The difference lies in the tableau S^- and the reading word condition in [AM]. It is an exercise in tableaux combinatorics that there is a bijection between the indices (S^-, S^+) of Theorem 6 and the corresponding indices of Conjecture 6.1 in [AM].

Proof Proof of Theorem 6: We reinterpret (15) in terms of tableaux. Let (R^-, R^+, S) be a triple of tableaux of partition shape with $\mathrm{sh}(S) = \sigma$ and $R^- * R^+ * S \equiv_K T$. If $\mathrm{sh}(R^-) = \rho$, then by [Ful97, Ch. 5, Cor. 2(i)], $c^{\lambda}_{\lambda^-,\rho'} = c^{\lambda'}_{(\lambda^-)',\rho}$ counts skew tableaux S^- of shape $(\lambda/\lambda^-)'$ with $S^- \equiv_K R^-$. Likewise, if $\mathrm{sh}(R^+) = \pi$, then $c^{\mu^+}_{\mu,\pi}$ counts skew tableaux S^+ of shape μ^+/μ with $S^+ \equiv_K R^+$. Now (15) may be written as

$$s_{\mu/\lambda} \cdot s_{\tau/\sigma} \;=\; \sum (-1)^{|S^-|} s_{\mu^+/\lambda^-}\ ,$$

summing over skew tableaux (R^-, R^+, S^-, S^+, S) with $R^\pm$ of partition shape, $\mathrm{sh}(S) = \sigma$, $R^- {*} R^+ {*} S \equiv_K T$, $\mathrm{sh}(S^+) = \mu^+/\mu$, $\mathrm{sh}(S^-) = (\lambda/\lambda^-)'$, and $S^\pm \equiv_K R^\pm$.

Finally, note that $R^\pm$ is the unique tableau of partition shape Knuth-equivalent to $S^\pm$. Since $S^- {*} S^+ {*} S$ is Knuth-equivalent to T (by transitivity of $\equiv_K$), we omit the unnecessary tableaux $R^\pm$ from the indices of summation and reach the statement of the theorem. □

4 Skew Littlewood–Richardson rule for Schur P- and Q-functions

The self-dual Hopf algebra of symmetric functions has a natural self-dual subalgebra Ω. This has dual bases the Schur P- and Q-functions [Mac95, III.8], which are indexed by *strict partitions* $\lambda\colon \lambda_1 > \cdots > \lambda_l > 0$. Write $\ell(\lambda) = l$ for the length of the partition λ. As in Section 3, the constants and skew functions in the structure equations

$$Q_\lambda \cdot Q_\mu \;=\; \sum_\nu g^{\nu}_{\lambda,\mu}\, Q_\nu \qquad\qquad \Delta(Q_\nu) \;=\; \sum_{\lambda,\mu} f^{\nu}_{\lambda,\mu} Q_\lambda \otimes Q_\mu \;=\; \sum_\mu Q_{\nu/\mu} \otimes Q_\mu \tag{17}$$

$$\Delta(Q_{\tau/\sigma}) \;=\; \sum_{\pi,\rho} f^{\tau}_{\pi,\rho,\sigma}\, Q_\pi \otimes Q_\rho \tag{18}$$

have combinatorial interpretations (see below). Also, each basis $\{P_\lambda\}$ and $\{Q_\lambda\}$ is *almost* self-dual in that $P_\lambda = 2^{-\ell(\lambda)} Q_\lambda$ and $g^{\nu}_{\lambda,\mu} = 2^{\ell(\lambda)+\ell(\mu)-\ell(\nu)} f^{\nu}_{\lambda,\mu}$.

The algebra Ω is generated by the special Q-functions $q_n = Q_{(n)} := \sum_{i+j=n} h_i e_j$ [Mac95, III, (8.1)]. This implies that $S(q_n) = (-1)^n q_n$, from which we deduce that

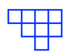
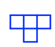
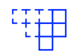
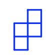
$$S(Q_\rho) = (-1)^{|\rho|} Q_\rho\,.$$

As $f^{\lambda}_{\lambda^-,\rho} = 0$ unless $|\rho| = |\lambda/\lambda^-|$, we may write the algebraic rule (11) as

$$Q_{\mu/\lambda} \cdot Q_{\tau/\sigma} \;=\; \sum_{\pi,\rho,\lambda^-,\mu^+} (-1)^{|\lambda/\lambda^-|}\, f^{\tau}_{\pi,\rho,\sigma}\, g^{\lambda}_{\lambda^-,\rho}\, g^{\mu^+}_{\mu,\pi}\; Q_{\mu^+/\lambda^-}\;, \tag{19}$$

with a similar identity holding for $P_{\mu/\lambda} \cdot P_{\tau/\sigma}$ (swapping $P \leftrightarrow Q$ and $f \leftrightarrow g$).

We formulate two combinatorial versions of (19). Strict partitions λ, μ are written as *shifted Young diagrams* (where row i begins in column i). Skew shifted shapes λ/μ are defined in the obvious manner:

if $\lambda \;=\; 431 \;=$  and $\mu \;=\; 31 \;=$ , then $\lambda/\mu \;=$  $=$ .

In what follows, tableaux means *semi-standard (skew) shifted tableaux on a marked alphabet* [Mac95, III.8]. We use shifted versions of the jeu-de-taquin and plactic equivalence from [Sag87] and [Ser], denoting the corresponding relations by $\equiv_{\text{SJ}}$ and $\equiv_{\text{SP}}$, respectively. Given tableaux R, S, T, we write $R * S \equiv_{\text{SP}} T$ when representative words u, v, w (built via "mread" [Ser, §2]) of the corresponding shifted plactic classes satisfy $uv \equiv_{\text{SP}} w$.

Stembridge notes (following [Ste89, Prop. 8.2]) that for a fixed tableau M of shape μ,

$$f^{\nu}_{\lambda,\mu} \;=\; \#\bigl\{\text{skew tableaux } L : \text{sh}(L) = \nu/\lambda \text{ and } L \equiv_{\text{SJ}} M\bigr\}. \tag{20}$$

Serrano has a similar description of these coefficients in terms of $\equiv_{\text{SP}}$. Fixing a tableau T of shape τ, it follows from [Ser, Cor. 1.15] that the coefficient $f^{\tau}_{\pi,\rho,\sigma}$ in $P_\pi \cdot P_\rho \cdot P_\sigma = \sum_\tau f^{\tau}_{\pi,\rho,\sigma} P_\tau$ is given by

$$f^{\tau}_{\pi,\rho,\sigma} \;=\; \#\bigl\{(P,R,S) : \text{sh}(P) = \pi,\ \text{sh}(R) = \rho,\ \text{sh}(S) = \sigma,\ \text{and } P * R * S \equiv_{\text{SP}} T\bigr\}. \tag{21}$$

If T is a tableau of shape λ, write $\ell(T)$ for $\ell(\lambda)$. The formula relating the g's and f's combines with (20) and (21) to give our next result.

Theorem 8 (Skew Littlewood–Richardson rule) *Let $\lambda, \mu, \sigma, \tau$ be strict partitions and fix a tableau T of shape τ. Then*

$$Q_{\mu/\lambda} \cdot Q_{\tau/\sigma} \;=\; \sum (-1)^{|\lambda/\lambda^-|}\, 2^{\ell(R^-)+\ell(R^+)+\ell(\lambda^-)+\ell(\mu)-\ell(\lambda)-\ell(\mu^+)}\; Q_{\mu^+/\lambda^-}\;, \tag{22}$$

the sum taken over quintuples (R^-, R^+, S^-, S^+, S) with $R^\pm$ of partition shape, $\text{sh}(S) = \sigma$, $R^- * R^+ * S \equiv_{\text{SP}} T$, $\text{sh}(S^+) = \mu^+/\mu$, $\text{sh}(S^-) = (\lambda/\lambda^-)$, *and* $S^\pm \equiv_{\text{SJ}} R^\pm$. □

Serrano conjectures an elegant combinatorial description [Ser, Conj. 2.12 and Cor. 2.13] of the structure constants $g^{\nu}_{\lambda,\mu}$ in (17): For any tableau M of shape μ, he conjectured

$$g^{\nu}_{\lambda,\mu} \;=\; \#\bigl\{\text{skew tableaux } L : \text{sh}(L) = \nu/\lambda \text{ and } L \equiv_{\text{SP}} M\bigr\}. \tag{23}$$

(Note that if S, T are tableaux, then $S \equiv_{\text{SP}} T$ does not necessarily imply that $S \equiv_{\text{SJ}} T$.) This leads to a conjectural reformulation of Theorem 8 in the spirit of Theorem 6.

Conjecture 1 (Conjectural Skew Littlewood–Richardson rule) *Let $\lambda, \mu, \sigma, \tau$ be strict partitions, and fix a tableau T of shape τ. Then*

$$Q_{\mu/\lambda} \cdot Q_{\tau/\sigma} \;=\; \sum (-1)^{|S^-|}\, Q_{\mu^+/\lambda^-}\ , \tag{24}$$

the sum taken over triples (S^-, S^+, S) of skew tableaux of respective shapes (λ/λ^-), μ^+/μ, and σ such that $S^- * S^+ * S \equiv_{\text{SP}} T$.

Proof: There is a unique shifted tableau R in any shifted plactic class [Ser, Thm. 2.8]. So the conditions $S^\pm \equiv_{\text{SP}} R^\pm$ and $R^- * R^+ * S \equiv_{\text{SP}} T$ in (21) and (23) may be replaced with the single condition $S^- * S^+ * S \equiv_{\text{SP}} T$. $\square$

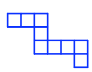
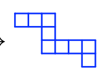
5 Skew Littlewood–Richardson rule for noncommutative Schur functions

The Hopf algebra of noncommutative symmetric functions was introduced, independently, in [GKL$^+$95, MR95] as the (graded) dual to the commutative Hopf algebra of quasisymmetric functions. We consider the dual bases (indexed by compositions) $\{F_\alpha\}$ of Gessel's quasisymmetric functions and $\{R_\alpha\}$ of noncommutative ribbon Schur functions. The structure constants in

$$R_\alpha \cdot R_\beta \;=\; \sum_\gamma b^\gamma_{\alpha,\beta}\, R_\gamma \qquad\qquad \Delta(R_\gamma) \;=\; \sum_{\alpha,\beta} c^\gamma_{\alpha,\beta} R_\alpha \otimes R_\beta \;=\; \sum_\beta R_{\gamma/\beta} \otimes R_\beta \tag{25}$$

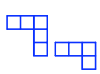
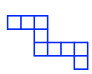
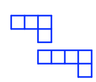
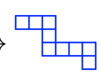
$$\Delta(R_{\tau/\sigma}) \;=\; \sum_{\pi,\rho,\sigma} c^\tau_{\pi,\rho,\sigma}\, R_\pi \otimes R_\rho \tag{26}$$

may be given combinatorial meaning via the descent map $\mathsf{d}\colon \mathfrak{S}_n \to \Gamma_n$ from permutations to compositions and a section of it $\mathsf{w}\colon \Gamma_n \to \mathfrak{S}_n$. These maps are linked via *ribbon diagrams*, edge-connected skew Young diagrams (written in the french style), with no 2×2 subdiagram present. By way of example,

$\mathsf{d}\colon\ 148623795 \to$  $\to 3141,$ $\qquad$ $\mathsf{w}\colon\ 3141 \to$  $\to 789623451.$

(In the intermediate step for $\mathsf{d}(w)$, new rows in the ribbon begin at descents of w. In the intermediate step for $\mathsf{w}(\alpha)$, the boxes in the ribbon are filled left-to-right, bottom-to-top.)

A ribbon α may be extended by a ribbon β in two ways: affixing β to the rightmost edge or bottommost edge of α (written $\alpha \triangleleft \beta$ and $\alpha \vartriangle \beta$, respectively):

$311 \triangleleft 31:$  $\to$  $,$ $\qquad$ $31 \vartriangle 41:$  $\to$  $.$

If a ribbon γ is formed from α and β in either of these two ways, we write $\gamma \in \alpha \Diamond \beta$. The coefficient $b^\gamma_{\alpha,\beta}$ is 1, if $\gamma \in \alpha \Diamond \beta$, and 0 otherwise. If $*$ is the shifted shuffle product on permutations (see (3.4) in [MR95]), then the coefficient $c^\gamma_{\alpha,\beta}$ is the number of words w in $\mathsf{w}(\alpha) * \mathsf{w}(\beta)$ such that $\mathsf{d}(w) = \gamma$. The coefficient $c^\tau_{\pi,\rho,\sigma}$ has the analogous description.

Antipode formulas for the distinquished bases were found, independently, in [Ehr96, MR95]:

$$S(F_\alpha) \;=\; (-1)^{|\alpha|} F_{\alpha'} \qquad \text{and} \qquad S(R_\alpha) \;=\; (-1)^{|\alpha|} R_{\alpha'}\,,$$

where α' is the conjugate of α (in the sense of french style skew partitions). For example, $(3141)' = 211311$. The descriptions of the antipode and structure constants in (25) and (26) give a formula for the product of two skew ribbon Schur functions.

Theorem 9 (Skew Littlewood–Richardson rule) *Let $\alpha, \beta, \sigma, \tau$ be compositions. Then*

$$R_{\beta/\alpha} \cdot R_{\tau/\sigma} \;=\; \sum (-1)^{|\rho|}\, R_{\beta^+/\alpha^-}\;,$$

the sum taken over factorizations $\alpha \in \alpha^- \Diamond \rho'$, extensions $\beta^+ \in \beta \Diamond \pi$, and words w in the shuffle product $\mathsf{w}(\pi) * \mathsf{w}(\rho) * \mathsf{w}(\sigma)$ *such that* $\mathsf{d}(w) = \tau$. □

Remark 10 The nonzero skew ribbon Schur functions $R_{\beta/\alpha}$ do not correspond to skew ribbon shapes in a simple way. For example, 111 is not a (connected) sub-ribbon of 221, yet $R_{221/111} = R_2 + R_{11} \neq 0$. Contrast this with the skew functions $F_{\beta/\alpha}$, where $F_{\beta/\alpha} \neq 0$ if and only if $\beta \in \omega \Diamond \alpha$ for some ribbon ω. That is, $F_{\beta/\alpha} = F_\omega$. Thus we may view α as the last $|\alpha|$ boxes of the ribbon β and β/α as the complementary ribbon ω. Interpreting $F_{\beta/\alpha} \cdot F_{\tau/\sigma}$ alternately as a product of ordinary functions or skew functions yields the curious identity

$$F_{\beta/\alpha} \cdot F_{\tau/\sigma} \;=\; \sum_\gamma c^{\gamma}_{\beta/\alpha,\tau/\sigma}\, F_\gamma \;=\; \sum_{\pi,\rho,\alpha^-,\beta^+} (-1)^{|\rho|}\, b^{\tau/\sigma}_{\pi,\rho}\, c^{\alpha}_{\alpha^-,\rho'}\, c^{\beta^+}_{\beta,\pi}\; F_{\beta^+/\alpha^-}\,. \tag{27}$$

6 Skew k-Pieri rule for k-Schur functions

Fix an integer $k \geq 1$. Let $\Lambda_{(k)}$ denote the Hopf subalgebra of the Hopf algebra of symmetric functions generated by the homogeneous symmetric functions $h_1, h_2, \ldots, h_k$. Let $\Lambda^{(k)}$ denote the Hopf-dual quotient Hopf algebra of symmetric functions. We consider the dual bases $\{s^{(k)}_\lambda\} \subset \Lambda_{(k)}$ and $\{F^{(k)}_\lambda\} \subset \Lambda^{(k)}$ of k-Schur functions and dual k-Schur functions of [LLMS, LM07], also called strong Schur functions and weak Schur functions in [LLMS]. The k-Schur functions were first introduced by Lapointe, Lascoux, and Morse in the context of Macdonald polynomials, and were later shown by Lam to represent Schubert classes in the affine Grassmannian of $\mathrm{SL}(k{+}1, \mathbb{C})$. We refer the reader to the references in [LLMS].

Here λ varies over all k-bounded partitions, that is, those partitions satisfying $\lambda_1 \leq k$. There is an involution $\lambda \mapsto \lambda^{\omega_k}$ on k-bounded partitions called *k-conjugation*. We have

$$S(s^{(k)}_\lambda) \;=\; (-1)^{|\lambda|} s^{(k)}_{\lambda^{\omega_k}} \qquad \text{and} \qquad S(F^{(k)}_\lambda) \;=\; (-1)^{|\lambda|} F^{(k)}_{\lambda^{\omega_k}}\,.$$

If $\lambda = (r)$ is a one-part partition, then $s^{(k)}_\lambda = h_r$ is a homogeneous symmetric function. We have the k-Pieri and dual k-Pieri rules [LLMS, LM07] (called weak and strong Pieri rules in [LLMS])

$$s^{(k)}_\lambda \cdot h_r \;=\; \sum_{\lambda \rightsquigarrow_r \mu} s^{(k)}_\mu \qquad \text{and} \qquad F^{(k)}_\lambda \cdot h_r \;=\; \sum_{\lambda \to_r \mu} F^{(k)}_\mu \tag{28}$$

for $r \leq k$. Here $\lambda \rightsquigarrow_r \mu$ denotes an *r-weak strip* connecting λ and μ—present if and only if both μ/λ and $\mu^{\omega_k}/\lambda^{\omega_k}$ are horizontal r-strips. The notation $\lambda \to_r \mu$ denotes an *r-strong strip* as introduced in [LLMS], which we will not define here. The terminology comes from the relationship with the weak and strong (Bruhat) orders of the affine symmetric group. We remark that there may be distinct strong strips $\lambda \to_r \mu$ and $(\lambda \to_r \mu)'$ which start and end at the same partition, so that the second Pieri rule of (28) may have multiplicities. (Strictly speaking, the strong strips in [LLMS] are built on $(k+1)$-cores, and our $\lambda \to_r \mu$ denotes the strips obtained after applying a bijection between $(k+1)$-cores and k-bounded partitions.)

We define skew functions $s^{(k)}_{\lambda/\mu}$ and $F^{(k)}_{\lambda/\mu}$ using (7) and (8). There is an explicit combinatorial description of $F^{(k)}_{\lambda/\mu}$ in terms of the weak tableaux of [LLMS], but only a conjectured combinatorial description of $s^{(k)}_{\lambda/\mu}$ [LLMS, Conj. 4.18(3)].

Theorem 11 (Skew k-Pieri (or weak Pieri) rule) *For k-bounded partitions λ, μ, and $r \le k$,*

$$s^{(k)}_{\mu/\lambda} \cdot h_r \;=\; \sum_{i+j=r} (-1)^j \sum_{\substack{\mu \rightsquigarrow_i \mu^+ \\ (\lambda^-)^{\omega_k} \rightsquigarrow_j \lambda^{\omega_k}}} s^{(k)}_{\mu^+/\lambda^-}$$

Proof: In Theorem 3, take $L_{\tau/\sigma} = h_r$. For $c^{\tau}_{\pi,\rho,\sigma}$, use the formula $\Delta(h_r) = \sum_{i+j=r} h_i \otimes h_j$, and for $b^{\lambda}_{\lambda^-,\rho^{\omega_k}}$ and $b^{\mu^+}_{\mu,\pi}$, use (28). □

Theorem 12 (Skew dual k-Pieri (or strong Pieri) rule) *For k-bounded partitions λ, μ, and $r \le k$,*

$$F^{(k)}_{\lambda/\mu} \cdot h_r \;=\; \sum_{i+j=r} (-1)^j \sum_{\substack{\lambda \to_i \lambda^+ \\ (\mu^-)^{\omega_k} \to_j \mu^{\omega_k}}} F^{(k)}_{\lambda^+/\mu^-}$$

Proof: *Identical to the proof of Theorem 11.* □

As an example, let $k = 2$, $r = 2$, $\mu = (2, 1, 1)$, and $\lambda = (1)$. Then Theorem 11 states that

$$s^{(2)}_{211/1} \cdot h_2 \;=\; s^{(2)}_{2211/1} - s^{(2)}_{2111}\,,$$

which one can verify using (28) and the expansions $s^{(2)}_{211/1} = s^{(2)}_{21} + s^{(2)}_{111}$ and $s^{(2)}_{2211/1} = 2\,s^{(2)}_{2111} + s^{(2)}_{221}$. Theorem 12 states that

$$F^{(2)}_{211/1} \cdot h_2 \;=\; 3\,F^{(2)}_{222/1} + 5\,F^{(2)}_{2211/1} + 3\,F^{(2)}_{21111/1} + 3\,F^{(2)}_{111111/1} - 2\,F^{(2)}_{221} - 3\,F^{(2)}_{2111} - 2\,F^{(2)}_{11111}.$$

One can verify that both sides are equal to $6\,F^{(2)}_{221} + 5\,F^{(2)}_{2111} + 4\,F^{(2)}_{11111}$.

7 Skew Pieri rule for affine Grassmannian of the symplectic group

Fix $n \ge 1$. The Hopf algebra Ω of Section 4 contains a Hopf subalgebra $\Omega_{(n)}$ generated by the Schur P-functions $P_1, P_3, \ldots, P_{2n-1}$. In [LSS10], it was shown that $\Omega_{(n)}$ is isomorphic to the homology ring $H_*(Gr_{\mathrm{Sp}(2n,\mathbb{C})})$ of the affine Grassmannian of the symplectic group $\mathrm{Sp}(2n, \mathbb{C})$. A distinguished basis

$\{P_w^{(n)}\} \subset \Omega_{(n)}$, representing the Schubert basis, was studied there. The symmetric functions $P_w^{(n)}$ are shifted versions of the k-Schur functions of Section 6.

The indexing set for the basis $\{P_w^{(n)}\}$ is the set $\widetilde{C}_n^0$ of affine Grassmannian type C permutations: they are the minimal length coset representatives of C_n in $\widetilde{C}_n$. A lower Bruhat order ideal $\mathcal{Z} \subset \widetilde{C}_n$ of the affine type C Weyl group is defined in [LSS10]. Let $\mathcal{Z}_j \subset \mathcal{Z}$ denote those $v \in \mathcal{Z}$ with length $\ell(v) = j$. For each $v \in \mathcal{Z}$, there is a nonnegative integer $c(v) \in \mathbb{Z}_{\geq 0}$, called the *number of components* of v. We note that $c(\mathrm{id}) = 0$. With this notation, for each $1 \leq j \leq 2n-1$, we have the Pieri rule [LSS10, Thms. 1.3 and 1.4]

$$P_w^{(n)} \cdot P_j \;=\; \sum_{v \in \mathcal{Z}_j} 2^{c(v)-1}\, P_{vw}^{(n)}\,, \tag{29}$$

where the sum is over all $v \in \mathcal{Z}_j$ such that $vw \in \widetilde{C}_n^0$, and $\ell(vw) = \ell(v) + \ell(w)$.

It follows from the discussion in Section 4 that the antipode acts on the P_j by $S(P_j) = (-1)^j P_j$. We define $P_{w/v}^{(n)}$ using (7).

Theorem 13 (Skew Pieri rule) *For $w, v \in \widetilde{C}_n^0$, and $r \leq 2n-1$,*

$$P_{w/v}^{(n)} \cdot P_r \;=\; \sum_{i+j=r} (-1)^j \sum_{\substack{u \in \mathcal{Z}_i \\ z \in \mathcal{Z}_j}} 2^{c(u)+c(z)-1}\, P_{uw/z^{-1}v}^{(n)}\,,$$

where the sum is over all $u \in \mathcal{Z}_i$ and $z \in \mathcal{Z}_j$ such that $uw, z^{-1}v \in \widetilde{C}_n^0$, $\ell(uw) = \ell(u) + \ell(w)$, and $\ell(z^{-1}v) + \ell(z) = \ell(v)$.

Proof: In Theorem 3, take $L_{\tau/\sigma} = P_r$ and use (29). For the constants $c^{\tau}_{\pi,\rho,\sigma}$, use the formula

$$\Delta(P_r) \;=\; 1 \otimes P_r + P_r \otimes 1 + 2 \sum_{0<j<r} P_{r-j} \otimes P_j\,.$$

If $0 < j < r$, the product $c^{\tau}_{\pi,\rho,\sigma} b^{\lambda}_{\lambda^-,\rho'} b^{\mu^+}_{\mu,\pi}$ in (11) becomes $2 \cdot 2^{c(u)-1} \cdot 2^{c(z)-1} = 2^{c(u)+c(z)-1}$. If $j = 0$ (resp., $j = r$), it becomes $1 \cdot 2^{c(u)-1} \cdot 1 = 2^{c(u)+c(z)-1}$ (resp., $1 \cdot 1 \cdot 2^{c(z)-1} = 2^{c(u)+c(z)-1}$). □

References

[AM] Sami H. Assaf and Peter R. McNamara. A Pieri rule for skew shapes. *J. Combin. Theory Ser. A*. To appear.

[Ehr96] Richard Ehrenborg. On posets and Hopf algebras. *Adv. Math.*, 119(1):1–25, 1996.

[Ful97] William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997.

[GKL+95] Israel M. Gelfand, Daniel Krob, Alain Lascoux, Bernard Leclerc, Vladimir S. Retakh, and Jean-Yves Thibon. Noncommutative symmetric functions. *Adv. Math.*, 112(2):218–348, 1995.

[LLMS] Thomas Lam, Luc Lapointe, Jennifer Morse, and Mark Shimozono. Affine insertion and Pieri rules for the affine Grassmannian. *Memoirs of the AMS*. To appear.

[LM07] Luc Lapointe and Jennifer Morse. A k-tableau characterization of k-Schur functions. *Adv. Math.*, 213(1):183–204, 2007.

[LSS10] Thomas Lam, Anne Schilling, and Mark Shimozono. Schubert polynomials for the affine Grassmannian of the symplectic group. *Math. Z.*, 264(4):765–811, 2010.

[Mac95] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995.

[Mon93] Susan Montgomery. *Hopf algebras and their actions on rings*, volume 82 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1993.

[MR95] Claudia Malvenuto and Christophe Reutenauer. Duality between quasi-symmetric functions and the Solomon descent algebra. *J. Algebra*, 177(3):967–982, 1995.

[Sag87] Bruce E. Sagan. Shifted tableaux, Schur Q-functions, and a conjecture of R. Stanley. *J. Combin. Theory Ser. A*, 45(1):62–103, 1987.

[Ser] Luis Serrano. The shifted plactic monoid. *Math. Z.* To appear.

[Ste89] John R. Stembridge. Shifted tableaux and the projective representations of symmetric groups. *Adv. Math.*, 74(1):87–134, 1989.

[Zel81] Andrey V. Zelevinsky. *Representations of finite classical groups. A Hopf algebra approach*, volume 869 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1981.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 245–256

QSym over Sym has a stable basis

Aaron Lauve[1] and Sarah K Mason[2†]

[1]*Department of Mathematics*
Texas A&M University
College Station, TX 778439

[2]*Department of Mathematics*
Wake Forest University
Winston-Salem, NC 27109

Abstract. We prove that the subset of quasisymmetric polynomials conjectured by Bergeron and Reutenauer to be a basis for the coinvariant space of quasisymmetric polynomials is indeed a basis. This provides the first constructive proof of the Garsia–Wallach result stating that quasisymmetric polynomials form a free module over symmetric polynomials and that the dimension of this module is $n!$.

Résumé. Nous prouvons que le sous-ensemble des polynômes quasisymétriques conjecturé par Bergeron et Reutenauer pour former une base pour l'espace coinvariant des polynômes quasisymétriques est en fait une base. Cela fournit la première preuve constructive du résultat de Garsia–Wallach indiquant que les polynômes quasisymétriques forment un module libre sur les polynômes symétriques et que la dimension de ce module est $n!$.

Keywords: quasisymmetric functions, symmetric functions, free modules, compositions, inverting compositions

1 Introduction

Quasisymmetric polynomials have held a special place in algebraic combinatorics since their introduction in [7]. They are the natural setting for many enumeration problems [16] as well as the development of Dehn–Somerville relations [1]. In addition, they are related in a natural way to Solomon's descent algebra of the symmetric group [14]. In this paper, we follow [2, Chapter 11] and view them through the lens of invariant theory. Specifically, we consider the relationship between the two subrings $Sym_n \subseteq QSym_n \subseteq \mathbb{Q}[\mathbf{x}]$ of symmetric and quasisymmetric polynomials in variables $\mathbf{x} = \mathbf{x}_n := \{x_1, x_2, \ldots, x_n\}$. Let J_n denote the ideal in $QSym_n$ generated by the elementary symmetric polynomials. In 2002, F. Bergeron and C. Reutenauer made a sequence of three successively finer conjectures concerning the quotient ring $QSym_n/J_n$. A. Garsia and N. Wallach were able to prove the first two in [6], but the third one remained open; we close it here (Corollary 6) with the help of a new basis for $QSym_n$ introduced in [8].

Acknowledgements

We would like to thank François Bergeron for sharing the preceding story with us and encouraging us to write this paper. Our approach follows an idea that he proposed during CanaDAM 2009. We also gratefully acknowledge several beneficial discussions we had with Adriano Garsia, Christophe Reutenauer, and Frank Sottile.

†Partially supported by NSF postdoctoral fellowship DMS-0603351.

1.1 Motivating context

Recall that Sym_n is the ring $\mathbb{Q}[\mathbf{x}]^{\mathfrak{S}_n}$ of invariants under the permutation action of $\mathfrak{S}_n$ on $\mathbf{x}$ and $\mathbb{Q}[\mathbf{x}]$. One of the crowning results in the invariant theory of $\mathfrak{S}_n$ is that the following true statements are equivalent:

(S1) $\mathbb{Q}[\mathbf{x}]^{\mathfrak{S}_n}$ is a polynomial ring, generated, say, by the elementary symmetric polynomials $\mathcal{E}_n = \{e_1(\mathbf{x}), \dots, e_n(\mathbf{x})\}$;

(S2) the ring $\mathbb{Q}[\mathbf{x}]$ is a free $\mathbb{Q}[\mathbf{x}]^{\mathfrak{S}_n}$-module;

(S3) the *coinvariant space* $\mathbb{Q}[\mathbf{x}]_{\mathfrak{S}_n} = \mathbb{Q}[\mathbf{x}]/\big(\mathcal{E}_n\big)$ has dimension $n!$ and is isomorphic to the regular representation of $\mathfrak{S}_n$.

See [11, §§17, 18] for details. Analogous statements hold on replacing $\mathfrak{S}_n$ by any pseudo-reflection group. Since all spaces in question are graded, we may add a fourth item to the list: the ***Hilbert series*** $H_q\big(\mathbb{Q}[\mathbf{x}]_{\mathfrak{S}_n}\big) = \sum_{k\geq 0} d_k\, q^k$, where d_k records the dimension of the kth graded piece of $\mathbb{Q}[\mathbf{x}]_{\mathfrak{S}_n}$, satisfies

(S4) $H_q\big(\mathbb{Q}[\mathbf{x}]_{\mathfrak{S}_n}\big) = H_q\big(\mathbb{Q}[\mathbf{x}]\big) \Big/ H_q\big(\mathbb{Q}[\mathbf{x}]^{\mathfrak{S}_n}\big)$.

Before we formulate the conjectures of Bergeron and Reutenauer, we recall another page in the story of Sym_n and the quotient space $\mathbb{Q}[\mathbf{x}]/\left(\mathcal{E}_n\right)$. The ring homomorphism ζ from $\mathbb{Q}[\mathbf{x}_{n+1}]$ to $\mathbb{Q}[\mathbf{x}_n]$ induced by the mapping $x_{n+1} \mapsto 0$ respects the rings of invariants (that is, $\zeta\colon Sym_{n+1} \twoheadrightarrow Sym_n$ is a ring homomorphism). Moreover, ζ respects the fundamental bases of monomial (m_λ) and Schur (s_λ) symmetric polynomials of Sym_n, indexed by partitions λ with at most n parts. For example,

$$\zeta(m_\lambda(\mathbf{x}_{n+1})) = \begin{cases} m_\lambda(\mathbf{x}_n), & \text{if } \lambda \text{ has at most } n \text{ parts,} \\ 0, & \text{otherwise.} \end{cases}$$

The stability of these bases plays a crucial role in representation theory [13]. Likewise, the associated stability of bases for the coinvariant spaces (e.g., of Schubert polynomials [4, 12, 15]) plays a role in the cohomology theory of flag varieties.

1.2 Bergeron–Reutenauer context

Given that $QSym_n$ is a polynomial ring [14] containing Sym_n, one might ask, by analogy with $\mathbb{Q}[\mathbf{x}]$, how $QSym_n$ looks as a module over Sym_n. This was the question investigated by Bergeron and Reutenauer [3]. (See also [2, §11.2].) They began by computing the quotient $P_n(q) := H_q\big(QSym_n\big) \big/ H_q\big(Sym_n\big)$ by analogy with (S4). To everyone's surprise, the result was a polynomial in q with nonnegative integer coefficients (so it could, conceivably, enumerate the graded space $QSym_n/J_n$). More astonishingly, sending q to 1 gave $P_n(1) = n!$. This led to the following two conjectures, subsequently proven in [6]:

(Q1) The ring $QSym_n$ is a free module over Sym_n;

(Q2) The dimension of the "coinvariant space" $QSym_n/J_n$ is $n!$.

In their efforts to solve the conjectures above, Bergeron and Reutenauer introduced the notion of "pure and inverting" compositions B_n with at most n parts. These compositions have the favorable property of being n-stable in that $\mathsf{B}_n \subseteq \mathsf{B}_{n+1}$ and that $\mathsf{B}_{n+1} \setminus \mathsf{B}_n$ are the pure and inverting compositions with exactly

$n+1$ parts. They were able to show that the pure and inverting "quasi-monomials" M_β (see Section 2) span $QSym_n/J_n$ and that they are $n!$ in number. However, the linear independence of these polynomials over Sym_n remained open. Their final conjecture, which we prove in Corollary 6, is as follows:

(Q3) The set of quasi-monomials $\{M_\beta : \beta \in \mathsf{B}_n\}$ is a basis for $QSym_n/J_n$.

The balance of this paper is organized as follows. In Section 2, we recount the details surrounding a new basis $\{\mathcal{S}_\alpha\}$ for $QSym_n$ called the quasisymmetric Schur polynomials. These behave particularly well with respect to the Sym_n action in the Schur basis. In Section 3, we give further details surrounding the "coinvariant space" $QSym_n/J_n$. These include a bijection between compositions α and pairs (λ, β), with λ a partition and β a pure and inverting composition, that informs our main results. Section 4 contains these results—a proof of (Q3), but with the quasi-monomials M_β replaced by the quasisymmetric Schur polynomials $\mathcal{S}_\beta$. We conclude in Section 5 with some corollaries to the proof. These include (Q3), as originally stated, as well as a version of (Q1) and (Q3) over the integers.

2 Quasisymmetric polynomials

A polynomial in n variables $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ is said to be *quasisymmetric* if and only if for each composition $(\alpha_1, \alpha_2, \ldots, \alpha_k)$, the monomial $x_1^{a_1} x_2^{a_2} \cdots x_k^{a_k}$ has the same coefficient as $x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_k}^{\alpha_k}$ for all sequences $1 \le i_1 < i_2 < \cdots < i_k \le n$. For example, $x_1^2 x_2 + x_1^2 x_3 + x_2^2 x_3$ is a quasisymmetric polynomial in the variables $\{x_1, x_2, x_3\}$. The ring of quasisymmetric polynomials in n variables is denoted $QSym_n$. (Note that every symmetric polynomial is quasisymmetric.)

It is easy to see that $QSym_n$ has a vector space basis given by the quasi-monomials

$$M_\alpha(\mathbf{x}) = \sum_{i_1 < \cdots < i_k} x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k},$$

for $\alpha = (\alpha_1, \ldots, \alpha_k)$ running over all compositions with at most n parts. It is less evident that $QSym_n$ is a ring, but see [10] for a formula for the product of two quasi-monomials. We write $\boldsymbol{l}(\alpha) = k$ for the *length* (number of parts) of α in what follows. We return to the quasi-monomial basis in Section 5, but for the majority of the paper, we focus on the basis of "quasisymmetric Schur polynomials" as its known multiplicative properties assist in our proofs.

2.1 The basis of quasisymmetric Schur polynomials

A *quasisymmetric Schur polynomial* $\mathcal{S}_\alpha$ is defined combinatorially through fillings of composition diagrams. Given a composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, its associated diagram is constructed by placing α_i boxes, or *cells*, in the i^{th} row from the top. (See Figure 1.) The cells are labeled using matrix notation; that is, the cell in the j^{th} column of the i^{th} row of the diagram is denoted (i, j). We abuse notation by writing α to refer to the diagram for α.

Given a composition diagram $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with largest part m, we define a *composition tableau* T of *shape* α to be a filling of the cells (i, j) of α with positive integers $T(i, j)$ such that

(CT1) entries in the rows of T weakly decrease when read from left to right,

(CT2) entries in the leftmost column of T strictly increase when read from top to bottom,

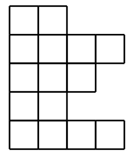


FIG. 1: The diagram associated to the composition $(2, 4, 3, 2, 4)$

(CT3) entries satisfy the *triple rule:*

Let (i, k) and (j, k) be two cells in the same column so that $i < j$. If $\alpha_i \geq \alpha_j$ then either $T(j, k) < T(i, k)$ or $T(i, k-1) < T(j, k)$. If $\alpha_i < \alpha_j$ then either $T(j, k) < T(i, k)$ or $T(i, k) < T(j, k+1)$.

Assign a *weight,* x^T to each composition tableau T by letting a_i be the number of times i appears in T and setting $x^T = \prod x_i^{a_i}$. The quasisymmetric Schur polynomial $\mathcal{S}_\alpha$ corresponding to the composition α is defined by

$$\mathcal{S}_\alpha(\mathbf{x}_n) = \sum_T x^T,$$

the sum being taken over all composition tableaux T of shape α with entries chosen from $[n]$. (See Figure 2.) Each polynomial $\mathcal{S}_\alpha$ is quasisymmetric and the collection $\{\mathcal{S}_\alpha : \boldsymbol{l}(\alpha) \leq n\}$ forms a basis for $QSym_n$ [8].

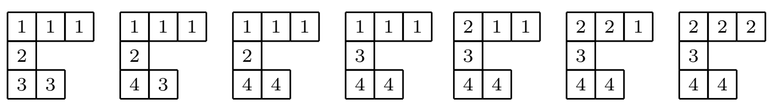


FIG. 2: The composition tableaux encoded in the polynomial $\mathcal{S}_{(3,1,2)}(\mathbf{x}_4) = x_1^3x_2x_3^2 + x_1^3x_2x_3x_4 + x_1^3x_2x_4^2 + x_1^3x_3x_4^2 + x_1^2x_2x_3x_4^2 + x_1x_2^2x_3x_4^2 + x_2^3x_3x_4^2$.

2.2 Sym action in the Quasisymmetric Schur polynomial basis

We need several definitions in order to describe the multiplication rule for quasisymmetric Schur polynomials found in [9]. The reverse of a partition λ is the composition λ^* obtained by reversing the order of its parts. Symbolically, if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ then $\lambda^* = (\lambda_k, \ldots, \lambda_2, \lambda_1)$. Let β be a composition, let λ be a partition, and let α be a composition obtained by adding $|\lambda|$ cells to β, possibly between adjacent rows of β. A filling of the cells of α is called a *Littlewood–Richardson composition tableau* of shape α/β if it satisfies the following rules:

(LR1) The i^{th} row from the bottom of β is filled with the entries $k + i$.

(LR2) The content of the appended cells is λ^*.

(LR3) The filling satisfies conditions (CT1) and (CT3) from Section 2.1.

(LR4) The entries in the appended cells, when read from top to bottom, column by column, from right to left, form a *reverse lattice word.* That is, one for which each prefix contains at least as many i's as $(i-1)$'s for each $1 < i \leq k$.

The following theorem provides a method for multiplying an arbitrary quasisymmetric Schur polynomial by an arbitrary Schur polynomial.

Theorem 1 ([9]) *In the expansion*

$$s_\lambda(\mathbf{x}) \cdot \mathcal{S}_\alpha(\mathbf{x}) = \sum_\gamma C^\gamma_{\lambda\,\alpha}\, \mathcal{S}_\gamma(\mathbf{x}), \tag{1}$$

the coefficient $C^\gamma_{\lambda\,\alpha}$ is the number of Littlewood–Richardson composition tableaux of shape γ/α with content λ^*.

3 The coinvariant space for quasisymmetric polynomials

Let $B \subseteq A$ be two $\mathbb{Q}$-algebras with A a free left module over B. This implies the existence of a subset $C \subseteq A$ with $A \simeq B \otimes C$ as vector spaces over $\mathbb{Q}$. In the classical setting of invariant theory (where B is the subring of invariants for some group action on A), this set C is identified as coset representatives for the quotient $A/(B_+)$, where (B_+) is the ideal in A generated by the positive part of the graded algebra $B = \bigoplus_{k\geq 0} B_k$.

Now suppose that A and B are graded rings. If A is free over B, then the Hilbert series of C is given as the quotient $H_q\big(A\big)\big/\, H_q\big(B\big)$. Let us try this with the choice $A = QSym_n$ and $B = Sym_n$. It is well-known that the Hilbert series for $QSym_n$ and Sym_n are given by

$$H_q\big(QSym_n\big) = 1 + \frac{q}{1-q} + \cdots + \frac{q^n}{(1-q)^n} \tag{2}$$

and

$$H_q\big(Sym_n\big) = \prod_{i=1}^{n} \frac{1}{1-q^i}\,. \tag{3}$$

Let $P_n(q) = \sum_{k\geq 0} p_k\, q^k$ denote the quotient of (2) by (3). It is easy to see that

$$P_n(q) = \prod_{i=1}^{n-1} \big(1 + q + \cdots + q^i\big) \sum_{i=0}^{n} q^i (1-q)^{n-i}\,,$$

and hence $P_n(1) = n!$. It is only slightly more difficult (see (0.13) in [6]) to show that $P_n(q)$ satisfies the recurrence relation

$$P_n(q) = P_{n-1}(q) + q^n\big([n]_q! - P_{n-1}(q)\big), \tag{4}$$

where $[n]_q!$ is the standard q-version of $n!$. Bergeron and Reutenauer use this recurrence to show that p_k is a nonnegative integer for all $k \geq 0$ and to produce a set of compositions B_n satisfying $p_k = \#\{\beta \in \mathsf{B}_n : |\beta| = k\}$ for all n. In particular, $|\mathsf{B}_n| = n!$.

Let J_n be the ideal in $QSym_n$ generated by all symmetric polynomials with zero constant term and call $R_n := QSym_n/J_n$ the *coinvariant space for quasisymmetric polynomials*. From the above discussion, R_n has dimension at most $n!$. If the set of quasi-monomials $\{M_\beta \in QSym_n : \beta \in \mathsf{B}_n\}$ are linearly independent over Sym_n, then it has dimension exactly $n!$ and $QSym_n$ becomes a free Sym_n module of the same dimension.

3.1 Destandardization of permutations

To produce a set B_n of compositions indexing a proposed basis of R_n, first recognize the $[n]_q!$ in (4) as the Hilbert series for the classical coinvariant space $\mathbb{Q}[\mathbf{x}]/(\mathcal{E}_n)$ from (S3). The standard set of compositions indexing this space are the *Artin monomials* $\{x_1^{\alpha_1}\cdots x_n^{\alpha_n} : 0 \leq \alpha_i \leq n-i\}$, but these do not fit into the desired recurrence (4) with n-stability. In [5], Garsia developed an alternative set of monomials indexed by permutations. His "descent monomials" (actually, the "reversed" descent monomials, see [6, §6]) were chosen as the starting point for the recursive construction of the sets B_n. Here we give a description in terms of "destandardized permutations."

In what follows, we view partitions and compositions as words in the alphabet $\mathbb{N} = \{0,1,2,\dots\}$. For example, we write 2543 for the composition $(2,5,4,3)$. The *standardization* $\mathrm{st}(w)$ of a word w of length k is a permutation in $\mathfrak{S}_k$ obtained by first replacing (from left to right) the ℓ_1 1s in w with the numbers $1,\dots,\ell_1$, then replacing (from left to right) the ℓ_2 2s in w with the numbers $\ell_1{+}1,\dots,\ell_1{+}\ell_2$, and so on. For example, $\mathrm{st}(121) = 132$ and $\mathrm{st}(2543) = 1432$. The *destandardization* $\mathbf{d}(\sigma)$ of a permutation $\sigma \in \mathfrak{S}_k$ is the lexicographically least word $w \in \mathbb{P}^k$ satisfying $\mathrm{st}(w) = \sigma$. For example, $\mathbf{d}(132) = 121$ and $\mathbf{d}(1432) = 1321$. Let $\mathsf{D}_{(n)}$ denote the compositions $\{\mathbf{d}(\sigma) : \sigma \in \mathfrak{S}_n\}$. Finally, given $\mathbf{d}(\sigma) = (\alpha_1,\dots,\alpha_k)$, let $\mathbf{r}(\sigma)$ denote the vector difference $(\alpha_1,\dots,\alpha_k) - 1^k$ (leaving in place any zeros created in the process). For example, $\mathbf{r}(132) = 010$ and $\mathbf{r}(1432) = 0210$. Up to a relabelling, the weak compositions $\mathbf{r}(\sigma)$ are the ones introduced by Garsia in [5].

Bergeron and Reutenauer define their sets B_n recursively in such a way that

- $\mathsf{B}_0 := \{0\}$,
- $1^n{+}\mathsf{B}_{n-1} \subseteq \mathsf{D}_{(n)}$ and $\mathsf{D}_{(n)}$ is disjoint from B_{n-1}, and
- $\mathsf{B}_n := \mathsf{B}_{n-1} \cup \mathsf{D}_{(n)} \setminus \big(1^n{+}\mathsf{B}_{n-1}\big)$.

Here, $1^n{+}\mathsf{B}_{n-1}$ denotes the vector sums $\{1^n + \mathbf{d} : \mathbf{d} \in \mathsf{B}_{n-1}\}$. Note that the compositions in $\mathsf{D}_{(n)}$ all have length n. Moreover, $1^{n+1}{+}\mathsf{D}_{(n)} \subseteq \mathsf{D}_{(n+1)}$. Indeed, if $\sigma = \sigma'1$ is a permutation in $\mathfrak{S}_{n+1}$ with suffix "1" in one-line notation, then $1^{n+1} + \mathbf{d}(\mathrm{st}(\sigma')) = \mathbf{d}(\sigma)$. That (4) enumerates B_n is immediate [6, Proposition 6.1]. We give the first few sets B_n and $\mathsf{D}_{(n)}$ in Figure 3.

$$
\begin{array}{ll}
\mathsf{D}_{(1)} = \{\underline{1}\} & \mathsf{B}_0 = \{0\} \\
\mathsf{D}_{(2)} = \{\underline{11}, 21\} & \mathsf{B}_1 = \{0\} \\
\mathsf{D}_{(3)} = \{\underline{111}, 211, 121, 221, 212, \underline{321}\} & \mathsf{B}_2 = \{0, 21\} \\
\mathsf{D}_{(4)} = \{\underline{1111}, 2111, 1211, 1121, 2211, 2121, 1221, 2112, 1212, 2221, 2212, 2122, & \mathsf{B}_3 = \{0, 21, 211, \\
\qquad\quad \underline{3211}, 3121, 1321, \underline{3221}, \underline{2321}, 3212, 2312, 2132, \underline{3321}, \underline{3231}, 3213, 4321\} & \qquad\quad 121, 221, 212\}
\end{array}
$$

FIG. 3: The sets $\mathsf{D}_{(n)}$ and B_n for small values of n. Compositions $1^n{+}\mathsf{B}_{n-1}$ are underlined in D_n.

3.2 Pure and inverting compositions

We now give an alternative description of the compositions in B_n that will be easier to work with in what follows. Call a composition α *inverting* if and only if for each $i > 1$ there exists a pair of indices $s < t$

such that $\alpha_s = i$ and $\alpha_t = i-1$. For example, 13112312 is inverting while 21123113 is not. Any composition α admits a unique factorization

$$\alpha = \gamma k^{i_k} \cdots 2^{i_2} 1^{i_1}, \quad (i_j \geq 1)$$

such that γ is a composition that does not contain any of the values from 1 to k, and k is maximal (but possibly zero). We say α is *pure* if and only if this maximal k is even. (Note that if the last part of a composition is not 1, then $k = 0$ and the composition is pure.) For example, 5435211 is pure with $k = 2$ while 3231 is impure since $k = 1$.

Proposition 2 (Bergeron and Reutenauer) *The set of inverting compositions of length n is precisely $\mathsf{D}_{(n)}$. The set of pure and inverting compositions of length at most n is precisely B_n.* □

Let $\mathsf{C}_{n,d}$ be the set of all compositions of d into at most n parts and set $\mathsf{PB}_{n,d} := \big\{(\lambda,\beta) : \lambda$ a partition, $\beta \in \mathsf{B}_n, |\lambda| + |\beta| = d,$ and $\boldsymbol{l}(\lambda) \leq n, \boldsymbol{l}(\beta) \leq n\big\}$. We define a map $\phi\colon \mathsf{PB}_{n,d} \to \mathsf{C}_{n,d}$ as follows. Given an arbitrary element (λ,β) be of $\mathsf{PB}_{n,d}$, $\phi((\lambda,\beta))$ is the composition obtained by adding λ_i to the i^{th} largest part of β for each $1 \leq i \leq \boldsymbol{l}(\lambda)$, where if $\beta_j = \beta_k$ and $j < k$, then β_j is considered smaller than β_k. If $\boldsymbol{l}(\lambda) > \boldsymbol{l}(\beta)$, append zeros after the last part to lengthen β before applying ϕ. (See Figure 4.)

$$\begin{array}{rcl} \lambda & = & 1\ 4\ 2\ 1\ 1\ 4\ 5\ 2\ 4\ 1\ 1 \\ \beta & = & 2\ 4\ 3\ 1\ 1\ 3\ 4\ 2\ 3 \\ \hline \phi(\lambda,\beta) & = & 3\ 8\ 5\ 2\ 2\ 7\ 9\ 4\ 7\ 1\ 1 \end{array}$$

FIG. 4: An example of the map $\phi\colon \mathsf{PB}_{13,49} \to \mathsf{C}_{13,49}$.

Proposition 3 *The map ϕ is a bijection between $\mathsf{PB}_{n,d}$ and $\mathsf{C}_{n,d}$*

Proof: We prove this by describing the inverse ϕ^{-1} algorithmically. Let α be an arbitrary composition in $\mathsf{C}_{n,d}$ and set $(\lambda,\beta) := (\emptyset,\alpha)$.

1. If β is pure and inverting, then $\phi^{-1}(\alpha) := (\lambda,\beta)$
2. If β is impure and inverting, then set $\phi^{-1}(\alpha) := (\lambda + (1^n), \beta - (1^n))$.
3. If β is not inverting, then let j be the smallest part of β such that there does not exist a pair of indices $s < t$ such that $\beta_s = j$ and $\beta_t = j-1$. Let m be the number of parts of β which are greater than or equal to j. Replace β by the composition obtained by subtracting 1 from each part greater than or equal to j and replace λ by the partition obtained by adding 1 to each of the first m parts.
4. Repeat until ϕ^{-1} is obtained. That is, until Step (1) or (2) above is followed.

To see that $\phi\phi^{-1} = \mathbb{1}$, consider an arbitrary composition α. If α is pure and inverting, then $\phi\phi^{-1}(\alpha) = \phi(\emptyset,\alpha) = \alpha$. If α is impure and inverting, then $\phi(\phi^{-1}(\alpha)) = \phi((1^{\boldsymbol{l}(\alpha)}, \alpha - (1^{\boldsymbol{l}(\alpha)}))) = \alpha$. Consider therefore a composition α which is not inverting. Note that the largest entry in α is decreased at each iteration of the procedure. Therefore the largest entry in the partition records the number of times the largest entry in α is decreased. Similarly, for each $i \leq \boldsymbol{l}(\lambda)$, the i^{th} largest entry in α is decreased by one

λ_i times. This means that the i^{th} largest part of α is obtained by adding λ_i to the i^{th} largest part of β and therefore our procedure ϕ^{-1} inverts the map ϕ. $\Box$

Figure 5 illustrates the algorithmic description of ϕ^{-1} as introduced in the proof of Proposition 3 on $\alpha = 38522794711$.

$$\begin{array}{cccccc}
\alpha \mapsto \binom{\lambda}{\beta}: & \begin{array}{c} \emptyset \\ 3\,\underline{8}\,\underline{5}\,2\,2\,\underline{7}\,\underline{9}\,\underline{\underline{4}}\,\underline{7}\,1\,1 \end{array} & \rightarrow & \begin{array}{c} 1\,1 \quad\; 1\,1\,1\,1 \\ 3\,\underline{7}\,4\,2\,2\,\underline{\underline{6}}\,\underline{8}\,3\,\underline{\underline{6}}\,1\,1 \end{array} & & \\
 & & & \downarrow & & \\
 & \begin{array}{c} 3\,1 \quad\; 3\,3\,1\,3 \\ 3\,5\,4\,2\,2\,4\,\underline{\underline{6}}\,3\,4\,1\,1 \end{array} & \leftarrow & \begin{array}{c} 2\,1 \quad\; 2\,2\,1\,2 \\ 3\,\underline{6}\,4\,2\,2\,\underline{\underline{5}}\,\underline{7}\,3\,\underline{\underline{5}}\,1\,1 \end{array} & & \\
 & \downarrow & & & & \\
 & \begin{array}{c} 3\,1 \quad\; 3\,4\,1\,3 \\ 3\,5\,4\,2\,2\,4\,5\,3\,4\,1\,1 \end{array} & \rightarrow & \begin{array}{c} 1\,4\,2\,1\,1\,4\,5\,2\,4\,1\,1 \\ 2\,4\,3\,1\,1\,3\,4\,2\,3 \end{array} & \rightarrow & \begin{pmatrix} 14211452411 \\ 243113423 \end{pmatrix}.
\end{array}$$

FIG. 5: The map $\phi^{-1}\colon \mathsf{C}_{13,49} \to \mathsf{PB}_{13,49}$ applied to $\alpha = 38522794711$. Parts j from Step 3 of the algorithm are marked with a double underscore.

4 Main Results

Let B_n be as in Section 3 and set $\mathcal{B}_n := \{\mathcal{S}_\beta : \beta \in \mathsf{B}_n\}$. We prove the following.

Theorem 4 *The set $\mathcal{B}_n$ is a basis for the Sym_n-module R_n.*

To prove this, we focus on the quasisymmetric polynomials $QSym_{n,d}$ in n variables of homogeneous degree d. Note that $QSym_n = \bigoplus_{d\geq 0} QSym_{n,d}$ and therefore if $\mathfrak{C}_{n,d}$ is a basis for $QSym_{n,d}$, then the collection $\bigcup_{d\geq 0} \mathfrak{C}_{n,d}$ is a basis for $QSym_n$.

First, we introduce a useful term order. Each composition α can be rearranged to form a partition $\boldsymbol{\lambda}(\alpha)$ by arranging the parts in weakly decreasing order. Recall the *lexicographic* order $\succeq$ on partitions of n, which states that $\lambda \succeq \mu$ if and only if the first nonzero entry in $\lambda - \mu$ is positive. For two compositions α and γ of n, we say that α is larger then γ in *revlex* order (written $\alpha \succ \gamma$) if and only if either

- $\boldsymbol{\lambda}(\alpha) \succeq \boldsymbol{\lambda}(\gamma)$, or
- $\boldsymbol{\lambda}(\alpha) = \boldsymbol{\lambda}(\gamma)$ and α is lexicographically larger than γ when reading right to left.

For instance, we have

$$4 \succeq 13 \succeq 31 \succeq 22 \succeq 112 \succeq 121 \succeq 211 \succeq 1111.$$

Remark: Extend revlex to weak compositions of length at most n by padding the beginning of α or γ with zeros as necessary, so $\boldsymbol{l}(\alpha) = \boldsymbol{l}(\gamma) = n$. Viewing these as exponent vectors for monomials in $\mathbf{x}$ provides a term ordering on $\mathbb{Q}[\mathbf{x}]$. However, it is not good term ordering in the sense that it is not multiplicative: given exponent vectors α, β, and γ with $\alpha \succeq \gamma$, it is not necessarily the case that $\alpha + \beta \succeq \gamma + \beta$. This is no doubt the trouble encountered in [3] and [6] when trying to prove the Bergeron–Reutenauer conjecture (Q3). We circumvent this difficulty by working with the Schur polynomials s_λ and the quasisymmetric

Schur polynomials $\mathcal{S}_\alpha$. We consider leading polynomials $\mathcal{S}_\gamma$ instead of leading monomials x^γ. The leading term $\mathcal{S}_\gamma$ in a product $s_\lambda \cdot \mathcal{S}_\alpha$ is readily found.

Proof of Theorem 4: We claim that the collection $\mathfrak{C}_{n,d} = \{s_\lambda \mathcal{S}_\beta : |\lambda| + |\beta| = d, \boldsymbol{l}(\lambda) \leq n, \boldsymbol{l}(\beta) \leq n,$ and $\beta \in \mathsf{B}_n\}$ is a basis for $QSym_{n,d}$, which in turn implies that $\mathcal{B}_n$ is a basis for R_n.

We define an ordering on $\mathfrak{C}_{n,d}$ by using the map ϕ and the revlex ordering on compositions. Note that $\mathfrak{C}_{n,d}$ is indexed by pairs (λ, β), where λ is a partition of some $k \leq d$ and β is a composition of $d-k$ which lies in B_n. We claim that the leading term in the quasisymmetric Schur polynomial expansion of $s_\lambda \mathcal{S}_\beta$ is the polynomial $S_{\phi(\lambda,\beta)}$. To see this, recall that the terms of $s_\lambda \mathcal{S}_\beta$ are given by Littlewood–Richardson composition tableaux of shape α/β and content λ^*, where α is an arbitrary composition shape obtained by appending $|\lambda|$ cells to the diagram of β so that conditions (CT1) and (CT3) are satisfied.

To form the largest possible composition (in revlex order), one first appends as many cells as possible to the longest row of β. (Again, the lower of two equal rows is considered longer.) This new longest row must end in an $L := \boldsymbol{l}(\lambda)$, since the reading word of the Littlewood–Richardson composition tableau must satisfy (LR4). No entry smaller than L can appear to the left of L in this row, since the row entries are weakly decreasing from left to right. Thus the maximum possible number of entries that could be added to the longest row of β is λ_1. Similarly, the maximum possible number of entries that can be added to the second longest row of β is λ_2 and so on. If $\boldsymbol{l}(\lambda) > \boldsymbol{l}(\beta)$, append the extra parts of λ to end of β.

We must show that this largest possible shape is indeed the shape of a Littlewood–Richardson composition tableau obtained by adding cells with content λ^* to the diagram of β. Place the entries of λ^* into the new cells as described above, with the following exception. If two augmented rows have the same length and the corresponding parts of λ^* are equal, then place the larger entries into the higher of the two rows.

Such a filling T of shape α satisfies (CT1) by construction. To prove that the filling $S(\lambda, \beta)$ satisfies (CT3), consider an arbitrary pair of cells (i,k) and (j,k) in the same column. If $\alpha_i \geq \alpha_j$ then $\beta_i \geq \beta_j$, since the entries from λ are appended to the rows of β from largest row to smallest row. Therefore if (i,k) is a cell in the diagram of β then $T(j,k) < T(i,k) = T(i,k-1)$ regardless of whether or not (j,k) is in the diagram of β. If (i,k) is not in the diagram of β then (j,k) cannot be in the diagram of β since $\beta_i \geq \beta_j$. Therefore $T(j,k) < T(i,k)$ since the smaller entry is placed into the shorter row, or the lower row if the rows have equal length.

If $\alpha_i < \alpha_j$ then $\beta_i \leq \beta_j$. If $T(i,k) \leq T(j,k)$ then (i,k) is not in the diagram of β. If $(j,k+1)$ is in the diagram of β then $T(i,k) < T(j,k+1)$ since the entries in the diagram of β are larger than the appended entries. Otherwise the cell $(j,k+1)$ is filled with a larger entry than (i,k) since the longer rows are filled with larger entries and $\alpha_j > \alpha_i$. Therefore the entries in $S(\lambda, \beta)$ satisfy (CT3).

To see that the reading word of such a filling satisfies (LR4), consider an entry i. We must show that an arbitrary prefix of the reading word contains at least as many i's as $(i-1)$'s. Let c_i be the rightmost column of T containing the letter i and let c_{i-1} be the rightmost column of T containing the letter $i-1$. If $c_i > c_{i-1}$ then every prefix will contain at least as many $i's$ as $(i-1)$'s since there will always be at least one i appearing before any pairs $i, i-1$ in reading order. If $c_i = c_{i-1}$, then the entry i will appear in a higher row than the entry $i-1$ and hence will be read first for each column containing both an i and an $i-1$. Therefore the reading word is a reverse lattice word and hence the filling constructed above is indeed a Littlewood–Richardson composition tableau. The shape of this Littlewood–Richardson composition tableau corresponds to the largest composition appearing as an index of a quasisymmetric Schur polynomial in the expansion of $s_\lambda \mathcal{S}_\beta$, so $\mathcal{S}_{\phi(\lambda,\beta)}$ is indeed the leading term in this expansion. Since ϕ is a bijection, the entries in $\mathfrak{C}_{n,d}$ span $QSym_{n,d}$ and are linearly independent. Therefore $\mathfrak{C}_{n,d}$ is a basis for $QSym_{n,d}$ and hence $\mathcal{B}_n$ is a basis for the Sym_n-module R_n. $\square$

The transition matrix between the basis $\mathfrak{C}_{3,4}$ and the quasisymmetric Schur polynomial basis for $QSym_{3,4}$ is given in Table 1. Note that this is an upper unitriangular matrix. In fact, this is true in general and therefore the algebra $QSym_n(\mathbb{Z})$ is a free module over $Sym_n(\mathbb{Z})$. A basis is given by $\{s_\lambda \mathcal{S}_\beta : \beta \in \Pi_n, \boldsymbol{l}(\lambda) \leq n, \text{ and } \boldsymbol{l}(\beta) \leq n\}$. Replacing $\mathcal{S}_\beta$ by M_β results in an alternative basis.

$$
\begin{array}{r c}
 & \begin{array}{ccccccc} 4 & 13 & 31 & 22 & 112 & 121 & 211 \end{array} \\
\begin{array}{r} s_4 \\ s_{31} \\ s_1 \cdot \mathcal{S}_{21} \\ s_{22} \\ s_{211} \\ \mathcal{S}_{121} \\ \mathcal{S}_{211} \end{array} &
\begin{pmatrix}
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 1 & 1 & \cdot & \cdot & 1 \\
\cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1
\end{pmatrix}
\end{array}
$$

TAB. 1: The transition matrix for $n = 3, d = 4$.

5 Corollaries and applications

5.1 Closing the Bergeron–Reuteuaner conjecture

In [8, Theorem 6.1], it was shown that the polynomials M_γ are related to the polynomials $\mathcal{S}_\alpha$ as follows:

$$\mathcal{S}_\alpha = \sum_\gamma K_{\alpha,\gamma} \, M_\gamma \,, \tag{5}$$

where $K_{\alpha,\gamma}$ counts the number of composition tableaux T of shape α and content γ.

Lemma 5 *In the expansion $M_\alpha = \sum_\gamma \tilde{K}_{\alpha,\gamma} \, \mathcal{S}_\gamma$, if $\tilde{K}_{\alpha,\gamma} \neq 0$ then $\alpha \succeq \gamma$. Also, $\tilde{K}_{\alpha,\alpha} = 1$.*

Proof: Theorem 6.1 and Proposition 6.7 in [8] reveal that the transition matrix $K = (K_{\alpha,\gamma})$ defined in (5) is upper-unitriangular with respect to $\succeq$. That is, $\alpha \prec \gamma$ implies $K_{\alpha,\gamma} = 0$ and $K_{\alpha,\alpha} = 1$. A closer analysis shows that, moreover, $K_{\alpha,\gamma} = 0$ when $\alpha \neq \gamma$ but $\boldsymbol{\lambda}(\alpha) = \boldsymbol{\lambda}(\gamma)$. The desired result for $\tilde{K} = K^{-1}$ follows readily from this fact. $\square$

We are ready to prove Conjecture (Q3). Let B_n and R_n be as in Section 4.

Corollary 6 *The set $\{M_\beta : \beta \in \mathsf{B}_n\}$ is a basis for the Sym_n-module R_n.*

Proof: We show that the collection $\mathfrak{M}_{n,d} = \{s_\lambda M_\beta : |\lambda| + |\beta| = d, \boldsymbol{l}(\lambda) \leq n, \boldsymbol{l}(\beta) \leq n, \text{ and } \beta \in \mathsf{B}_n\}$ is a basis for $QSym_{n,d}$, which in turn implies that $\{M_\beta : \beta \in \mathsf{B}_n\}$ is a basis for R_n. We first claim that the leading term in the quasisymmetric Schur polynomial expansion of $s_\lambda M_\beta$ is indexed by the composition $\phi(\lambda, \beta)$. The corollary will easily follow.

Applying Lemma 5, we may write $s_\lambda M_\beta$ as

$$s_\lambda M_\beta = s_\lambda \mathcal{S}_\beta + \sum_{\beta \succ \gamma} \tilde{K}_{\beta,\gamma}\, s_\lambda \mathcal{S}_\gamma\,.$$

Given any composition γ, the leading term of $s_\lambda S_\gamma$ is indexed by $\phi(\lambda,\gamma)$. This follows by the reasoning used in the proof of Theorem 4. We prove the claim by showing that $\beta \succ \gamma \implies \phi(\lambda,\beta) \succ \phi(\lambda,\gamma)$.

Assume first that $\boldsymbol{\lambda}(\beta) = \boldsymbol{\lambda}(\gamma)$. Let i be the greatest integer such that $\beta_i > \gamma_i$. The map ϕ adds λ_j cells to β_i and λ_k cells to γ_i, where $\lambda_j \geq \lambda_k$. Therefore $\beta_i + \lambda_j > \gamma_i + \lambda_k$. Since the parts of $\phi(\lambda,\beta)$ and $\phi(\lambda,\gamma)$ are equal after part i, we have $\phi(\lambda,\beta) \succeq \phi(\lambda,\gamma)$.

Next assume that $\boldsymbol{\lambda}(\beta) \succ \boldsymbol{\lambda}(\gamma)$. Consider the smallest i such that the i^{th} largest part β_j of β is not equal to the i^{th} largest part γ_k of γ. The map ϕ adds λ_i cells to β_j and to γ_k, so that $\beta_j + \lambda_i > \gamma_k + \lambda_i$. Since the largest $i-1$ parts of $\phi(\lambda,\beta)$ and $\phi(\lambda,\gamma)$ are equal, we have $\boldsymbol{\lambda}(\phi(\lambda,\beta)) \succ \boldsymbol{\lambda}(\phi(\lambda,\gamma))$.

We now use the claim to complete the proof. Following the proof of Theorem 4, we arrange the products $s_\lambda M_\beta$ as row vectors written in the basis of quasisymmetric Schur polynomials. The claim shows that the corresponding matrix is upper-unitriangular. Thus $\mathfrak{M}_{n,d}$ forms a basis for $QSym_{n,d}$, as desired. □

5.2 Triangularity

It was shown in Section 4 that the transition matrix between the bases $\mathfrak{C}$ and $\{\mathcal{S}_\alpha\}$ is triangular with respect to the revlex ordering. Here, we show that a stronger condition holds: it is triangular with respect to a natural partial ordering on compositions. Every composition α has a corresponding partition $\boldsymbol{\lambda}(\alpha)$ obtained by arranging the parts of α in weakly decreasing order. A partition λ is said to *dominate* a partition μ iff $\sum_{i=1}^k \lambda_i \geq \sum_{i=1}^k \mu_i$ for all k. Let $C^\alpha_{\lambda,\beta}$ be the coefficient of $\mathcal{S}_\alpha$ in the expansion of the product $s_\lambda\, \mathcal{S}_\beta$.

Theorem 7 *If $\boldsymbol{\lambda}(\alpha)$ is not dominated by $\boldsymbol{\lambda}(\phi(\lambda,\beta))$, then $C^\alpha_{\lambda,\beta} = 0$.*

Proof: Let (λ,β) be an arbitrary element of $\mathsf{PB}_{n,d}$ and let α be an arbitrary element of $\mathsf{C}_{n,d}$. Set $\gamma := \phi(\lambda,\beta)$. If $\alpha \succeq \gamma$ then $C^\alpha_{\lambda,\beta} = 0$ (by the proof of Theorem 4) and we are done. Hence, assume that $\alpha \not\succeq \phi(\lambda,\beta) = \gamma$ and that $\boldsymbol{\lambda}(\alpha)$ is not dominated by $\boldsymbol{\lambda}(\gamma)$. Let k be the smallest positive integer such that $\sum_{i=1}^k \boldsymbol{\lambda}(\alpha)_i > \sum_{i=1}^k \boldsymbol{\lambda}(\gamma)_i$. (Such an integer exists since $\boldsymbol{\lambda}(\alpha)$ is not dominated by $\boldsymbol{\lambda}(\gamma)$.) Therefore $\sum_{i=1}^k \boldsymbol{\lambda}(\alpha)_i - \sum_{i=1}^k \boldsymbol{\lambda}(\beta)_i > \sum_{i=1}^k \boldsymbol{\lambda}(\gamma)_i - \sum_{i=1}^k \boldsymbol{\lambda}(\beta)_i$ and there are more entries in the longest k rows of α/β then there are in the longest k rows of γ/β. This implies that there are more than $\sum_{i=1}^k \lambda_i$ entries from α/β contained in the longest k rows of α, since there are $\sum_{i=1}^k \lambda_i$ entries in the longest k rows of γ/β. This implies that in a Littlewood–Richardson composition tableau of shape α/β, the longest k rows must contain an entry less than $L-k+1$ where $L = \boldsymbol{l}(\lambda)$.

The rightmost entry in the i^{th} longest row of α/β must be $L-i+1$ for otherwise the filling would not satisfy the reverse lattice condition. This means that the longest k rows of α must contain only entries greater than or equal to $L-i+1$, which contradicts the assertion that an entry less than $L-k+1$ appears among the k longest rows of α. Therefore there is no such Littlewood–Richardson composition tableau of shape α and so $C^\alpha_{\lambda,\beta} = 0$. □

References

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

[2] F. Bergeron. *Algebraic combinatorics and coinvariant spaces*. CMS Treatises in Mathematics. Canadian Mathematical Society, Ottawa, ON, 2009.

[3] F. Bergeron and C. Reutenauer. The coinvariant space for quasisymmetric polynomials. *Unpublished manuscript*.

[4] S. Fomin and A. N. Kirillov. Combinatorial B_n-analogues of Schubert polynomials. *Trans. Amer. Math. Soc.*, 348(9):3591–3620, 1996.

[5] A. M. Garsia. Combinatorial methods in the theory of Cohen-Macaulay rings. *Adv. in Math.*, 38(3):229–266, 1980.

[6] A. M. Garsia and N. Wallach. Qsym over Sym is free. *J. Combin. Theory Ser. A*, 104(2):217–263, 2003.

[7] I. M. Gessel. Multipartite P-partitions and inner products of skew Schur functions. In *Combinatorics and algebra (Boulder, Colo., 1983)*, volume 34 of *Contemp. Math.*, pages 289–317. Amer. Math. Soc., Providence, RI, 1984.

[8] J. Haglund, K. Luoto, S. K. Mason, and S. van Willigenburg. Quasisymmetric Schur functions. *J. Combin. Theory Ser. A. to appear*, arXiv:0810.2489.

[9] J. Haglund, K. Luoto, S. K. Mason, and S. van Willigenburg. Refinements of the Littlewood–Richardson rule. *Trans. Amer. Math. Soc. to appear*, arXiv:0908.3540.

[10] M. Hazewinkel. Symmetric functions, noncommutative symmetric functions, and quasisymmetric functions. *Acta Applicandae Mathematicae*, 75(1):55–83, 2003.

[11] R. Kane. *Reflection groups and invariant theory*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 5. Springer-Verlag, New York, 2001.

[12] A. Lascoux and M.-P. Schützenberger. Polynômes de Schubert. *C. R. Acad. Sci. Paris Sér. I Math.*, 294(13):447–450, 1982.

[13] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995.

[14] C. Malvenuto and C. Reutenauer. Duality between quasi-symmetric functions and the Solomon descent algebra. *J. Algebra*, 177(3):967–982, 1995.

[15] L. Manivel. *Symmetric functions, Schubert polynomials and degeneracy loci*, volume 6 of *SMF/AMS Texts and Monographs*. American Mathematical Society, Providence, RI, 2001. Translated from the 1998 French original by John R. Swallow, Cours Spécialisés [Specialized Courses], 3.

[16] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 257–267

An Algebraic Analogue of a Formula of Knuth

Lionel Levine

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

Abstract. We generalize a theorem of Knuth relating the oriented spanning trees of a directed graph G and its directed line graph $\mathcal{L}G$. The sandpile group is an abelian group associated to a directed graph, whose order is the number of oriented spanning trees rooted at a fixed vertex. In the case when G is regular of degree k, we show that the sandpile group of G is isomorphic to the quotient of the sandpile group of $\mathcal{L}G$ by its k-torsion subgroup. As a corollary we compute the sandpile groups of two families of graphs widely studied in computer science, the de Bruijn graphs and Kautz graphs.

Résumé. Nous généralisons un théorème de Knuth qui relie les arbres couvrants dirigés d'un graphe orienté G au graphe adjoint orienté $\mathcal{L}G$. On peut associer à tout graphe orienté un groupe abélien appelé groupe du tas de sable, et dont l'ordre est le nombre d'arbres couvrants dirigés enracinés en un sommet fixé. Lorsque G est régulier de degré k, nous montrons que le groupe du tas de sable de G est isomorphe au quotient du groupe du tas de sable de $\mathcal{L}G$ par son sous-groupe de k-torsion. Comme corollaire, nous déterminons les groupes de tas de sable de deux familles de graphes étudiées en informatique: les graphes de de Bruijn et les graphes de Kautz.

Keywords: critical group, de Bruijn graph, iterated line digraph, Kautz graph, matrix-tree theorem, oriented spanning tree, weighted Laplacian

---

1 Introduction

In this extended abstract we discuss some new generalizations of an enumerative formula of Knuth [10]. Proofs omitted here due to space constraints can be found in [11].

Let $G = (V, E)$ be a finite directed graph, which may have loops and multiple edges. Each edge $e \in E$ is directed from its source vertex $s(e)$ to its target vertex $t(e)$. The *directed line graph* $\mathcal{L}G = (E, E_2)$ has as vertices the edges of G, and as edges the set

$$E_2 = \{(e_1, e_2) \in E \times E \mid s(e_2) = t(e_1)\}.$$

For example, if G has just one vertex and n loops, then $\mathcal{L}G$ is the complete directed graph on n vertices. If G has two vertices and no loops, then $\mathcal{L}G$ is a bidirected complete bipartite graph.

An *oriented spanning tree* of G is a subgraph containing all of the vertices of G, having no directed cycles, in which one vertex, the *root*, has outdegree 0, and every other vertex has outdegree 1. The number $\kappa(G)$ of oriented spanning trees of G is sometimes called the *complexity* of G.

Our first result relates the numbers $\kappa(\mathcal{L}G)$ and $\kappa(G)$. Let $\{x_e\}_{e\in E}$ and $\{x_v\}_{v\in V}$ be indeterminates, and consider the polynomials

$$\kappa^{edge}(G, \mathbf{x}) = \sum_T \prod_{e \in T} x_e$$

$$\kappa^{vertex}(G, \mathbf{x}) = \sum_T \prod_{e \in T} x_{t(e)}.$$

The sums are over all oriented spanning trees T of G.

Write

$$\text{indeg}(v) = \#\{e \in E \mid t(e) = v\}$$

$$\text{outdeg}(v) = \#\{e \in E \mid s(e) = v\}$$

for the indegree and outdegree of vertex v in G. We say that v is a *source* if $\text{indeg}(v) = 0$.

Theorem 1.1 *Let $G = (V, E)$ be a finite directed graph with no sources. Then*

$$\kappa^{vertex}(\mathcal{L}G, \mathbf{x}) = \kappa^{edge}(G, \mathbf{x}) \prod_{v \in V} \left(\sum_{s(e)=v} x_e \right)^{\text{indeg}(v)-1}. \tag{1}$$

Note that since the vertex set of $\mathcal{L}G$ coincides with the edge set of G, both sides of (1) are polynomials in the same set of variables $\{x_e\}_{e \in E}$. Setting all $x_e = 1$ yields the product formula

$$\kappa(\mathcal{L}G) = \kappa(G) \prod_{v \in V} \text{outdeg}(v)^{\text{indeg}(v)-1} \tag{2}$$

due in a slightly different form to Knuth [10]. Special cases of (2) include Cayley's formula n^{n-1} for the number of rooted spanning trees of the complete graph K_n, as well as the formula $(m+n)m^{n-1}n^{m-1}$ for the number of rooted spanning trees of the complete bipartite graph $K_{m,n}$. These are respectively the cases that G has just one vertex with n loops, or G has just two vertices a and b with m edges directed from a to b and n edges directed from b to a.

Suppose now that G is strongly connected, that is, for any $v, w \in V$ there are directed paths in G from v to w and from w to v. Then associated to any vertex v_* of G is an abelian group $K(G, v_*)$, the *sandpile group*, whose order is the number of oriented spanning trees of G rooted at v_*. Its definition and basic properties are reviewed in section 3. Other common names for this group are the critical group, Picard group, Jacobian, and group of components. In the case when G is *Eulerian* (that is, $\text{indeg}(v) = \text{outdeg}(v)$ for all vertices v) the groups $K(G, v_*)$ and $K(G, v'_*)$ are isomorphic for any $v_*, v'_* \in V$, and we often denote the sandpile group just by $K(G)$.

When G is Eulerian, we show that there is a natural map from the sandpile group of $\mathcal{L}G$ to the sandpile group of G, descending from the $\mathbb{Z}$-linear map

$$\phi : \mathbb{Z}^E \to \mathbb{Z}^V$$

which sends $e \mapsto t(e)$.

Let k be a positive integer. We say that G is *balanced k-regular* if $\text{indeg}(v) = \text{outdeg}(v) = k$ for every vertex v.

Theorem 1.2 *Let $G = (V, E)$ be a strongly connected Eulerian directed graph, fix $e_* \in E$ and let $v_* = t(e_*)$. The map ϕ descends to a surjective group homomorphism*

$$\bar{\phi} : K(\mathcal{L}G, e_*) \to K(G, v_*).$$

Moreover, if G is balanced k-regular, then $\ker(\bar{\phi})$ is the k-torsion subgroup of $K(\mathcal{L}G, e_*)$.

This result extends to directed graphs some of the recent work of Berget, Manion, Maxwell, Potechin and Reiner [1] on undirected line graphs. If $G = (V, E)$ is an undirected graph, the (undirected) line graph line(G) of G has vertex set E and edge set

$$\{\{e, e'\} \,|\, e, e' \in E,\ e \cap e' \neq \emptyset\}.$$

The results of [1] relate the sandpile groups of G and line(G). The undirected case is considerably more subtle, because although there is still a natural map $K(\text{line } G) \to K(G)$ when G is regular, this map may fail to be surjective.

A particularly interesting family of directed line graphs are the *de Bruijn graphs* DB_n, defined recursively by

$$DB_n = \mathcal{L}(DB_{n-1}), \qquad n \geq 1,$$

where DB_0 is the graph with just one vertex and two loops. The 2^n vertices of DB_n can be identified with binary words $b_1 \dots b_n$ of length n; two such sequences b and b' are joined by a directed edge (b, b') if and only if $b'_i = b_{i+1}$ for all $i = 1, \dots, n-1$.

Using Theorem 1.2, we obtain the full structure of the sandpile groups of the de Bruijn graphs.

Theorem 1.3

$$K(DB_n) = \bigoplus_{j=1}^{n-1} (\mathbb{Z}/2^j\mathbb{Z})^{2^{n-1-j}}.$$

Closely related to the de Bruijn graphs are the *Kautz graphs*, defined by

$$\text{Kautz}_1 = (\{1, 2, 3\}, \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\})$$

and

$$\text{Kautz}_n = \mathcal{L}(\text{Kautz}_{n-1}), \qquad n \geq 2.$$

The Kautz graphs are useful in network design because they have close to the maximum possible number of vertices given their diameter and degree [7] and because they contain many short vertex-disjoint paths between any pair of vertices [5]. The following result gives the sandpile group of Kautz_n.

Theorem 1.4

$$K(\text{Kautz}_n) = (\mathbb{Z}/3\mathbb{Z}) \oplus (\mathbb{Z}/2^{n-1}\mathbb{Z})^2 \oplus \bigoplus_{j=1}^{n-2} (\mathbb{Z}/2^j\mathbb{Z})^{3 \cdot 2^{n-2-j}}.$$

Bidkhori and Kishore [2] have recently generalized Theorems 1.3 and 1.4 to m-regular de Bruijn and Kautz graphs.

The remainder of the paper is organized as follows. In section 2, we discuss some interesting variants and special cases of Theorem 1.1. Section 3 begins by defining the sandpile group, and moves on from there to the proof of Theorem 1.2. In section 4 we enumerate spanning trees of iterated line digraphs. Huaxiao, Fuji and Qiongxiang [9] prove that for a balanced k-regular directed graph G on N vertices,

$$\kappa(\mathcal{L}^n G) = \kappa(G) k^{(k^n-1)N}.$$

Theorem 4.1 generalizes this formula to an arbitrary directed graph G having no sources. Section 4 also contains the proofs of Theorems 1.3 and 1.4.

2 Spanning Trees

In this section we discuss a few variants and special cases of Theorem 1.1. We omit the proof due to space constraints. See [11] for a proof using the matrix-tree theorem. Very recently, Bidkhori and Kishore [2] have found a bijective proof, and used it to resolve a question of Stanley about de Bruijn sequences.

Theorem 1.1 enumerates all oriented spanning trees of $\mathcal{L}G$, while in many applications one wants to enumerate spanning trees with a fixed root. Given a vertex $v_* \in V$, let

$$\kappa^{edge}(G, v_*, \mathbf{x}) = \sum_{\text{root}(T)=v_*} \prod_{e \in T} x_e$$

and

$$\kappa^{vertex}(G, v_*, \mathbf{x}) = \sum_{\text{root}(T)=v_*} \prod_{e \in T} x_{t(e)}.$$

The following variant of Theorem 1.1 enumerates spanning trees of $\mathcal{L}G$ with a fixed root e_* in terms of spanning trees of G with root $w_* = s(e_*)$.

Theorem 2.1 *Let $G = (V, E)$ be a finite directed graph, and let $e_* = (w_*, v_*)$ be an edge of G. If* $\text{indeg}(v) \geq 1$ *for all vertices $v \in V$, and* $\text{indeg}(v_*) \geq 2$*, then*

$$\kappa^{vertex}(\mathcal{L}G, e_*, \mathbf{x}) = \kappa^{edge}(G, w_*, \mathbf{x})\, x_{e_*} \left(\sum_{s(e)=v_*} x_e \right)^{\text{indeg}(v_*)-2} \times$$
$$\times \prod_{v \neq v_*} \left(\sum_{s(e)=v} x_e \right)^{\text{indeg}(v)-1}.$$

Setting all $x_e = 1$ in Theorem 2.1 yields the enumeration

$$\kappa(\mathcal{L}G, e_*) = \frac{\kappa(G, w_*)}{\text{outdeg}(v_*)} \pi(G) \tag{3}$$

where $\kappa(G, w_*)$ is the number of oriented spanning trees of G rooted at w_*, and

$$\pi(G) = \prod_{v \in V} \text{outdeg}(v)^{\text{indeg}(v)-1}.$$

It is interesting to compare this formula to the theorem of Knuth [10], which in our notation reads

$$\kappa(\mathcal{L}G, e_*) = \left(\kappa(G, v_*) - \frac{1}{\text{outdeg}(v_*)} \sum_{\substack{t(e)=v_* \\ e \neq e_*}} \kappa(G, s(e))\right) \pi(G). \tag{4}$$

To see directly why the right sides of (3) and (4) are equal, we define a *unicycle* to be a spanning subgraph of G which contains a unique directed cycle, and in which every vertex has outdegree 1. If vertex v_* is on the unique cycle of a unicycle U, we say that U goes through v_*.

Lemma 2.2

$$\kappa^{edge}(G, v_*, \mathbf{x}) \sum_{s(e)=v_*} x_e = \sum_{t(e)=v_*} \kappa^{edge}(G, s(e), \mathbf{x})\, x_e.$$

Proof: Removing e gives a bijection from unicycles containing a fixed edge e to spanning trees rooted at $s(e)$. If U is a unicycle through v_*, then the cycle of U contains a unique edge e with $s(e) = v_*$ and a unique edge e' with $t(e') = v_*$, so both sides are equal to

$$\sum_U \prod_{e \in U} x_e$$

where the sum is over all unicycles U through v_*. □

Setting all $x_e = 1$ in Lemma 2.2 yields

$$\text{outdeg}(v_*)\kappa(G, v_*) = \sum_{t(e)=v_*} \kappa(G, s(e)).$$

Hence the factor appearing in front of $\pi(G)$ in Knuth's formula (4) is equal to $\kappa(G, w_*)/\text{outdeg}(v_*)$.

We conclude this section by discussing some interesting examples and special cases of Theorem 1.1.

- **Deletion and contraction**. Fix $e \in E$ and set $x_f = 1$ for all $f \neq e$. The coefficient of x_e^ℓ in $\kappa^{vertex}(\mathcal{L}G, \mathbf{x})$ then counts the number of oriented spanning trees T of $\mathcal{L}G$ with $\text{indeg}_T(e) = \ell$. If $v = s(e)$ has indegree k and outdegree m, then this coefficient is given by

$$\prod_{w \neq v} \text{outdeg}(w)^{\text{indeg}(w)-1} \left(\binom{k-1}{\ell} \kappa(G \setminus e)(m-1)^{k-1-\ell} + \right.$$
$$\left. + \binom{k-1}{\ell-1} \kappa(G/e)(m-1)^{k-\ell} \right).$$

  Here $G \setminus e$ and G/e are respectively the graphs resulting from deleting and contracting the edge e. (There is more than one sensible way to define contraction for directed graphs. By G/e we mean the graph obtained from G by first deleting all edges f with $s(f) = s(e)$, and then identifying the vertices $s(e)$ and $t(e)$.)

- **Complete graph**. Taking G to be the graph with one vertex and n loops, so that $\mathcal{L}G$ is the complete directed graph $\vec{K}_n$ on n vertices, we obtain the classical formula

$$\kappa^{vertex}(\vec{K}_n) = (x_1 + \ldots + x_n)^{n-1}.$$

For a generalization to forests, see [15, Theorem 5.3.4]. Note that oriented spanning trees of $\vec{K}_n$ are in bijection with rooted spanning trees of the complete undirected graph K_n, by forgetting orientation.

- **Complete bipartite graph**. Taking G to have two vertices, a and b, with m edges directed from a to b and n edges directed from b to a, we obtain

$$\begin{aligned}\kappa^{vertex}(\vec{K}_{m,n}) = (x_1 + \ldots + x_m)^{n-1}(y_1 + \ldots + y_n)^{m-1} \cdot \\ \cdot (x_1 + \ldots + x_m + y_1 + \ldots + y_n)\end{aligned}$$

where $\vec{K}_{m,n} = \mathcal{L}G$ is the bidirected complete bipartite graph on $m+n$ vertices. The variables $x_1, \ldots, x_m$ correspond to vertices in the first part, and $y_1, \ldots, y_n$ correspond to vertices in the second part. As with the complete graph, oriented spanning trees of $\vec{K}_{m,n}$ are in bijection with rooted spanning trees of the undirected complete bipartite graph $K_{m,n}$ by forgetting orientation.

- **De Bruijn graphs**. The spanning tree enumerators for the first few de Bruijn graphs are

$$\kappa^{vertex}(DB_1) = x_0 + x_1;$$

$$\kappa^{vertex}(DB_2) = (x_{00} + x_{01})(x_{10} + x_{11})(x_{01} + x_{10});$$

$$\begin{aligned}\kappa^{vertex}(DB_3) = (x_{000} + x_{001})(x_{010} + x_{011})(x_{100} + x_{101})(x_{110} + x_{111}) \times \\ \times \big(x_{011}x_{110}x_{100} + x_{010}x_{110}x_{100} + x_{110}x_{101}x_{001} + x_{110}x_{100}x_{001} + \\ + x_{100}x_{001}x_{011} + x_{101}x_{001}x_{011} + x_{001}x_{010}x_{110} + x_{001}x_{011}x_{110}\big).\end{aligned}$$

3 Sandpile Groups

Let $G = (V, E)$ be a strongly connected finite directed graph, loops and multiple edges allowed. Consider the free abelian group $\mathbb{Z}^V$ generated by the vertices of G; we think of its elements as formal linear combinations of vertices with integer coefficients. For $v \in V$ let

$$\Delta_v = \sum_{s(e)=v} (t(e) - v) \in \mathbb{Z}^V$$

where the sum is over all edges $e \in E$ such that $s(e) = v$. Fixing a vertex $v_* \in V$, let L_V be the subgroup of $\mathbb{Z}^V$ generated by v_* and $\{\Delta_v\}_{v \neq v_*}$. The *sandpile group* $K(G, v_*)$ is defined as the quotient group

$$K(G, v_*) = \mathbb{Z}^V / L_V.$$

The $V \times V$ integer matrix whose column vectors are $\{\Delta_v\}_{v \in V}$ is called the *Laplacian* of G. Its principal minor omitting the row and column corresponding to v_* counts the number $\kappa(G, v_*)$ of oriented spanning trees of G rooted at v_*. (This is the matrix-tree theorem, [15, Theorem 5.6.4].) Since this minor is also the index of L_V in $\mathbb{Z}^V$, we have

$$\#K(G, v_*) = \kappa(G, v_*).$$

Recall that G is *Eulerian* if $\text{indeg}(v) = \text{outdeg}(v)$ for every vertex v. If G is Eulerian, then the groups $K(G, v_*)$ and $K(G, v'_*)$ are isomorphic for any vertices v_* and v'_* [8, Lemma 4.12]. In this case we usually denote the sandpile group just by $K(G)$.

The sandpile group arose independently in several fields, including arithmetic geometry [12, 13], statistical physics [4] and algebraic combinatorics [3]. Often it is defined for an undirected graph G; to translate this definition into the present setting of directed graphs, replace each undirected edge by a pair of directed edges oriented in opposite directions. Sandpiles on directed graphs were first studied in [14]. For a survey of the basic properties of sandpile groups of directed graphs and their proofs, see [8].

The goal of this section is to relate the sandpile groups of an Eulerian graph G and its directed line graph $\mathcal{L}G$. To that end, let $\mathbb{Z}^E$ be the free abelian group generated by the edges of G. For $e \in E$ let

$$\Delta_e = \sum_{s(f)=t(e)} (f - e) \in \mathbb{Z}^E.$$

Fix an edge $e_* \in E$, and let $v_* = t(e_*)$. Let $L_E \subset \mathbb{Z}^E$ be the subgroup generated by e_* and $\{\Delta_e\}_{e \neq e_*}$. Then the sandpile group associated to $\mathcal{L}G$ and e_* is

$$K(\mathcal{L}G, e_*) = \mathbb{Z}^E / L_E.$$

Note that $\mathcal{L}G$ may not be Eulerian even when G is Eulerian.

Lemma 3.1 *Let $\phi : \mathbb{Z}^E \to \mathbb{Z}^V$ be the $\mathbb{Z}$-linear map sending $e \mapsto t(e)$. If G is Eulerian, then ϕ descends to a surjective group homomorphism*

$$\bar{\phi} : K(\mathcal{L}G, e_*) \to K(G, v_*).$$

Proof: To show that ϕ descends, it suffices to show that $\phi(L_E) \subset L_V$. For any $e \in E$, we have

$$\phi(\Delta_e) = \sum_{s(f)=t(e)} (t(f) - t(e)) = \Delta_{t(e)}.$$

The right side lies in L_V by definition if $t(e) \neq v_*$. Moreover, since G is Eulerian,

$$\sum_{v \in V} \Delta_v = \sum_{e \in E} (t(e) - s(e)) = \sum_{v \in V} (\text{indeg}(v) - \text{outdeg}(v)) v = 0,$$

so $\Delta_{v_*} = -\sum_{v \neq v_*} \Delta_v$ also lies in L_V. Finally, $\phi(e_*) = v_* \in L_V$, and hence $\phi(L_E) \subset L_V$.

Since G is strongly connected, every vertex has at least one incoming edge, so ϕ is surjective, and hence $\bar{\phi}$ is surjective. □

Let k be a positive integer. We say that G is *balanced k-regular* if $\text{indeg}(v) = \text{outdeg}(v) = k$ for every vertex v. Note that any balanced k-regular graph is Eulerian; and if G is balanced k-regular, then its directed line graph $\mathcal{L}G$ is also balanced k-regular. In particular, this implies

$$\sum_{e \in E} \Delta_e = 0$$

so that $\Delta_{e_*} \in L_E$.

Now consider the $\mathbb{Z}$-linear map

$$\psi : \mathbb{Z}^V \to \mathbb{Z}^E$$

sending $v \mapsto \sum_{s(e)=v} e$. For a group Γ, write $k\Gamma = \{kg | g \in \Gamma\}$. The following lemma is proved in [11].

Lemma 3.2 *If G is balanced k-regular, then ψ descends to a group isomorphism*

$$\bar{\psi} : K(G) \xrightarrow{\simeq} k\, K(\mathcal{L}G).$$

Proof of Theorem 1.2: If G is Eulerian, then ϕ descends to a surjective homomorphism of sandpile groups by Lemma 3.1. If G is balanced k-regular, then $\bar{\psi}$ is injective by Lemma 3.2, so

$$\ker(\bar{\phi}) = \ker(\bar{\psi} \circ \bar{\phi}).$$

Moreover for any edge $e \in E$

$$(\psi \circ \phi)(e) = \sum_{s(f)=t(e)} f = ke + \Delta_e.$$

Hence $\bar{\psi} \circ \bar{\phi}$ is multiplication by k, and $\ker(\bar{\phi})$ is the k-torsion subgroup of $K(\mathcal{L}G)$.

4 Iterated Line Graphs

Let $G = (V, E)$ be a finite directed graph, loops and multiple edges allowed. The *iterated line digraph* $\mathcal{L}^n G = (E_n, E_{n+1})$ has as vertices the set

$$E_n = \{(e_1, \ldots, e_n) \in E^n \,|\, s(e_{i+1}) = t(e_i),\ i = 1, \ldots, n-1\}$$

of directed paths of n edges in G. The edge set of $\mathcal{L}^n G$ is E_{n+1}, and the incidence is defined by

$$\begin{aligned} s(e_1, \ldots, e_{n+1}) &= (e_1, \ldots, e_n); \\ t(e_1, \ldots, e_{n+1}) &= (e_2, \ldots, e_{n+1}). \end{aligned}$$

(We also set $E_0 = V$, and $\mathcal{L}^0 G = G$.) For example, the de Bruijn graph DB_n is $\mathcal{L}^n(DB_0)$, where DB_0 is the graph with one vertex and two loops.

Our next result relates the number of spanning trees of G and $\mathcal{L}^n G$. Given a vertex $v \in V$, let

$$p(n, v) = \#\{(e_1, \ldots, e_n) \in E_n \,|\, t(e_n) = v\}$$

be the number of directed paths of n edges in G ending at vertex v.

Theorem 4.1 *Let $G = (V, E)$ be a finite directed graph with no sources. Then*

$$\kappa(\mathcal{L}^n G) = \kappa(G) \prod_{v \in V} \text{outdeg}(v)^{p(n,v)-1}.$$

Proof: For any $j \geq 0$, by Theorem 1.1 applied to $\mathcal{L}^j G$ with all edge weights 1,

$$\begin{aligned}\frac{\kappa(\mathcal{L}^{j+1}G)}{\kappa(\mathcal{L}^j G)} &= \prod_{(e_1,\ldots,e_j)\in E_j} \text{outdeg}(t(e_j))^{\text{indeg}(s(e_1))-1} \\ &= \prod_{v\in V} \text{outdeg}(v)^{p(j+1,v)-p(j,v)}.\end{aligned}$$

Taking the product over $j = 0, \ldots, n-1$ yields the result. □

When G is balanced k-regular, we have $p(n, v) = k^n$ for all vertices v, so we obtain as a special case of Theorem 4.1 the result of Huaxiao, Fuji and Qiongxiang [9, Theorem 1]

$$\kappa(\mathcal{L}^n G) = \kappa(G) k^{(k^n-1)\#V}.$$

In particular, taking $G = DB_0$ yields the classical formula

$$\kappa(DB_n) = 2^{2^n-1}.$$

Since DB_n is Eulerian, the number $\kappa(DB_n, v_*)$ of oriented spanning trees rooted at v_* does not depend on v_*, so

$$\kappa(DB_n, v_*) = 2^{-n}\kappa(DB_n) = 2^{2^n-n-1}. \tag{5}$$

This familiar number counts *de Bruijn sequences* of order $n+1$ (Eulerian tours of DB_n) up to cyclic equivalence. De Bruijn sequences are in bijection with oriented spanning trees of DB_n rooted at a fixed vertex v_*; for more on the connection between spanning trees and Eulerian tours, see [6] and [15, section 5.6].

Perhaps less familiar is the situation when G is not regular. As an example, consider the graph

$$G = (\{0, 1\}, \{(0, 0), (0, 1), (1, 0)\}).$$

The vertices of its iterated line graph $\mathcal{L}^n G$ are binary words of length $n + 1$ containing no two consecutive 1's. The number of such words is the Fibonacci number F_{n+3}, and the number of words ending in 0 is F_{n+2}. By Theorem 4.1, the number of oriented spanning trees of $\mathcal{L}^n G$ is

$$\kappa(\mathcal{L}^n G) = 2 \cdot 2^{p(n,0)-1} = 2^{F_{n+2}}.$$

Next we turn to the proofs of Theorems 1.3 and 1.4. If a and b are positive integers, we write $\mathbb{Z}_b^a$ for the group $(\mathbb{Z}/b\mathbb{Z}) \oplus \ldots \oplus (\mathbb{Z}/b\mathbb{Z})$ with a summands.

Proof of Theorem 1.3: Induct on n. From (5) we have

$$\#K(DB_n) = 2^{2^n-n-1}$$

hence

$$K(DB_n) = \mathbb{Z}_2^{a_1} \oplus \mathbb{Z}_4^{a_2} \oplus \mathbb{Z}_8^{a_3} \oplus \ldots \oplus \mathbb{Z}_{2^m}^{a_m}$$

for some nonnegative integers m and $a_1, \ldots, a_m$ satisfying

$$\sum_{j=1}^{m} j a_j = 2^n - n - 1. \tag{6}$$

By Lemma 3.2 and the inductive hypothesis,

$$\begin{aligned} \mathbb{Z}_2^{a_2} \oplus \mathbb{Z}_4^{a_3} \oplus \ldots \oplus \mathbb{Z}_{2^{m-1}}^{a_m} &\simeq 2K(DB_n) \\ &\simeq K(DB_{n-1}) \\ &\simeq \mathbb{Z}_2^{2^{n-3}} \oplus \mathbb{Z}_4^{2^{n-4}} \oplus \ldots \oplus \mathbb{Z}_{2^{n-2}}. \end{aligned}$$

hence $m = n - 1$ and

$$a_2 = 2^{n-3}, a_3 = 2^{n-4}, \ldots, a_{n-1} = 1.$$

Solving (6) for a_1 now yields $a_1 = 2^{n-2}$.

For p prime, by carrying out the same argument on a general balanced p-regular directed graph G on N vertices, we find that

$$K(\mathcal{L}^n G) \simeq \tilde{K} \oplus \bigoplus_{j=1}^{n-1} (\mathbb{Z}_{p^j})^{p^{n-1-j}(p-1)^2 N} \oplus (\mathbb{Z}_{p^n})^{(p-1)N-r-1} \oplus \bigoplus_{j=1}^{m} (\mathbb{Z}_{p^{n+j}})^{a_j}$$

where

$$\text{Sylow}_p(K(G)) = (\mathbb{Z}_p)^{a_1} \oplus \ldots \oplus (\mathbb{Z}_{p^m})^{a_m};$$

$$\tilde{K} = K(G)/\text{Sylow}_p(K(G));$$

$$r = a_1 + \ldots + a_m.$$

In particular, taking $G = \text{Kautz}_1$ with $p = 2$, we have $K(G) = \tilde{K} = \mathbb{Z}_3$, and we arrive at Theorem 1.4.

5 Concluding Remark

Theorem 1.2 describes a map from the sandpile group $K(\mathcal{L}G, e_*)$ to the group $K(G, v_*)$ when G is an Eulerian directed graph and $e_* = (w_*, v_*)$ is an edge of G. There is also a suggestive numerical relationship between the orders of the sandpile groups $K(\mathcal{L}G, e_*)$ and $K(G, w_*)$, which holds even when G is not Eulerian: by equation (3) we have

$$\kappa(G, w_*) \,|\, \kappa(\mathcal{L}G, e_*)$$

whenever G satisfies the hypothesis of Theorem 2.1. This observation leads us to ask whether $K(G, w_*)$ can be expressed as a subgroup or quotient group of $K(\mathcal{L}G, e_*)$.

References

[1] A. Berget, A. Manion, M. Maxwell, A. Potechin and V. Reiner, The critical group of a line graph, *Ann. Combin.*, to appear. `http://arxiv.org/abs/0904.1246`.

[2] H. Bidkhori and S. Kishore, Counting the spanning trees of a directed line graph. `http://arxiv.org/abs/0910.3442`

[3] N. L. Biggs, Chip-firing and the critical group of a graph, *J. Algebraic Combin.* **9**, no. 1 (1999), 25–45.

[4] D. Dhar, Self-organized critical state of sandpile automaton models, *Phys. Rev. Lett.* **64** (1990), 1613–1616.

[5] D.-Z. Du, Y.-D. Lyuu, and F. D. Hsu, Line digraph iterations and connectivity analysis of de Bruijn and Kautz graphs, *IEEE Trans. Comput.* **42**, no. 5 (1993), 612–616.

[6] T. van Aardenne-Ehrenfest and N. G. de Bruijn, Circuits and trees in oriented linear graphs, *Simon Stevin* **28** (1951), 203–217.

[7] M. A. Fiol, J. L. A. Yebra and I. A. De Miquel, Line digraph iterations and the (d, k)-digraph problem, *IEEE Trans. Comput.* **33**, no. 5 (1984), 400–403.

[8] A. E. Holroyd, L. Levine, K. Mészáros, Y. Peres, J. Propp and D. B. Wilson, Chip-firing and rotor-routing on directed graphs, *In and out of equilibrium 2*, 331–364, Progr. Probab. **60**, Birkhäuser, 2008. `http://arxiv.org/abs/0801.3306`

[9] Z. Huaxiao, Z. Fuji and H. Qiongxiang, On the number of spanning trees and Eulerian tours in iterated line digraphs, *Discrete Appl. Math.* **73**, no. 1 (1997), 59–67.

[10] D. E. Knuth, Oriented subtrees of an arc digraph, *J. Comb. Theory* **3** (1967), 309–314.

[11] L. Levine, Sandpile groups and spanning trees of directed line graphs, *J. Comb. Theory A*, to appear. `http://arxiv.org/abs/0906.2809`

[12] D. J. Lorenzini, Arithmetical graphs, *Math. Ann.* **285**, no. 3 (1989), 481–501.

[13] D. J. Lorenzini, A finite group attached to the Laplacian of a graph, *Discrete Math.* **91**, no. 3 (1991), 277–282.

[14] E. R. Speer, Asymmetric abelian sandpile models. *J. Statist. Phys.* **71** (1993), 61–74.

[15] R. P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, 1999.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 268–279

Pattern avoidance in alternating permutations and tableaux (extended abstract)

Joel Brewster Lewis

Massachusetts Institute of Technology
77 Massachusetts Avenue, Room 2-333
Cambridge, MA 02139

Abstract. We give bijective proofs of pattern-avoidance results for a class of permutations generalizing alternating permutations. The bijections employed include a modified form of the RSK insertion algorithm and recursive bijections based on generating trees. As special cases, we show that the sets $A_{2n}(1234)$ and $A_{2n}(2143)$ are in bijection with standard Young tableaux of shape $\langle 3^n \rangle$.

Alternating permutations may be viewed as the reading words of standard Young tableaux of a certain skew shape. In the last section of the paper, we study pattern avoidance in the reading words of standard Young tableaux of any skew shape. We show bijectively that the number of standard Young tableaux of shape λ/μ whose reading words avoid 213 is a natural μ-analogue of the Catalan numbers. Similar results for the patterns 132, 231 and 312.

Résumé. Nous présentons des preuves bijectives de résultats pour une classe de permutations à motifs exclus qui généralisent les permutations alternantes. Les bijections utilisées reposent sur une modification de l'algorithme d'insertion "RSK" et des bijections récursives basées sur des arbres de génération. Comme cas particuliers, nous montrons que les ensembles $A_{2n}(1234)$ et $A_{2n}(2143)$ sont en bijection avec les tableaux standards de Young de la forme $\langle 3^n \rangle$.

Une permutation alternante peut être considérée comme le mot de lecture de certain skew tableau. Dans la dernière section de l'article, nous étudions l'évitement des motifs dans les mots de lecture de skew tableaux genéraux. Nous montrons bijectivement que le nombre de tableaux standards de forme λ/μ dont les mots de lecture évitent 213 est un μ-analogue naturel des nombres de Catalan. Des résultats analogues sont valables pour les motifs 132, 231 et 312.

Resumen. Presentamos pruebas biyectivas de resultados de "evasión de patrones" para una clase de permutaciones que generalizan permutaciones alternantes. Las biyecciónes utilizadas incluyen una modificación del algoritmo de inserción de RSK y una biyección recursiva basada en árboles generatrices. Mostramos, como casos especiales, que los conjuntos $A_{2n}(1234)$ y $A_{2n}(2143)$ están en biyección con los tableaux de Young estándares de forma $\langle 3^n \rangle$.

Las permutaciones alternantes pueden ser entendidas como palabras de lectura de tableaux de Young estándares de cierta forma sezgada. En la ultima sección del articulo, expandimos nuestro estudio al considerar evasión de patrones en las palabras de lectura de tableaux de Young estándares de cualquier forma sezgada. Mostramos biyectivamente que el número de tableaux de Young estándares de forma λ/μ cuyas palabras de lectura evitan 213 es un μ-anólogo de los números de Catalán y resultados similares para los patrones 132, 231 y 312.

Keywords: Alternating permutations, permutation patterns, RSK, generating trees, Young tableaux

1 Introduction

A classical problem asks for the number of permutations that avoid a certain permutation pattern. This problem has received a great deal of attention (see *e.g.*, [12, 3]) and has led to a number of interesting variations including the enumeration of special classes of pattern-avoiding permutations (*e.g.*, involutions [12] and derangements [9]). One such variation, first studied by Mansour in [8], is the enumeration of *alternating* permutations avoiding a given pattern or collection of patterns. Alternating permutations have the intriguing property [8, 15, 4] that for any pattern of length three, the number of alternating permutations of a given length avoiding that pattern is given by a Catalan number. This property is doubly interesting because it is shared by the class of all permutations. This coincidence suggests that pattern avoidance in alternating permutations and in usual permutations may be closely related and so motivates the study of pattern avoidance in alternating permutations.

In this paper, we extend the study of pattern avoidance in alternating permutations to patterns of length four. In particular, we show that the number of alternating permutations of length $2n$ avoiding either of the patterns 1234 or 2143 is $\frac{2\cdot(3n)!}{n!(n+1)!(n+2)!}$. This is the first enumeration of a set of pattern-avoiding alternating permutations for a single pattern of length four. In the case of 1234, we give a direct bijective proof using a variation of RSK, while in the case of 2143 we give a recursive generating tree bijection.

Most of our bijections work in a more general setting in which we replace alternating permutations with the set $\mathcal{L}_{n,k}$ of reading words of standard Young tableaux of certain nice skew shapes. (These permutations are enumerated with no pattern restriction in [1].) Inspired by the idea of permutations as reading words of tableaux, we give an enumeration of standard skew Young tableaux of *any* fixed shape whose reading words avoid certain patterns. In particular, this provides a uniform argument to enumerate permutations in S_n and permutations in $\mathcal{L}_{n,k}$ that avoid either 132 or 213. That such a bijection should exist is far from obvious, and it raises the possibility that there is substantially more to be said in this area. In the remainder of this introduction, we provide a more detailed summary of results.

Both the set of all permutations of a given length and the set of alternating permutations of a given length can be expressed as the set of reading words of the standard Young tableaux of a particular skew shape (essentially a difference of two staircases). We define a class $\mathcal{L}_{n,k} \subseteq S_{nk}$ of permutations such that $\mathcal{L}_{n,1} = S_n$ is the set of all permutations of length n, $\mathcal{L}_{n,2}$ is the set of alternating permutations of length $2n$, and for each k $\mathcal{L}_{n,k}$ is the set of reading words of the standard Young tableaux of a certain skew shape. In Section 2, we provide definitions of all the most important objects in this paper. In Sections 3 and 4, we use bijective proofs to derive enumerative pattern avoidance results for $\mathcal{L}_{n,k}$. In Section 3 we give a simple bijection between elements of $\mathcal{L}_{n,k}$ with no $(k+1)$-term increasing subsequence and standard Young tableaux of rectangular shape $\langle k^n \rangle$. In Section 4 we exhibit two bijections between the elements of $\mathcal{L}_{n,k}$ with no $(k+2)$-term increasing subsequence and standard Young tableaux of rectangular shape $\langle (k+1)^n \rangle$, one of which is a modified version of the famous RSK bijection and the other of which is a generating tree approach that also yields an enumeration of alternating permutations avoiding 2143.

In Section 5, we broaden our study to arbitrary skew shapes and so initiate the study of pattern avoidance in reading words of skew tableaux of any shape. In Section 5.1, we show bijectively that the number of tableaux of shape λ/μ (under a technical restriction on the possible shapes that sacrifices no generality – see Note 2) whose reading words avoid 213 can be easily computed from the shape. Notably, the resulting value does not depend on λ and is in fact a natural μ-generalization of the Catalan numbers. Replacing 213 with $132, 231$ or 312 leads to similar results.

For a complete version of this extended abstract, see [6] and [5].

2 Definitions

A *permutation* w of length n is a word containing each of the elements of $[n] = \{1, 2, \ldots, n\}$ exactly once. The set of permutations of length n is denoted S_n. Given a word $w = w_1 w_2 \cdots w_n$ and a permutation $p = p_1 \cdots p_k \in S_k$, we say that w *contains the pattern* p if there exists a set of indices $1 \le i_1 < i_2 < \ldots < i_k \le n$ such that the subsequence $w_{i_1} w_{i_2} \cdots w_{i_k}$ of w is order-isomorphic to p, *i.e.*, $w_{i_\ell} < w_{i_m}$ if and only if $p_\ell < p_m$. Otherwise, w is said to *avoid* p. Given a pattern p and a set S of permutations, we denote by $S(p)$ the set of elements of S that avoid p. For example, $S_n(123)$ is the set of permutations of length n avoiding the pattern 123, *i.e.*, the set of permutations with no three-term increasing subsequence.

A permutation $w = w_1 w_2 \cdots w_n$ is *alternating* if $w_1 < w_2 > w_3 < w_4 > \ldots$. (Note that in the terminology of [13], these "up-down" permutations are *reverse alternating* while alternating permutations are "down-up" permutations. Luckily, this convention doesn't matter: any pattern result on either set can be translated into a result on the other via *complementation*, *i.e.*, by considering w^c such that $w^c_i = n + 1 - w_i$. Then results for the pattern 123 would be replaced by results for 321 and so on.) We denote by A_n the set of alternating permutations of length n.

For $n, k \ge 1$, let $\mathcal{L}_{n,k}$ be the set of permutations $w = w_{1,1} w_{1,2} \cdots w_{1,k} w_{2,1} \cdots w_{n,k}$ in S_{nk} that satisfy the conditions

L1. $w_{i,j} < w_{i,j+1}$ for all $1 \le i \le n$, $1 \le j \le k - 1$, and

L2. $w_{i,j+1} > w_{i+1,j}$ for all $1 \le i \le n - 1$, $1 \le j \le k - 1$.

Note in particular that $\mathcal{L}_{n,1} = S_n$ (we have no restrictions in this case) and $\mathcal{L}_{n,2} = A_{2n}$. For any k and n, $\mathcal{L}_{n,k}(12 \cdots k) = 0$. Thus, for monotone pattern-avoidance in $\mathcal{L}_{n,k}$ we should consider patterns of length $k + 1$ or longer. The set $\mathcal{L}_{n,k}$ has been enumerated by Baryshnikov and Romik [1], and the formulas that result are quite simple for small values of k.

Note 1 *If $w = w_{1,1} \cdots w_{n,k} \in S_{nk}$ satisfies L1 and also avoids $12 \cdots (k+1)(k+2)$ then it automatically satisfies L2, since a violation $w_{i,j+1} < w_{i+1,j}$ of L2 leads immediately to a $(k + 2)$-term increasing subsequence $w_{i,1} < \ldots < w_{i,j+1} < w_{i+1,j} < \ldots < w_{i+1,k}$. In particular, we can also describe $\mathcal{L}_{n,k}(1 \cdots (k + 2))$ (respectively, $\mathcal{L}_{n,k}(1 \cdots (k + 1))$) as the set of permutations in $S_{nk}(1 \cdots (k + 2))$ (respectively, $S_{nk}(1 \cdots (k + 1))$) whose descent set is (or in fact, is contained in) $\{k, 2k, \ldots, (n - 1)k\}$.*

A *partition* is a weakly decreasing, finite sequence of nonnegative integers. We consider two partitions that differ only in the number of trailing zeroes to be the same. We write partitions in sequence notation, as $\langle \lambda_1, \lambda_2, \ldots, \lambda_n \rangle$, or to save space, with exponential notation instead of repetition of equal elements. Thus, the partition $\langle 5, 5, 3, 3, 2, 1 \rangle$ may be abbreviated $\langle 5^2, 3^2, 2, 1 \rangle$. If the sum of the entries of λ is equal to m then we write $\lambda \vdash m$.

Given a partition $\lambda = \langle \lambda_1, \lambda_2, \ldots \rangle$, the *Young diagram* of shape λ is the left-justified array of $\lambda_1 + \ldots + \lambda_n$ boxes with λ_1 in the first row, λ_2 in the second row, and so on. We will identify each partition with its Young diagram and speak of them interchangeably. A *skew Young diagram* λ/μ is the diagram that results when we remove the boxes of μ from those of λ, when both are arranged so that their first rows and first columns coincide. If λ/μ is a skew Young diagram with n boxes, a *standard Young tableau* of shape λ/μ is a filling of the boxes of λ/μ with $[n]$ so that each element appears in exactly one box, and entries increase along rows and columns. We identify boxes in a (skew) Young diagram using matrix coordinates, so the box in the first row and second column is numbered $(1, 2)$. We denote by $\mathrm{sh}(T)$ the

				2	3	6
			1	5	9	
		4	11	12		
	8	13	15			
7	10	14				

Fig. 1: A standard skew Young tableau (in English notation, *i.e.*, with the first row on top) whose reading word is the permutation $7\,10\,14\,8\,13\,15\,4\,11\,12\,1\,5\,9\,2\,3\,6 \in \mathcal{L}_{5,3}$.

shape of the standard Young tableau T, by $\mathrm{SYT}(\lambda)$ the set of standard Young tableaux of shape λ and by $f^\lambda = |\,\mathrm{SYT}(\lambda)|$ the size of this set.

Given a standard Young tableau T, the *reading word* of T is the permutation that consists of the entries of the last row read from left to right, then the next-to-last row, and so on. For example, the reading words of the tableaux of shape $\langle n, n-1, \ldots, 2, 1\rangle / \langle n-1, n-2, \ldots, 1\rangle$ are all of S_n, and similarly $\mathcal{L}_{n,k}$ is equal to the set of reading words of standard skew Young tableaux of shape $\langle n+k-1, n+k-2, \ldots, k\rangle / \langle n-1, n-2, \ldots, 1\rangle$, as illustrated in Figure 1. The other "usual" reading order, from right to left then top to bottom in English notation, is simply the reverse of our reading order. Consequently, any pattern-avoidance result in our case carries over to the other reading order by taking the *reverse* of all permutations and patterns involved, *i.e.*, by replacing $w = w_1 \ldots w_n$ with $w^r = w_n \cdots w_1$.

We make note of two operations on Young diagrams and tableaux. Given a partition λ, the *conjugate partition* λ' is defined so that the ith row of λ' has the same length as the ith column of λ for all i. Similarly, the conjugate of a skew Young diagram λ/μ is defined by $(\lambda/\mu)' = \lambda'/\mu'$. Given a standard skew Young tableau T of shape λ/μ, the conjugate tableau T' of shape $(\lambda/\mu)'$ is defined to have the entry a in box (i, j) if and only if T has the entry a in box (j, i). Geometrically, all these operations can be described as "reflection through the main diagonal." Given a skew Young diagram λ/μ, rotation by $180°$ gives a new diagram $(\lambda/\mu)^*$. Given a tableaux T with n boxes, we can form T^*, the *rotated-complement* of T, by rotating T by $180°$ and replacing the entry i with $n+1-i$ for each i. Observe that the reading word of T^* is exactly the reverse-complement of the reading word of T.

The *Schensted insertion algorithm*, or equivalently the *RSK correspondence*, is an extremely powerful tool relating permutations to pairs of standard Young tableaux. For a description of the bijection and a proof of its correctness and some of its properties, we refer the reader to [14, Chapter 7]. Our use of notation follows that source, so in particular we denote by $T \leftarrow i$ the tableau that results when we (row-) insert i into the tableau T. Particular properties of RSK will be quoted as needed.

3 The pattern $12 \cdots (k+1)$

In this section we give the simplest of the bijections in this paper.

Proposition 3.1 *There is a bijection between $\mathcal{L}_{n,k}(12 \cdots (k+1))$ and the set of standard Young tableaux of shape $\langle k^n \rangle$.*

We have $f^{\langle n\rangle} = f^{\langle 1^n\rangle} = 1$ and $f^{\langle n,n\rangle} = f^{\langle 2^n\rangle} = \frac{1}{n+1}\binom{2n}{n} = C_n$, the nth Catalan number. By the hook-length formula [14, 11] we have

$$f^{\langle k^n\rangle} = \frac{(kn)! \cdot 1! \cdot 2! \cdots (k-1)!}{n! \cdot (n+1)! \cdots (n+k-1)!}.$$

So Proposition 3.1 says $|\mathcal{L}_{n,k}(1\cdots(k+1))| = f^{\langle k^n\rangle}$. For $k = 1$, this is the uninspiring result $|S_n(12)| = 1$. For $k = 2$, it tells us $|A_{2n}(123)| = C_n$, a result that Stanley [15] attributes to Deutsch and Reifegerste.

Proof idea: The bijection is to identify the permutation $w = w_{1,1}\cdots w_{n,k} \in \mathcal{L}_{n,k}(12\cdots k(k+1))$ with the tableau $T \in \mathrm{SYT}(\langle k^n\rangle)$ given by $T_{i,j} = w_{n+1-i,j}$. It is not difficult to verify that the conditions on w correspond precisely to the conditions that T be a standard Young tableau and vice-versa. □

Both directions of this bijection are more commonly seen with other names. The map that sends $w \mapsto T$ is actually the Schensted insertion algorithm used in the RSK correspondence. (For any $w \in \mathcal{L}_{n,k}(1\cdots(k+1))$, the recording tableau is the tableau whose first row contains the $\{1,\ldots,k\}$, second row contains $\{k+1,\ldots,2k\}$, and so on.) The map that sends $T \mapsto w$ is the reading-word map as defined in Section 2.

4 The pattern $12\cdots(k+2)$

There are several nice proofs of the equality $|S_n(123)| = C_n$ including a clever application of the RSK algorithm [14, Problem 6.19(ee)]. In this section, we give two bijective proofs of the following generalization of this result:

Theorem 4.1 *There is a bijection between $\mathcal{L}_{n,k}(12\cdots(k+2))$ and the set of standard Young tableaux of shape $\langle(k+1)^n\rangle$ and so*

$$|\mathcal{L}_{n,k}(12\cdots(k+2))| = f^{\langle(k+1)^n\rangle}.$$

For $k = 1$ this is a rederivation of the equality $|S_n(123)| = C_n$ while for $k = 2$ it implies

Corollary 4.2 *We have $|A_{2n}(1234)| = f^{\langle 3^n\rangle} = \dfrac{2(3n)!}{n!(n+1)!(n+2)!}$ for all $n \geq 0$.*

We believe this to be the first computation of any expression of the form $A_{2n}(\pi)$ or $A_{2n+1}(\pi)$ for $\pi \in S_4$. One can derive the complementary result for $|A_{2n+1}(1234)|$ using similar methods.

The first of our two bijections makes use of a modification of Schensted insertion, and the key idea for the modification appears in [10] (in the context of *doubly-alternating permutations*). The second bijection makes use of generating trees; its proof involves a number of technical results that we omit in this extended abstract.

4.1 A bijection using a modified version of RSK

In this section, we prove Theorem 4.1 using a modification of the RSK insertion algorithm. Recall that the RSK is a bijection between S_n and pairs (P, Q) of standard Young tableaux such that $\mathrm{sh}(P) = \mathrm{sh}(Q) \vdash n$ with the following properties:

Theorem 4.3 ([14, 7.11.2(b)]) *If P is a standard Young tableau and $j < k$ then the insertion path of j in $P \leftarrow j$ lies strictly to the left of the insertion path of k in $(P \leftarrow j) \leftarrow k$, and the latter insertion path does not extend below the former.*

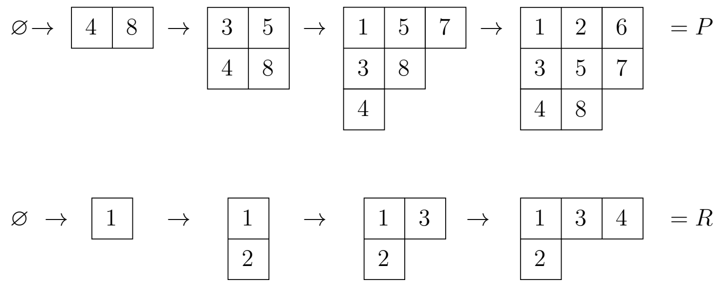


Fig. 2: An application of our modified version of RSK to the permutation $48351726 \in \mathcal{L}_{4,2}(1234)$. Note that only every other insertion step is shown in the construction of P.

Theorem 4.4 ([14, 7.23.11]) *If $w \in S_n$ and $w \xrightarrow{RSK} (P, Q)$ with $\mathrm{sh}(P) = \mathrm{sh}(Q) = \lambda$, then λ_1 is the length of the longest increasing subsequence in w.*

Now we describe a bijection from $\mathcal{L}_{n,k}(12\cdots(k+2))$ to pairs (P, R) of standard Young tableau such that P has nk boxes, R has n boxes, and the shape of R can be rotated 180° and joined to the shape of P to form a rectangle of shape $\langle (k+1)^n \rangle$. (In other words, $\mathrm{sh}(P)'_i + \mathrm{sh}(R)'_{k+2-i} = n$ for $1 \leq i \leq k+1$.) Observe that the set of such pairs of tableaux is in natural bijection with the set of standard Young tableaux of shape $\langle (k+1)^n \rangle$: given a tableau of shape $\langle (k+1)^n \rangle$, break off the portion of the tableau filled with $nk+1, \ldots, n(k+1)$, rotate it 180° and replace each value i that appears in it with $nk+n+1-i$.

Given a permutation $w = w_{1,1}w_{1,2}\cdots w_{1,k}w_{2,1}\cdots w_{n,k}$, let $P_0 = \varnothing$ and for $1 \leq i \leq n$ let $P_i = (\cdots((P_{i-1} \leftarrow w_{i,1}) \leftarrow w_{i,2})\cdots) \leftarrow w_{i,k}$. Define $P = P_n$, so P is the usual RSK insertion tableau for w. Define R as follows: set $R_0 = \varnothing$ and $\lambda_i = \mathrm{sh}(P_i)$. Observe that by Theorem 4.3, λ_i/λ_{i-1} is a horizontal strip of size k and that by Theorem 4.4, λ_i/λ_{i-1} stretches no further right than the $(k+1)$th column. Thus there is a unique j such that λ_i/λ_{i-1} has boxes in the ℓth column for all $\ell \in [k+1] \setminus \{j\}$. Let R_i be the shape that arises from R_{i-1} by adding a box filled with i in the $(k+2-j)$th column, and define $R = R_n$. This map is illustrated in Figure 2.

Proposition 4.5 *The algorithm just described is a bijection between $\mathcal{L}_{n,k}(12\cdots(k+2))$ and pairs (P, R) of standard Young tableaux such that P has nk boxes, R has n boxes, and $\mathrm{sh}(R)$ can be rotated and joined to $\mathrm{sh}(P)$ to form a rectangle of shape $\langle (k+1)^n \rangle$.*

Proof: By construction, P is a standard Young tableau with nk boxes and R is a shape with n boxes filled with $[n]$ such that we may rotate R by 180° and join it to P in order to get a rectangle of shape $\langle (k+1)^n \rangle$. Moreover, we have from standard properties of RSK that each P_i is of partition shape and by construction that the corresponding R_i may be rotated 180° and joined to P_i to form a rectangle, so each of the R_i (including R itself) is a partition shape. Finally, the unique box in R_i but not in R_{i-1} is filled with i, which is larger than the entry in any box in R_{i-1}, so R is a standard Young tableau.

We have left to show that this process is a bijection, *i.e.*, we need that this map is invertible and that its inverse takes pairs of tableaux of the given sort to permutations with the appropriate restrictions. Invertibility is straightforward, since from a pair (P, R) of standard Young tableaux of appropriate shapes we can construct a pair of standard Young tableaux (P, Q) of the same shape such that $w \mapsto (P, R)$ under our algorithm exactly when $w \xrightarrow{\text{RSK}} (P, Q)$: if R has entry i in column $k+2-j$, place the entries

$ki-k+1, ki-k+2, \ldots, ki$ respectively into columns $1, \ldots, j-1, j+1, \ldots, k+1$ of Q. Moreover, by Theorem 4.3 we have that the preimage under RSK of this pair (P, Q) must consist of n runs of k elements each in increasing order, *i.e.*, it must satisfy L1, and by Theorem 4.4 it must have no increasing subsequence of length $k+2$. Then by the remarks in Section 2 following the definition of $\mathcal{L}_{n,k}$ we have that the preimage satisfies L2 as well. This completes the proof. $\square$

4.2 A second approach using generating trees

Given a sequence $\{\Sigma_n\}_{n\geq 1}$ of nonempty sets with $|\Sigma_1| = 1$, a *generating tree* for this sequence is a rooted, labeled tree such that the vertices at level n are the elements of Σ_n and the label of each vertex determines the multiset of labels of its children. In other words, a generating tree is one particular type of recursive structure in which heredity is determined by some local data. We are particularly interested in generating trees for which the labels are (much) simpler than the objects they are labeling. In this case, we may easily describe a generating tree by giving the label L_1 of the *root vertex* (the element of Σ_1) and the *succession rule* $L \mapsto S$ that gives the set S of labels of the children in terms of the label L of the parent. Generating trees have proven to be an effective tool for finding bijections between different classes of pattern-avoiding permutations (see, *e.g.*, [16, 2]). In this section, we describe how generating trees can be used to give a second proof of Theorem 4.1 and to enumerate 2143-avoiding alternating permutations.

4.2.1 A tree for $\mathcal{L}_{n,k}$

There is a natural generating tree structure on $\bigcup_{n\geq 1} \mathcal{L}_{n,k}$: given a permutation $v \in \mathcal{L}_{n,k}$, its children are precisely the permutations $w \in \mathcal{L}_{n+1,k}$ such that the prefix of w of length nk is order-isomorphic to v. Since pattern containment is transitive, the subset $\bigcup_{n\geq 1} \mathcal{L}_{n,k}(p)$ of these permutations that avoid the pattern (or set of patterns) p is the set of vertices of a connected subtree. We now consider this restricted tree for the pattern $p = 12\cdots(k+2)$.

Given a permutation $w = w_1 w_2 \cdots w_{nk} \in \mathcal{L}_{n,k}(1\cdots(k+2))$, we associate a label $(a_2, \ldots, a_{k+1})$, where a_j is the smallest entry of w that is the largest entry in a j-term increasing subsequence, or $nk+1$ if there is no such entry. (Note that a_j could equivalently be defined as the last-occurring entry of w that is the largest term in a j-term increasing subsequence of w but is not the largest term in a $(j+1)$-term increasing subsequence.) Thus, for example, the unique permutation $12\cdots k \in \mathcal{L}_{1,k}(1\cdots(k+2))$ has label $(2, \ldots, k+1)$, while the permutation $136245 \in \mathcal{L}_{2,3}(12345)$ has label $(2, 4, 5)$.

Some relations among label entries are straightforward. For example, observe that if $(a_2, \ldots, a_{k+1})$ is the label of any permutation $u \in \mathcal{L}_{n,k}(1\cdots(k+2))$ then $2 \leq a_2 < \ldots < a_{k+1} \leq nk+1$. The following result (whose proof, which consists of many technical details and little insight, is omitted) characterizes the labels of children based on the labels of a parent.

Proposition 4.6 *Suppose that $u \in \mathcal{L}_{n,k}(12\cdots(k+2))$ has label $(a_2, \ldots, a_{k+1})$. Then for any k-tuple $(b_2, \ldots, b_{k+1})$ such that*

$$2 \leq b_2 < b_3 < \ldots < b_{k+1} \leq (n+1)k+1 \qquad \textit{and} \qquad b_j \leq a_j + j - 1 \qquad \textit{for all } j,$$

there is a unique child $w \in \mathcal{L}_{n+1,k}(12\cdots(k+2))$ of u with label $(b_2, \ldots, b_{k+1})$, and u has no other children.

4.2.2 A tree for Young tableaux

There is a natural generating tree on the set $\bigcup_{n\geq 1} \mathrm{SYT}(\langle (k{+}1)^n\rangle)$ of rectangular standard Young tableaux with $k+1$ columns: let a tableau T be the child of a tableau S if S is order-isomorphic to T with its first row removed.

Given a tableau $S \in \mathrm{SYT}(\langle (k+1)^n\rangle)$ with first row $(1, s_2, s_3, \ldots, s_{k+1})$, assign to it the label $(s_2, \ldots, s_{k+1})$. Thus, for example, that the unique tableau in $\mathrm{SYT}(\langle k+1\rangle)$ has label $(2, 3, \ldots, k+1)$, while the tableau

$$\begin{array}{|c|c|c|c|}\hline 1 & 2 & 4 & 5 \\ \hline 3 & 6 & 7 & 8 \\ \hline \end{array} \in \mathrm{SYT}(\langle 4, 4\rangle)$$

has label $(2, 4, 5)$. It's easy to see that if $(s_2, \ldots, s_{k+1})$ is the label of a tableau $T \in \mathrm{SYT}(\langle (k+1)^n\rangle)$ then $2 \leq s_2 < s_3 < \ldots < s_{k+1} \leq n(k+1) - (n-1) = nk+1$. Without too much effort, one can also show the following result:

Proposition 4.7 *Suppose that $S \in \mathrm{SYT}(\langle (k+1)^n\rangle)$ has label $(s_2, \ldots, s_{k+1})$. Then for any k-tuple $(t_2, \ldots, t_{k+1})$ such that*

$$2 \leq t_2 < t_3 < \ldots < t_{k+1} \leq (n+1)k+1 \qquad \textit{and} \qquad t_j \leq s_j + j - 1 \qquad \textit{for all } j,$$

there is a unique child $T \in \mathrm{SYT}(\langle (k+1)^{n+1}\rangle)$ of S with label $(t_2, \ldots, t_{k+1})$, and S has no other children.

Theorem 4.1 follows immediately from Propositions 4.6 and 4.7.

4.2.3 A tree for 2143-avoiding alternating permutations

If, as in [5], we restrict our focus to alternating permutations (*i.e.*, to $A_{2n} = \mathcal{L}_{n,2}$), brute-force computations suggest that there may be several patterns $p \in S_4$ such that $|A_{2n}(p)| = |A_{2n}(1234)|$ for all n. In this section we use generating trees to show that 2143 is one such pattern.

Given any permutation $w \in S_n$ and any $c \in [n+1]$, denote by $w \circ c$ the unique permutation in S_{n+1} whose last entry is c and whose first n entries are order-isomorphic to w. If $w = w_1 w_2 \cdots w_{2n} \in A_{2n}(2143)$, say that a value $c \in [2n+1]$ is *active* for w if $w \circ c$ avoids 2143. To each $w \in A_{2n}(2143)$, assign the label (a, b) where $a = w_{2n-1} + 1$ and b is equal to the number of values in $[n+1]$ that are active for w. The following result shows that with this labeling, the generating tree for $\bigcup_{n\geq 1} A_{2n}(2143)$ obeys a simple succession rule.

Proposition 4.8 *Suppose that $u \in A_{2n}(2143)$ has label (a, b). Then for any ordered pair (x, y) such that*

$$2 \leq x \leq a+1 \qquad \textit{and} \qquad x < y \leq b+2,$$

there is a unique child $w \in A_{2n+2}(2143)$ with label (x, y), and u has no other children.

One can easily verify that these conditions are equivalent to those of Propositions 4.6 and 4.7 in the case $k = 2$. Therefore, we may conclude with the following result.

Theorem 4.9 *For all $n \geq 1$ we have*

$$|A_{2n}(1234)| = |A_{2n}(2143)| = \frac{2 \cdot (3n)!}{n! \cdot (n+1)! \cdot (n+2)!}.$$

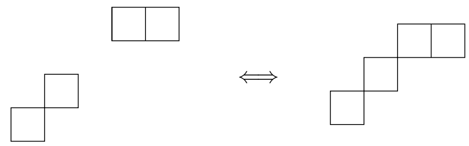


Fig. 3: Moving separated components gives a new shape but leaves the set of reading words of tableaux unchanged.

5 Pattern avoidance in reading words of tableaux of skew shapes

So far, we have considered permutations that arise as the reading words of standard skew Young tableaux of particular nice shapes. In this section, we expand our study to include pattern avoidance in the reading words of standard Young tableaux of *any* skew shape. As is the case for pattern avoidance in other settings, it is relatively simple to handle the case of small patterns (in our case, patterns of length three or less), but it appears to be quite difficult to prove exact results for larger patterns.

As we have seen, this new type of pattern avoidance encompasses pattern avoidance for the set of all permutations via the shape $\langle n, n-1 \ldots, 1\rangle / \langle n-1, n-2, \ldots, 1\rangle$, for alternating permutations via the shape $\langle n+1, n, \ldots, 2\rangle / \langle n-1, n-2, \ldots, 1\rangle$ and three other similar shapes, and for $\mathcal{L}_{n,k}$ for any k via the shape illustrated in Figure 1; it also incorporates other natural problems such as the enumeration of pattern-avoiding permutations with prescribed descent set (when the skew shape is a ribbon). Thus, on one hand the strength of our results is constrained by what is tractable to prove in these circumstances, while on the other hand any result we are able to prove in this context applies quite broadly.

Note 2 *We make the following general assumption on our Young diagrams: we will only ever be interested in diagrams λ/μ such that the inner (north-west) boundary of λ/μ contains the entire outer (south-east) boundary of μ. For example, the shape $\langle 4, 2, 1\rangle / \langle 2, 1\rangle$ meets this condition, while the shape $\langle 5, 2, 2, 1\rangle / \langle 3, 2, 1\rangle$ does not.*

Observe that imposing this restriction does not affect the universe of possible enumerative results: for a shape λ/μ failing this condition we can find a new shape λ'/μ' that passes it and has an identical set of reading words by moving the various disconnected components of λ/μ on the plane. For example, for $\lambda/\mu = \langle 5, 2, 2, 1\rangle / \langle 3, 2, 1\rangle$ we have $\lambda'/\mu' = \langle 4, 2, 1\rangle / \langle 2, 1\rangle$ – just slide disconnected sections of the tableau together until they share a corner. This example is illustrated in Figure 3.

5.1 The patterns 213 *and* 132

The equality $|S_n(213)| = |S_n(132)| = C_n$ is a simple recursive result. In [8] it was shown that $|A_{2n}(132)| = |A_{2n+1}(132)| = C_n$ (and so by reverse-complementation also $|A_{2n}(213)| = C_n$), and a bijective proof of this fact with implications for multiple-pattern avoidance was given in [7]. Here we extend this result to the reading words of tableaux of any fixed shape.

Theorem 5.1 *The number of tableaux of skew shape λ/μ whose reading words avoid the pattern* 213 *is equal to the number of partitions whose Young diagram is contained in that of μ (subject to Note 2).*

Note that this is a natural μ-generalization of the Catalan numbers: the outer boundaries of shapes contained in $\langle n-1, n-2, \ldots, 1\rangle$ are essentially Dyck paths of length $2n$ missing their first and last steps.

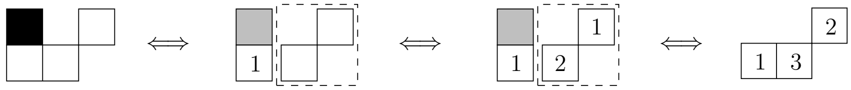


Fig. 4: Our bijection applied to the pair $(\langle 3,2\rangle/\langle 2\rangle, \langle 1\rangle)$ to generate a standard Young tableau.

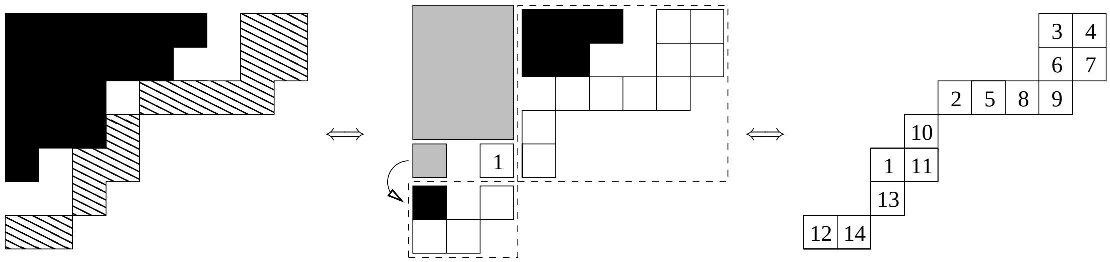


Fig. 5: A partial example: an application of our bijection to generate a standard Young tableau from the pair $(\langle 9,9,8,4,4,3,2\rangle/\langle 7,7,4,3,2,2\rangle, \langle 6,5,3,3,1\rangle)$.

Proof idea: We begin with a warm-up and demonstrate the claim in the case that μ is empty. In this case, the Proposition states that there is a unique standard Young tableau of a given shape $\lambda = \langle \lambda_1, \lambda_2, \ldots\rangle$ whose reading word avoids the pattern 213. In order to show this, we note that the reading word of every straight (*i.e.*, non-skew) tableau ends with an increasing run of length λ_1 and that the first entry of this run is 1. Since the reading word is 213-avoiding, each entry following the 1 must be smaller than every entry preceding the 1 and so this run consists of the values from 1 to λ_1. Applying the same argument to the remainder of the tableau (now with the minimal element $\lambda_1 + 1$), we see that the only possible filling is the one we get by filling the first row of the tableau with the smallest possible entries, then the second row with the smallest remaining entries, and so on. On the other hand, the reading word of the tableau just described is easily seen to be 213-avoiding, so we have our result in this case.

For the general case we give a recursive bijection. We recommend that the reader consult Figures 4 and 5 to most easily understand what follows.

Suppose we have a tableau T of shape λ/μ with entry 1 in position (i,j), an inner corner. Divide T into two pieces, one consisting of rows 1 through i with the box (i,j) removed, the other consisting of rows numbered $i+1, i+2$, etc. Let T_1 be the tableau order-isomorphic to the first part and let T_2 be the tableau order-isomorphic to the second part. Let $\nu = \langle \nu_1, \ldots, \nu_i\rangle$ be the result of applying this construction recursively to T_1 and let $\iota = \langle \iota_1, \iota_2, \ldots\rangle$ be the result of applying this construction recursively to T_2. Then the partition τ associated to T is given by $\tau = \langle \nu_1 + j, \ldots, \nu_i + j, \iota_1, \iota_2, \ldots\rangle$. That is, τ consists of all boxes (k,l) with $k < i$ and $l \leq j$ together with the result of applying our process to the right of this rectangle and the result of applying it below the rectangle, with the latter piece shifted up one row. By construction, τ is partition whose Young diagram fits inside μ.

To invert this process, start with a pair $(\lambda/\mu, \tau)$ of a skew and a non-skew shape such that τ fits inside μ. Let i be the largest index such that $\tau_{i-1} > \mu_i$, or let $i = 1$ if no such index exists. We divide τ and λ/μ into two pieces. For τ, we first remove the rectangle of shape $\langle (\mu_i + 1)^{i-1}\rangle$, leaving a partition to the right of the rectangle of shape $\nu_1 = \langle \tau_1 - \mu_i - 1, \tau_2 - \mu_i - 1, \ldots, \tau_{i-1} - \mu_i - 1\rangle$

and a second partition below the rectangle of shape $\nu_2 = \langle \tau_i, \tau_{i+1}, \ldots \rangle$. For λ/μ, we begin by filling the box $(i, \mu_i + 1)$ with the entry 1. Then we take the boxes to the right of this entry as one skew shape $\alpha_1/\beta_1 = \langle \lambda_1 - \mu_i - 1, \lambda_2 - \mu_i - 1, \ldots, \lambda_i - \mu_i - 1 \rangle / \langle \mu_1 - \mu_i - 1, \mu_2 - \mu_i - 1, \ldots, \mu_{i-1} - \mu_i - 1 \rangle$ and the boxes below it as our second skew shape $\alpha_2/\beta_2 = \langle \lambda_{i+1}, \lambda_{i+2}, \ldots \rangle / \langle \mu_{i+1}, \mu_{i+2}, \ldots \rangle$. Note that ν_2 fits inside β_2 and that ν_1 fits inside β_1 by the choice of i. Thus we may apply this construction recursively with the pairs $(\alpha_1/\beta_1, \nu_1)$ and $(\alpha_2/\beta_2, \nu_2)$, filling α_1/β_1 with the values $2, \ldots, |s_1| + 1$ and filling α_2/β_2 with the values $|\alpha_1/\beta_1| + 2, \ldots, |\lambda/\mu| = |\alpha_1/\beta_1| + |\alpha_2/\beta_2| + 1$. (Observe that this coincides with what we did in the first paragraph for $\mu = \varnothing$.)

One can prove by a simple inductive argument that these maps are mutually-inverse bijections between the sets in question. $\square$

Corollary 5.2 *We have that* $|\mathcal{L}_{n,k}(213)| = C_n$ *for all* $n, k \geq 1$.

Note that knowing the number of tableaux of each shape whose reading words avoid 213 automatically allows us to calculate for any shape the number of tableaux of that shape whose reading words avoid 132: if $\lambda = \langle \lambda_1, \ldots, \lambda_k \rangle$ and μ is contained in λ, the operation $T \mapsto T^*$ of rotation and complementation is a bijection between tableaux of shape λ/μ and tableaux of shape $\langle \lambda_1 - \mu_k, \lambda_1 - \mu_{k-1}, \ldots, \lambda_1 - \mu_1 \rangle / \langle \lambda_1 - \lambda_k, \lambda_1 - \lambda_{k-1}, \ldots, \lambda_1 - \lambda_2 \rangle$. Moreover, the reading word of T^* is the reversed-complement of the reading word of T, so the reading word of T avoids 132 if and only if the reading word of T^* avoids 213. This argument establishes the following corollary of Theorem 5.1:

Corollary 5.3 *The number of tableaux of skew shape* λ/μ *whose reading words avoid the pattern* 132 *is equal to the number of partitions whose Young diagram is contained in that of the partition* $\langle \lambda_1 - \lambda_k, \lambda_1 - \lambda_{k-1}, \ldots, \lambda_1 - \lambda_2 \rangle$.

Corollary 5.4 *We have that* $|\mathcal{L}_{n,k}(132)| = C_n$ *for all* $n, k \geq 1$.

5.2 The patterns 312 *and* 231

If the shape λ/μ contains a square, every tableau of that shape contains as a sub-tableau four entries

a	b
c	d

with $a < b < d$ and $a < c < d$, and the reading word of every such tableau is of the form $\ldots cd \ldots ab \ldots$. But any such permutation contains both an instance dab of the pattern 312 an instance cda of the pattern 231. Thus, the number of tableaux of shape λ/μ whose reading words avoid 312 or 231 is zero unless λ/μ contains no square, *i.e.*, unless λ/μ is contained in a ribbon. In this case, for a tableau T of shape λ/μ with reading word w we have that the reading word of the conjugate tableau T' is exactly the reverse w^r of w. Since w avoids 312 if and only if w^r avoids 213, we may apply Theorem 5.1 to deduce the following result.

Proposition 5.5 *If skew shape* λ/μ *is contained in a ribbon then the number of tableaux of shape* λ/μ *whose reading words avoid the pattern* 312 *is equal to the number of partitions whose Young diagram is contained in that of* μ. *Otherwise, the number of such tableaux is* 0.

Analogous arguments give the following result.

Corollary 5.6 *If skew shape λ/μ is contained in a ribbon then the number of tableaux of shape λ/μ whose reading words avoid the pattern 231 is equal to the number of partitions whose Young diagram is contained in that of the partition $\langle \lambda_1 - \lambda_k, \lambda_1 - \lambda_{k-1}, \ldots, \lambda_1 - \lambda_2 \rangle$. Otherwise, the number of such tableaux is* 0.

In the special case of $\mathcal{L}_{n,k}$ this says that for $k \geq 3$ and $n \geq 2$ we have $\mathcal{L}_{n,k}(231) = \mathcal{L}_{n,k}(312) = \varnothing$ while for $1 \leq k \leq 2$ we have that $|\mathcal{L}_{n,k}(231)|$ and $|\mathcal{L}_{n,k}(312)|$ are Catalan numbers [4, 15].

References

[1] Y. Baryshnikov and D. Romik. Enumeration formulas for Young tableaux in a diagonal strip. *Israel J. of Mathematics*, to appear.

[2] M. Bousquet-Mélou. Four classes of pattern-avoiding permutations under one roof: generating trees with two labels. *Electronic J. Combinatorics*, 9:R19, 2003.

[3] I. M. Gessel. Symmetric functions and P-recursiveness. *J. Combinatorial Theory, Series A*, 53:257–285, 1990.

[4] G. Hong. Catalan numbers in pattern-avoiding permutations. *MIT Undergraduate J. Mathematics*, 10:53–68, 2008.

[5] J. B. Lewis. Generating trees and pattern avoidance in alternating permutations. Available online at `arXiv:1005.4046v1`.

[6] J. B. Lewis. Pattern avoidance and RSK-like algorithms for alternating permutations and Young tableaux. Available online at `arXiv:0909.4966v2`.

[7] J. B. Lewis. Alternating, pattern-avoiding permutations. *Electronic J. Combinatorics*, 16:N7, 2009.

[8] T. Mansour. Restricted 132-alternating permutations and Chebyshev polynomials. *Annals of Combinatorics*, 7:201–227, 2003.

[9] T. Mansour and A. Robertson. Refined restricted permutations avoiding subsets of patterns of length three. *Annals of Combinatorics*, 6:407–418, 2002.

[10] E. Ouchterlony. Pattern avoiding doubly alternating permutations. *Proc. FPSAC 2006*. Available online at `http://garsia.math.yorku.ca/fpsac06/papers/83.pdf`.

[11] B. Sagan. *The Symmetric Group*. Springer-Verlag, 2001.

[12] R. Simion and F. W. Schmidt. Restricted permutations. *European J. Combinatorics*, 6:383–406, 1985.

[13] R. P. Stanley. *Enumerative Combinatorics, Volume 1*. Cambridge University Press, 1997.

[14] R. P. Stanley. *Enumerative Combinatorics, Volume 2*. Cambridge University Press, 2001.

[15] R. P. Stanley. Catalan number addendum to *Enumerative Combinatorics*. Available online at `http://www-math.mit.edu/~ rstan/ec/catadd.pdf`, 2009.

[16] J. West. Generating trees and forbidden subsequences. *Discrete Mathematics*, 157:363–372, 1996.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 280–288

Unitary Matrix Integrals, Primitive Factorizations, and Jucys-Murphy Elements

Sho Matsumoto[1] and Jonathan Novak[2,3]

[1]*Graduate School of Mathematics, Nagoya University, Furocho, Chikusa-ku, Nagoya, 464-8602, Japan*
[2]*Department of Combinatorics & Optimization, University of Waterloo, Waterloo, Canada*
[3]*The Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720-5070, USA*

Abstract. A factorization of a permutation into transpositions is called "primitive" if its factors are weakly ordered. We discuss the problem of enumerating primitive factorizations of permutations, and its place in the hierarchy of previously studied factorization problems. Several formulas enumerating minimal primitive and possibly non-minimal primitive factorizations are presented, and interesting connections with Jucys-Murphy elements, symmetric group characters, and matrix models are described.

Résumé. Une factorisation en transpositions d'une permutation est dite "primitive" si ses facteurs sont ordonnés. Nous discutons du problème de l'énumération des factorisations primitives de permutations, et de sa place dans la hiérarchie des problèmes de factorisation précédemment étudiés. Nous présentons plusieurs formules énumérant certaines classes de factorisations primitives, et nous soulignons des connexions intéressantes avec les éléments Jucys-Murphy, les caractéres des groupes symétriques, et les modèles de matrices.

Keywords: Primitive factorizations, Jucys-Murphy elements, matrix integrals.

1 Introduction

1.1 Polynomial integrals on unitary groups

Let $U(N)$ denote the group of $N \times N$ complex unitary matrices $U = [u_{ij}]_{1 \leq i,j \leq N}$. By a *polynomial function* on $U(N)$ we mean a function of the form

$$p(U) = \sum_{m,n \geq 0} \sum_{I,J,I',J'} c(I,J,I',J') U_{IJ} \overline{U}_{I'J'}, \tag{1}$$

where

$$\begin{aligned} I &= (i_1, \dots, i_m) \quad & I' &= (i'_1, \dots, i'_n) \\ J &= (j_1, \dots, j_m) \quad & J' &= (j'_1, \dots, j'_n) \end{aligned} \tag{2}$$

are multi-indices,

$$U_{IJ} \overline{U}_{I'J'} = u_{i_1 j_1} \dots u_{i_m j_m} \overline{u_{i'_1 j'_1} \dots u_{i'_n j'_n}} \tag{3}$$

is the corresponding monomial in matrix entries, and only finitely many of the coefficients $c(I,J,I',J') \in \mathbb{C}$ are non-zero. A *polynomial integral* over $U(N)$ is the integral of a polynomial function on $U(N)$ against the normalized Haar measure.

The computation of polynomial integrals over $U(N)$ is of interest from many points of view, including mathematical physics (nuclear physics, lattice gauge theory, quantum transport and quantum information), random matrix theory (matrix models, asymptotic freeness of random matrices), number theory (stochastic models of the Riemann zeta function), and algebraic combinatorics (integral representations of structure constants in the ring of symmetric functions), see [10] for references to the large body of literature on matrix integrals of this type. Nevertheless, the evaluation of such integrals is a problem of substantial complexity that is not yet fully understood.

We wish to develop a general theory of polynomial integrals over $U(N)$. By linearity of the integral, we have

$$\int_{U(N)} p(U)dU = \sum_{m,n\geq 0} \sum_{I,J,I',J'} c(I,J,I',J') \int_{U(N)} U_{IJ}\overline{U}_{I'J'}dU, \tag{4}$$

so we consider the problem of evaluating *monomial integrals*

$$\int_{U(N)} U_{IJ}\overline{U}_{I'J'}dU. \tag{5}$$

Monomial integrals are already of great interest in mathematical physics, see e.g. [3]. An easy argument involving the invariance of Haar measure shows that (5) can be non-zero only for $m = n$ (i.e. the multi-indices I, J are of the same length as the multi-indices I', J'). Furthermore, when $m = n \leq N$ (i.e. the degree of the monomial to be integrated is at most the dimension of the matrices being integrated over), the integral (5) can be decomposed into a double sum over the symmetric group $S(n)$ of the form

$$\int_{U(N)} U_{IJ}\overline{U}_{I'J'}dU = \sum_{(\sigma,\tau)\in S(n)\times S(n)} [I = \sigma(I')][J = \tau(J')]\mathsf{W}_{\sigma\tau}. \tag{6}$$

This integration formula has two ingredients: a combinatorial "Wick-like" rule — sum over pairs of permutations (σ, τ) such that σ maps the multi-index I' to the multi-index I and τ maps the multi-index J' to the multi-index J — together with a certain "weight" $\mathsf{W}_{\sigma\tau}$ associated to each admissible pair of permutations. These weights have a remarkable combinatorial interpretation as generating functions enumerating certain factorizations in the symmetric group; the resulting connections with algebraic combinatorics are the focus of this extended abstract prepared by the authors for FPSAC 2010.

1.2 Primitive factorizations and Weingarten numbers

Let $S(\infty)$ denote the group of finitary permutations of the natural numbers $\{1, 2, 3, \dots\}$, with $S(n) \leq S(\infty)$ the subgroup of permutations of $[1, n] = \{1, \dots, n\}$. An ordered sequence of transpositions

$$(s_1\ t_1)(s_2\ t_2)\dots(s_k\ t_k), \quad s_i < t_i, \tag{7}$$

is said to be a *factorization* of $\pi \in S(\infty)$ if

$$\pi = (s_1\ t_1) \circ (s_2\ t_2) \circ \dots \circ (s_k\ t_k). \tag{8}$$

A factorization is called *primitive* (more precisely, *right primitive*) if the inequalities

$$t_1 \le t_2 \le \cdots \le t_k \tag{9}$$

hold in (8). Consider the quantities

$$\begin{aligned} h_{k,\pi}(n) &= \#\{\text{factorizations of } \pi \text{ into } k \text{ transpositions from } S(n)\} \\ w_{k,\pi}(n) &= \#\{\text{primitive factorizations of } \pi \text{ into } k \text{ transpositions from } S(n)\}. \end{aligned} \tag{10}$$

The numbers $h_{k,\pi}(n)$ are known as (disconnected) *Hurwitz numbers*, and are of much interest in enumerative geometry, see e.g. [12]. We will call the numbers $w_{k,\pi}(n)$ *Weingarten numbers*, see [2, 10] for the origin of this name. The primitive factorizations counted by Weingarten numbers have previously been considered by combinatorialists, both in relation to the enumeration of chains in noncrossing partition lattices [1, 14] and for their own sake [5]. Our approach to polynomial integrals over unitary groups is based on the remarkable fact that the weights appearing in the integration formula (6) are generating functions for Weingarten numbers.

Theorem 1 ([10, 11]) *For any $n \le N$ and $\pi \in S(n)$ we have*

$$N^n \mathsf{W}_{\sigma\tau} = \sum_{k \ge 0} w_{k,\pi}(n) \left(\frac{-1}{N}\right)^k,$$

where $\pi = \sigma \circ \tau^{-1}$.

2 Jucys-Murphy elements

2.1 Centrality

Let $C_\mu \subset S(\infty)$ denote the conjugacy class of permutations of reduced cycle type μ (μ is a Young diagram). For instance, $C_{(0)}$ is the class of the identity permutation, $C_{(1)}$ is the class of transpositions, and more generally $C_{(r)}$ is the class of $(r+1)$-cycles. Note that $|\mu|$ is the minimal length of a factorization of π into transpositions. The conjugacy classes of $S(n)$ are $C_\mu(n) := C_\mu \cap S(n)$. Let $\mathcal{Z}(n)$ denote the centre of the group algebra $\mathbb{C}[S(n)]$. Then $\{C_\mu(n)\}$ is the canonical basis of $\mathcal{Z}(n)$, where $C_\mu(n)$ is identified with the formal sum of its elements, so $\mathcal{Z}(n)$ is referred to as the *class algebra* of $S(n)$.

Multiplying k copies of the class of transpositions, we obtain

$$\underbrace{C_{(1)}(n)C_{(1)}(n)\dots C_{(1)}(n)}_{k \text{ times}} = \sum_\mu h_{k,\mu}(n) C_\mu(n), \tag{11}$$

where clearly $h_{k,\mu}(n) = h_{k,\pi}(n)$ for any $\pi \in C_\mu(n)$. In other words, $h_{k,\pi}(n)$ depends on π only up to conjugacy class. That this also holds for Weingarten numbers is not so obvious. To see that Weingarten numbers are central, we consider the enumeration of *strictly primitive factorizations*, i.e. factorizations

$$\pi = (s_1\ t_1) \circ (s_2\ t_2) \circ \cdots \circ (s_k\ t_k) \tag{12}$$

such that $t_1 < t_2 < \cdots < t_k$. One may show by a direct combinatorial argument that any permutation $\pi \in C_\mu$ admits a *unique* strictly primitive factorization, and that this unique factorization has length $|\mu|$.

This combinatorial fact can be written algebraically as follows. Consider the *Jucys-Murphy* elements $J_1, J_2, \dots, J_t, \dots \in \mathbb{C}[S(\infty)]$ defined by

$$J_t = \sum_{s<t} (s\ t). \tag{13}$$

Let Ξ_n denote the alphabet $\{\{J_1, J_2, \dots, J_n, 0, 0, \dots\}\}$. Then

$$\begin{aligned} e_k(\Xi_n) &= \sum_{\pi \in S(n)} \#\{\text{length } k \text{ strictly primitive factorizations of } \pi\}\pi \\ h_k(\Xi_n) &= \sum_{\pi \in S(n)} \#\{\text{length } k \text{ primitive factorizations of } \pi\}\pi, \end{aligned} \tag{14}$$

where

$$\begin{aligned} e_k &= \sum_{t_1 < t_2 < \dots < t_k} x_{t_1} x_{t_2} \dots x_{t_k} \\ h_k &= \sum_{t_1 \le t_2 \le \dots \le t_k} x_{t_1} x_{t_2} \dots x_{t_k} \end{aligned} \tag{15}$$

are the elementary and complete symmetric functions in commuting variables $x_1, x_2, \dots$. The fact that each $\pi \in C_\mu$ admits a unique strictly primitive factorization, and that this factorization has length π, translates into the identity

$$e_k(\Xi_n) = \sum_{|\mu|=k} C_\mu(n) \in \mathcal{Z}(n), \tag{16}$$

which was first obtained by Jucys [8] (see also [4]). On the other hand, the algebra Λ of symmetric functions is precisely the polynomial algebra $\Lambda = \mathbb{C}[e_1, e_2, \dots]$ in the elementary symmetric functions, so we conclude that the substitution $f \mapsto f(\Xi_n)$ defines a specialization $\Lambda \to \mathcal{Z}(n)$ from the algebra of symmetric functions to the class algebra. In particular, $h_k(\Xi_n) \in \mathcal{Z}(n)$, and we can write

$$h_k(\Xi_n) = \sum_\mu w_{k,\mu}(n) C_\mu(n), \tag{17}$$

where $w_{k,\mu}(n) = w_{k,\pi}(n)$ for any $\pi \in C_\mu(n)$.

2.2 Character theory

Since any permutation is either even or odd, the Hurwitz and Weingarten numbers $h_{k,\mu}(n), w_{k,\mu}(n)$ can only be non-zero for k of the form $k = |\mu| + 2g$ for integer $g \ge 0$. We thus introduce the notation

$$\begin{aligned} \tilde{h}_{g,\mu}(n) &:= h_{|\mu|+2g,\mu}(n) \\ \tilde{w}_{g,\mu}(n) &:= w_{|\mu|+2g,\mu}(n). \end{aligned} \tag{18}$$

In particular, Theorem 1 reads

$$(-1)^{|\mu|} N^{n+|\mu|} \mathsf{W}_{\sigma\tau} = \sum_{g \ge 0} \frac{\tilde{w}_{g,\mu}(n)}{N^{2g}}, \tag{19}$$

where $n \leq N$ and $\sigma \circ \tau^{-1} \in C_\mu(n)$. Using the character theory of $S(n)$, Jackson [7] and Shapiro-Shapiro-Vainshtein [13] obtained the remarkable formula

$$\tilde{h}_{g,(n-1)}(n) = \frac{1}{n!}\sum_{j=0}^{n-1}(-1)^j\binom{n-1}{j}\left(\binom{n}{2}-jn\right)^{n-1+2g} \tag{20}$$

for the number of factorizations of a full cycle (i.e. an element of $C_{(n-1)}(n)$) into $n-1+2g$ transpositions. Here we will explain how properties of Jucys-Murphy elements in irreducible representations of $\mathbb{C}[S(n)]$ may be used to obtain an analogous formula for the Weingarten number $\tilde{w}_{g,(n-1)}(n)$.

Our point of departure is the remarkable expansion

$$f(\Xi_n) = \sum_{\lambda\vdash n}\frac{f(A_\lambda)}{H_\lambda}\chi^\lambda \tag{21}$$

obtained by Jucys [8], of the symmetric function $f \in \Lambda$ evaluated at Ξ_n in terms of the characters

$$\chi^\lambda := \sum_\mu \chi^\lambda(C_\mu(n))C_\mu(n) \tag{22}$$

of the irreducible (complex, finite-dimensional) representations of $\mathbb{C}[S(n)]$. Here A_λ denotes the alphabet of contents of the Young diagram λ, and H_λ is the product of its hook-lengths. This can be viewed as an analogue of the formula of Burnside which expresses the connection coefficients of $\mathcal{Z}(n)$ in terms of irreducible characters.

Consider the ordinary generating function

$$\Phi(z;n) = \sum_{k\geq 0} h_k(\Xi_n)z^k, \tag{23}$$

which is an element of the algebra $\mathcal{Z}(n)[[z]]$ of formal power series in one indeterminate z over the class algebra $\mathcal{Z}(n)$. Plugging in the character expansion

$$h_k(\Xi_n) = \sum_{\lambda\vdash n}\frac{h_k(A_\lambda)}{H_\lambda}\chi^\lambda \tag{24}$$

and changing order of summation, we obtain

$$\Phi(z;n) = \sum_{\lambda\vdash n}\frac{\chi^\lambda}{H_\lambda\prod_{\square\in\lambda}(1-c(\square)z)}, \tag{25}$$

where $c(\square)$ denotes the content of a cell $\square \in \lambda$ and we have made us of the generating function

$$\sum_{k\geq 0} h_k(x_1,x_2,\dots)z^k = \prod_{i\geq 1}\frac{1}{1-x_iz} \tag{26}$$

for the complete symmetric functions. Thus we obtain the formula

$$\Phi_\mu(z;n) = \sum_{\lambda\vdash n}\frac{\chi^\lambda(C_\mu(n))}{H_\lambda\prod_{\square\in\lambda}(1-c(\square)z)} \tag{27}$$

for the ordinary generating function

$$\Phi_{\mu}(z;n) = \sum_{k\geq 0} w_{k,\mu}(n) z^k \tag{28}$$

of Weingarten numbers. Note that by Theorem 1, this corresponds to the character expansion

$$\mathsf{W}_{\sigma\tau} = \sum_{\lambda\vdash n} \frac{\chi^{\lambda}(C_{\mu}(n))}{H_{\lambda}\prod_{\Box\in\lambda}(N+c(\Box))} \tag{29}$$

(where $n \leq N$ and $\sigma\circ\tau^{-1} \in C_{\mu}(n)$), which is well known in the physics literature and was first rigorously obtained in [2] by a different argument.

Up until this point, the partition μ has been generic, but now we restrict to the special case $\mu = (n-1)$, the class of a full cycle in $S(n)$. A classical result from representation theory informs us that the trace of $C_{(n-1)}(n)$ in an irreducible representation can only be non-zero in "hook" representations:

$$\chi^{\lambda}(C_{(n-1)}(n)) = \begin{cases} (-1)^r, \text{ if } \lambda = (n-r,1^r) \\ 0, \text{ otherwise} \end{cases}. \tag{30}$$

Now, the content alphabet of a hook diagram may be obtained immediately,

$$A_{(n-r,1^r)} = \{0,1,\ldots,n-r-1\} \sqcup \{-1,\ldots,-r\}. \tag{31}$$

so that

$$\Phi_{(n-1)}(z;n) = \sum_{r=0}^{n-1} \frac{(-1)^r}{H_{(n-r,1^r)}\prod_{i=1}^{n-r-1}(1-iz)\prod_{j=1}^{r}(1+jz)}. \tag{32}$$

For example, if $n=4$, this is a rational function of the form

$$\begin{aligned}\Phi_{(3)}(z;n) =& \frac{\text{const.}}{(1-z)(1-2z)(1-3z)} + \frac{\text{const.}}{(1-z)(1-2z)(1+z)} \\ &+ \frac{\text{const.}}{(1-z)(1+z)(1+2z)} + \frac{\text{const.}}{(1+z)(1+2z)(1+3z)}.\end{aligned} \tag{33}$$

Thus, as an irreducible rational function, $\Phi_{(n-1)}(z;n)$ has the form

$$\Phi_{(n-1)}(z;n) = \frac{\sum_{i=0}^{n-1} c_i z^i}{\prod_{i=1}^{n-1}(1-i^2z^2)} \tag{34}$$

where $c_0,\ldots,c_{n-1} \in \mathbb{C}$ are some constants to be determined momentarily.

Before finding the above coefficients, let us consider the generating function

$$\frac{1}{\prod_{i=1}^{n}(1-i^2u)} = \sum_{g\geq 0} h_g(1^2,\ldots,n^2)u^g. \tag{35}$$

The coefficients in this generating function are complete symmetric functions evaluated on the alphabet $\{1^2, \dots, n^2\}$ of square integers. Reason dictates that they ought to be close relatives of the Stirling numbers

$$S(n+g, n) = h_g(1, \dots, n). \tag{36}$$

The Stirling number $S(a, b)$ has the following combinatorial interpretation: it counts the number of partitions

$$\{1, \dots, a\} = V_1 \sqcup \dots \sqcup V_b \tag{37}$$

of an a-element set into b disjoint non-empty subsets. Stirling numbers are given by the explicit formula

$$S(a, b) = \sum_{j=0}^{b} (-1)^{b-j} \frac{j^a}{j!(b-j)!}. \tag{38}$$

The numbers

$$T(n+g, n) = h_g(1^2, \dots, n^2) \tag{39}$$

are known as *central factorial numbers*. The central factorial numbers were studied classically by Carlitz and Riordan, see [15, Exercise 5.8] for references. They have the following combinatorial interpretation: $T(a, b)$ counts the number of partitions

$$\{1, 1', \dots, a, a'\} = V_1 \sqcup \dots \sqcup V_b \tag{40}$$

of a set of a marked and a unmarked points into b disjoint non-empty subsets such that[(i)], for each block V_j, if i is the least integer such that either i or i' appears in V_j, then $\{i, i'\} \subseteq V_j$. Central factorial numbers are given by the explicit formula

$$T(a, b) = 2 \sum_{j=0}^{b} (-1)^{b-j} \frac{j^{2a}}{(b-j)!(b+j)!}. \tag{41}$$

Now let us determine the unknown constants $c_0, \dots, c_{n-1}$. By the above discussion, the generating function $\Phi_{(n-1)}(z; n)$ has the form

$$\Phi_{(n-1)}(z; n) = (c_0 + c_1 z + \dots + c_{n-1} z^{n-1}) \sum_{g \geq 0} T(n-1+g, n-1) z^{2g}. \tag{42}$$

On the other hand,

$$\begin{aligned} \Phi_{(n-1)}(z; n) &= \sum_{k \geq 0} w_{k,(n-1)}(n) z^k \\ &= \sum_{g \geq 0} \tilde{w}_{g,(n-1)}(n) z^{n-1+2g} \\ &= \tilde{w}_{0,(n-1)}(n) z^{n-1} + \tilde{w}_{1,(n-1)}(n) z^{n+1} + \dots . \end{aligned} \tag{43}$$

[(i)] Bálint Virág (personal communication) gave a colourful description of this condition, which is actually quite a useful mnemonic: "the most important guy gets to bring his wife."

Consequently, we must have $c_0 = \cdots = c_{n-2} = 0, c_{n-1} = \tilde{w}_{0,(n-1)}(n)$, the number of primitive factorizations of the cyclic permutation $\xi[1,n] = (1\ 2\ \ldots\ n)$ into the minimal number $n-1$ of transpositions. It is not difficult to show (see [5, 10]) bijectively that the number of minimal primitive factorizations of the cycle $\xi[1,n]$ is the Catalan number $\mathrm{Cat}_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}$. In fact, a stronger result from [10] asserts that the number $\tilde{w}_{0,\mu}(n)$ of minimal primitive factorizations of an arbitrary permutation of reduced cycle type μ is a product of Catalan numbers,

$$\tilde{w}_{0,\mu}(n) = \prod_{i=1}^{\ell(\mu)} \mathrm{Cat}_{\mu_i}, \tag{44}$$

so that the function

$$\pi \mapsto \#\{\text{minimal primitive factorizations of } \pi\} \tag{45}$$

is a central multiplicative function on $S(\infty)$ (note that, via Theorem 1, this result corresponds to the first-order estimate

$$(-1)^{|\mu|} N^{n+|\mu|} \mathsf{W}_{\sigma\tau} = \prod_{i=1}^{\ell(\mu)} \mathrm{Cat}_{\mu_i} + O\left(\frac{1}{N^2}\right), \tag{46}$$

where $\sigma \circ \tau^{-1} \in C_\mu(n)$). Thus we have proved the following analogue of (20) for primitive factorizations.

Theorem 2 *For any $g \geq 0$, the number of primitive factorizations of a full cycle from $S(n)$ into $n-1+2g$ transpositions is*

$$\tilde{w}_{g,(n-1)}(n) = \mathrm{Cat}_{n-1} \cdot T(n-1+g, n-1),$$

where $T(a,b)$ denotes the Carlitz-Riordan central factorial number. Equivalently, we have the generating function

$$\Phi_{(n-1)}(z;n) = \frac{\mathrm{Cat}_{n-1}\, z^{n-1}}{(1-1^2z^2)\ldots(1-(n-1)^2z^2)}.$$

Via Theorem 1, Theorem 2 corresponds to the exact integration formula

$$\mathsf{W}_{\sigma\tau} = \frac{(-1)^{n-1}\,\mathrm{Cat}_{n-1}}{N(N^2-1^2)\ldots(N^2-(n-1)^2)}, \quad \sigma \circ \tau^{-1} \in C_\mu(n), \tag{47}$$

which was first stated by Collins in [2].

3 Conclusion

We have discussed the close relationship between the problem of computing polynomial integrals over unitary groups and the enumeration of primitive factorizations of permutations. In particular, the problem was completely solved for full cycles, and the central factorial numbers of Carlitz and Riordan made a surprising appearance and were given a new combinatorial interpretation. It seems that Hurwitz numbers and Weingarten numbers are remarkably similar in character. For example, writing (20) and Theorem 2 in terms of exponential generating functions yields

$$\begin{aligned}\tilde{h}_{g,(n-1)}(n) &= n^{n-2} n^{2g} \binom{n-1+2g}{n-1} \left[\frac{z^{2g}}{(2g)!}\right] \left(\frac{\sinh z/2}{z/2}\right)^{n-1} \\ \tilde{w}_{g,(n-1)}(n) &= \mathrm{Cat}_{n-1} \binom{2n-2+2g}{2n-2} \left[\frac{z^{2g}}{(2g)!}\right] \left(\frac{\sinh z/2}{z/2}\right)^{2n-2}.\end{aligned} \tag{48}$$

On one hand, the multiplicative form of Theorem 2 suggests the existence of an underlying bijective explanation, and on the other computer calculations performed by Valentin Féray (personal communication) suggest that such a bijection could be very complex. Further similarities between Hurwitz numbers and Weingarten numbers are the subject of work in progress [6]. Let us finish by pointing out that the first author has extended many of the results presented here to the setting of polynomial integrals over orthogonal groups [9].

References

[1] P. Biane, Parking functions of types A and B, Electron. J. Combin. 9 (2002) #N7.

[2] B. Collins, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability, Int. Math. Res. Not. 17 (2003) 953–982.

[3] B. De Wit, G. 't Hooft, Nonconvergence of the $1/N$ expansion for $SU(N)$ gauge fields on a lattice, Phys. Lett. B 69 (1977) 61-64.

[4] P. Diaconis and C. Greene, Applications of Murphy's elements, Stanford University Technical Report no.335 (1989) 1-22.

[5] D. A. Gewurz, F. Merola, Some factorisations counted by Catalan numbers, European J. Combin. 27 (2006) 990-994.

[6] I. P. Goulden, M. Guay-Paquet, D. M. R. Jackson, J. Novak, in preparation.

[7] D. M. Jackson, Some combinatorial problems associated with products of conjugacy classes of the symmetric group, J. Combin. Theory Ser. A. 49 (1988) 363-369.

[8] A. Jucys, Symmetric polynomials and the center of the symmetric group ring, Rep. Math. Phys. 5 (1974) 107-112.

[9] S. Matsumoto, Jucys-Murphy elements, orthogonal matrix integrals, and Jack measures arXiv:1001.2345v1 (2010) 35 pages.

[10] S. Matsumoto, J. Novak, *Jucys-Murphy elements and unitary matrix integrals*, arXiv:0905.1992v2 (2009) 43 pages.

[11] J. Novak, Jucys-Murphy elements and the Weingarten function, Banach Center Publ. 89 (2010) 231-235.

[12] A. Okounkov, R. Pandharipande, Gromov-Witten theory, Hurwitz theory, and completed cycles, Ann. of Math. 163 (2006) 517-560.

[13] B. Shapiro, M. Shapiro, A. Vainshtein, Ramified covers of S^2 with one degenerate branching point and the enumeration of edge-ordered graphs, Adv. in Math. Sci. (AMS Transl.) 34 (1997) 219-228.

[14] R. P. Stanley, Parking functions and noncrossing partitions, Electron. J. Combin. 4 (1997) R20.

[15] R. P. Stanley, Enumerative Combinatorics, Vol. 1 & 2, Cambridge Studies in Advanced Mathematics, 49 & 62, Cambridge University Press, Cambridge, 1997 & 1999.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 289–300

Zonotopes, toric arrangements, and generalized Tutte polynomials

Luca Moci[1]

[1] *Università di Roma Tre, Dip. di Matematica, Largo San Leonardo Murialdo, 1, 00146, Roma (Italy).*

Abstract. We introduce a multiplicity Tutte polynomial $M(x, y)$, which generalizes the ordinary one and has applications to zonotopes and toric arrangements. We prove that $M(x, y)$ satisfies a deletion-restriction recurrence and has positive coefficients. The characteristic polynomial and the Poincaré polynomial of a toric arrangement are shown to be specializations of the associated polynomial $M(x, y)$, likewise the corresponding polynomials for a hyperplane arrangement are specializations of the ordinary Tutte polynomial. Furthermore, $M(1, y)$ is the Hilbert series of the related discrete Dahmen-Micchelli space, while $M(x, 1)$ computes the volume and the number of integral points of the associated zonotope.

Résumé. On introduit un polynôme de Tutte avec multiplicité $M(x, y)$, qui généralise le polynôme de Tutte ordinaire et a des applications aux zonotopes et aux arrangements toriques. Nous prouvons que $M(x, y)$ satisfait une récurrence de "deletion-restriction" et a des coefficients positifs. Le polynôme caractéristique et le polynôme de Poincaré d'un arrangement torique sont des spécialisations du polynôme associé $M(x, y)$, de même que les polynômes correspondants pour un arrangement d'hyperplans sont des spécialisations du polynôme de Tutte ordinaire. En outre, $M(1, y)$ est la série de Hilbert de l'espace discret de Dahmen-Micchelli associé, et $M(x, 1)$ calcule le volume et le nombre de points entiers du zonotope associé.

Keywords: Tutte polynomial, zonotope, integral points, toric arrangement, characteristic polynomial, Dahmen-Micchelli, partition function

1 Introduction

The Tutte polynomial is an invariant naturally associated to a matroid and encoding many of its features, such as the number of bases and their *internal and external activity* ([21], [3], [6]). If the matroid is defined by a finite list of vectors, it is natural to consider the arrangement obtained by taking the hyperplane orthogonal to each vector. To the poset of the intersections of the hyperplanes one associates its *characteristic polynomial*, which provides a rich combinatorial and topological description of the arrangement ([19], [22]). This polynomial can be obtained as a specialization of the Tutte polynomial.

Let T be a complex torus (i.e., a multiplicative group $(\mathbb{C}^*)^n$ of n-tuples of nonzero complex numbers) and take a finite list of characters: $X \subset Hom(T, \mathbb{C}^*)$. Then we consider the arrangement of hypersurfaces in T obtained by taking the kernel of each element of the list X. To understand the geometry of this *toric arrangement* one needs to describe the poset $\mathcal{C}(X)$ of the *layers*, i.e. connected components of the intersections of the hypersurfaces ([5], [9], [15], [18]). Clearly this poset depends also on the arithmetics of X, and not only on its linear algebra: for example, the kernel of the identity character λ of $\mathbb{C}^*$ is the point $t = 1$, but the kernel of 2λ has equation $t^2 = 1$, hence is made of two points. Therefore we have no chance to get the characteristic polynomial of $\mathcal{C}(X)$ as a specialization of the ordinary Tutte polynomial $T(x, y)$ of X. In this paper we define a polynomial $M(x, y)$ that specializes to the characteristic polynomial of $\mathcal{C}(X)$ (Theorem 5.5) and to the Poincaré polynomial of the complement $\mathcal{R}_X$ of the toric arrangement (Theorem 5.6). In particular $M(1, 0)$ equals the Euler characteristic of $\mathcal{R}_X$, and also the number of connected components of the complement of the arrangement in the *compact torus* $\overline{T} = (\mathbb{S}^1)^n$.

We call $M(x, y)$ the *multiplicity Tutte polynomial* of X, since it coincides with $T(x, y)$ when X is unimodular, and in general it satisfies the same *deletion-restriction* recurrence that holds for $T(x, y)$. By this formula (Theorem 3.3) we prove that $M(x, y)$ has positive coefficients (Theorem 3.4).

Actually a similar polynomial can be defined more generally for matroids, if we enrich their structure in order to encode some "arithmetic data"; we call such objects *multiplicity matroids*. We hope to develop in a future paper an axiomatic theory of these matroids, as well as applications to graph theory. In the present paper the focus is on the case of a list X of vectors in $\mathbb{Z}^n$. Given such a list, we consider two finite dimensional vector spaces: a space of polynomials $D(X)$, defined by differential equations, and a space of quasipolynomials $DM(X)$, defined by difference equations. These spaces were introduced by Dahmen and Micchelli to study respectively *box splines* and *partition functions*, and are deeply related respectively with the hyperplane arrangement and the toric arrangement defined by X, as explained in the forthcoming book [6]. In particular, $T(1, y)$ is known to be the Hilbert series of $D(X)$; then we prove that $M(1, y)$ is the Hilbert series of $DM(X)$ (Theorem 6.3).

On the other hand, by Theorem 4.1 the coefficients of $M(x, 1)$ count integral points in some faces of a convex polytope, the *zonotope* defined by X. The relations between arrangements, zonotopes and Dahmen-Micchelli spaces is being studied intensively in the very last years: see for example [6], [10], [7], [1], [11], . In particular $M(1, 1)$ equals the volume of the zonotope (Proposition 2.1), while $M(2, 1)$ is the number of its integral points (Proposition 4.2).

Finally we focus on the case in which X is a root system: then we show some connections with the theory of Weyl groups (see for instance Corollary 7.3).

Remark 1.1 *This paper is an extended abstract of [17], which contains more details and all the proofs, which are omitted here.*

2 Multiplicity matroids and multiplicity Tutte polynomials

We start recalling the notions we are going to generalize.

A *matroid* $\mathfrak{M}$ is a pair (X, I), where X is a finite set and I is a family of subsets of X (called the *independent sets*) with the following properties:

1. The empty set is independent;
2. Every subset of an independent set is independent;
3. Let A and B be two independent sets and assume that A has more elements than B. Then there exists an element $a \in A \setminus B$ such that $B \cup \{a\}$ is still independent.

A maximal independent set is called a *basis*. The last axiom implies that all bases have the same cardinality, which is called the *rank* of the matroid. Every $A \subseteq X$ has a natural structure of matroid, defined by considering a subset of A independent if and only if it is in I. Then each $A \subseteq X$ has a rank which we denote by $r(A)$.

The *Tutte polynomial* of the matroid is then defined as

$$T(x,y) \doteq \sum_{A \subseteq X} (x-1)^{r(X)-r(A)}(y-1)^{|A|-r(A)}.$$

From the definition it is clear that $T(1,1)$ equals the number of bases of the matroid.

In the next sections we will recall the main example of matroid and some properties of its Tutte polynomial.

We now introduce the following definitions.

A *multiplicity matroid* $\mathfrak{M}$ is a triple (X, I, m), where (X, I) is a matroid and m is a function (called *multiplicity*) from the family of all subsets of X to the positive integers.

We say that m is the *trivial multiplicity* if it is identically equal to 1.

We define the *multiplicity Tutte polynomial* of a multiplicity matroid as

$$M(x,y) \doteq \sum_{A \subseteq X} m(A)(x-1)^{r(X)-r(A)}(y-1)^{|A|-r(A)}.$$

Let us remark that we can endow every matroid with the trivial multiplicity, and then $M(x,y) = T(x,y)$.

Let X be a finite list of vectors spanning a real vector space U, and I be the family of its linearly independent subsets; then (X, I) is a matroid, and the rank of a subset A is just the dimension of the spanned subspace. We denote by $T_X(x,y)$ the associated Tutte polynomial.

We associate to the list X a *zonotope*, that is a convex polytope in U defined as follows:

$$\mathcal{Z}(X) \doteq \left\{ \sum_{x \in X} t_x x, 0 \le t_x \le 1 \right\}.$$

Zonotopes play an important role in the theory of hyperplane arrangements, and also in that of *splines*, a class of functions studied in Approximation Theory. (see [6]).

We recall that a *lattice* Λ of rank n is a discrete subgroup of $\mathbb{R}^n$ which spans the real vector space $\mathbb{R}^n$. Every such Λ can be generated from some basis of the vector space by forming all linear combinations with integral coefficients; hence the group Λ is isomorphic to $\mathbb{Z}^n$. We will use the word *lattice* always with this meaning, and not in the combinatorial sense (poset with *join* and *meet*).

Then let X be a finite list of elements in a lattice Λ, and let I and r be as above. We denote by $\langle A \rangle_{\mathbb{Z}}$ and $\langle A \rangle_{\mathbb{R}}$ respectively the sublattice of Λ and the subspace of $\Lambda \otimes \mathbb{R}$ spanned by A. Let us define $\Lambda_A \doteq \Lambda \cap \langle A \rangle_{\mathbb{R}}$: this is the largest sublattice of Λ in which $\langle A \rangle_{\mathbb{Z}}$ has finite index. Then we define m as this index:

$$m(A) \doteq [\Lambda_A : \langle A \rangle_{\mathbb{Z}}].$$

This defines a multiplicity matroid and then a multiplicity Tutte polynomial $M_X(x,y)$, which is the main subject of this paper. We start by showing the relations with the zonotope $\mathcal{Z}(X)$ generated by X in $U \doteq \Lambda \otimes \mathbb{R}$.

We already observed that $T_X(1,1)$ equals the number of bases that can be extracted from X; on the other hand we have:

Proposition 2.1 *$M_X(1,1)$ equals the volume of the zonotope $\mathcal{Z}(X)$.*

Further relations between the polynomial $M_X(x,y)$ and the zonotope $\mathcal{Z}(X)$ will be shown in Section 4.

3 Deletion-restriction formula and positivity

The central idea that inspired Tutte in defining the polynomial $T(x,y)$, was to find the most general invariant satisfying a recurrence known as *deletion-restriction*. Such recurrence allows to reduce the computation of the Tutte polynomial to some trivial cases. We will explain this algorithm in the case above, i.e. when the matroid is defined by a list of vectors, and we will show that in this case also the polynomial $M(x,y)$ satisfies a similar recursion.

3.1 Lists of vectors

Let X be a finite list of elements spanning a vector space U, and let $v \in X$ be a nonzero element. We define two new lists: the list $X_1 \doteq X \setminus \{v\}$ of elements of U and the list X_2 of elements of $U/\langle v\rangle$ obtained by reducing X_1 modulo v. Assume that v is dependent in X, i.e. $v \in \langle X_1\rangle_{\mathbb{R}}$. Then we have the following well-known formula:

Theorem 3.1

$$T_X(x,y) = T_{X_1}(x,y) + T_{X_2}(x,y)$$

It is now clear why we defined X as a list, and not as a set: even if we start with X made of (nonzero) distinct elements, in X_2 some vector may appear many times (and some vector may be zero).

By this recurrence we get:

Theorem 3.2 *$T_X(x,y)$ is a polynomial with positive coefficients.*

3.2 Lists of elements in finitely generated abelian groups.

We now want to show a similar recursion for the polynomial $M_X(x,y)$. Inspired by [8], we notice that in order to do this, we need to work in a larger category. Indeed, whereas the quotient of a vector space by a subspace is still a vector space, the quotient of a lattice by a sublattice is not a lattice, but a *finitely generated abelian group*. For example in the 1-dimensional case, the quotient of $\mathbb{Z}$ by $m\mathbb{Z}$ is the cyclic group of order m.

Then let Γ be a finitely generated abelian group. For every subset S of Γ we denote by $\langle S\rangle$ the generated subgroup. We recall that Γ is isomorphic to the direct product of a lattice Λ and of a finite group Γ_t, which is called the *torsion subgroup* of Γ. We denote by π the projection $\pi : \Gamma \to \Lambda$.

Let X be a finite subset of Γ; for every $A \subseteq X$ we set $\Lambda_A \doteq \Lambda \cap \big\langle \pi(A)\big\rangle_{\mathbb{R}}$ and $\Gamma_A \doteq \Lambda_A \times \Gamma_t$. In other words, Γ_A is the largest subgroup of Γ in which $\langle A\rangle$ has finite index.

Then we define $m(A) \doteq \big[\Gamma_A : \langle A\rangle\big]$. We also define $r(A)$ as the rank of $\pi(A)$. In this way we defined a multiplicity matroid, to which is associated a multiplicity Tutte polynomial:

$$M_X(x,y) \doteq \sum_{A\subseteq X} m(A)(x-1)^{r(X)-r(A)}(y-1)^{|A|-r(A)}.$$

It is clear that if Γ is a lattice, these definitions coincide with the ones given in the previous sections.

If on the opposite hand Γ is a finite group, $M(x,y)$ is a polynomial in which only the variable y appears; furthermore this polynomial, evaluated at $y=1$, gives the order of Γ. Indeed the only summand that does not vanish is the contribution of the empty set, which generates the trivial subgroup.

Now let $\lambda \in X$ be a nonzero element such that $\pi(\lambda) \in \big\langle \pi\big(X \setminus \{\lambda\}\big)\big\rangle_{\mathbb{R}}$. We set $X_1 \doteq X \setminus \{\lambda\} \subset \Gamma$ and we denote by $\overline{A}$ the image of every $A \subseteq X$ under the natural projection $\Gamma \longrightarrow \Gamma/\langle\lambda\rangle$. We denote by X_2 the subset $\overline{X_1}$ of $\Gamma/\langle\lambda\rangle$. Then we have the following deletion-restriction formula.

Theorem 3.3

$$M_X(x,y) = M_{X_1}(x,y) + M_{X_2}(x,y).$$

By this recurrence we prove:

Theorem 3.4 *$M_X(x,y)$ is a polynomial with positive coefficients.*

4 Integral points in zonotopes

Let X be a finite list of vectors contained in a lattice Λ and generating the vector space $U = \Lambda \otimes \mathbb{R}$. We say that a point of U is *integral* if it is contained in Λ. In this section we prove that $M_X(2,1)$ equals the number of integral points of the zonotope $\mathcal{Z}(X)$. Moreover we compare this number with the volume. In order to do that, we have to move the zonotope to a "generic position"; we proceed as follows. Following [6, Section 1.3], we define the *cut-locus* of the couple (Λ, X) as the union of all hyperplanes in U that are translations, under elements of Λ, of the linear hyperplanes spanned by subsets of X. Then let $\underline{\varepsilon}$ be a vector of U which does not lie in the cut-locus and has length $\varepsilon << 0$. Let $\mathcal{Z}(X) - \underline{\varepsilon}$ be the polytope obtained translating $\mathcal{Z}(X)$ by $-\underline{\varepsilon}$, and let $\mathfrak{I}(X)$ be the set of its integral points:

$$\mathfrak{I}(X) \doteq (\mathcal{Z}(X) - \underline{\varepsilon}) \cap \Lambda.$$

It is intuitive (and proved in [6, Prop 2.50]) that this number equals the volume:

$$|\mathfrak{I}(X)| = vol\,(\mathcal{Z}(X)) = M_X(1,1)$$

by Proposition 2.1. We now prove a stronger result. Let us choose $\underline{\varepsilon}$ so that $\mathcal{Z}(X) - \underline{\varepsilon}$ contanins the origin $\underline{0}$. We partition $\mathfrak{I}(X)$ as follows: set $\mathfrak{I}_n(X) = \{\underline{0}\}$, and for every $k = n-1, \ldots, 0$, let $\mathfrak{I}_k(X)$ be the set of elements of $\mathfrak{I}(X)$ that are contained in some $k-$codimensional face of $\mathcal{Z}(X)$ and that are not contained in $\mathfrak{I}_h(X)$ for $h > k$.

Then we have:

Theorem 4.1

$$M_X(x,1) = \sum_{k=0}^{n} |\mathfrak{I}_k(X)|\; x^k.$$

Furthermore we prove:

Proposition 4.2

$$M_X(2,1) = |\mathcal{Z}(X) \cap \Lambda|$$

Example 4.3 *Consider the list in* $\mathbb{Z}^2$

$$X = \{(3,3),(1,-1),(2,0)\}.$$

Then

$$M_X(x,y) = (x-1)^2 + (3+1+2)(x-1) + (6+6+2) + 2(y-1).$$

Hence

$$M_X(x,1) = x^2 + 4x + 9$$

and $M_X(2,1) = 21$. *Indeed the zonotope* $\mathcal{Z}(X)$ *has volume* 14 *and contains* 21 *integral points, 14 of which lying in* $\mathcal{Z}(X) - \underline{\varepsilon}$. *The sets* $\mathfrak{I}_2(X)$, $\mathfrak{I}_1(X)$, *and* $\mathfrak{I}_0(X)$ *contain 1, 4 and 9 points respectively.*

5 Application to arrangements

In this Section we describe some geometrical objects related to the lists considered in Section 2.2, and show that many of their features are encoded in the polynomials $T_X(x,y)$ and $M_X(x,y)$.

5.1 Recall on hyperplane arrangements

Let X be a finite list of elements of a vector space U. Then in the dual space $V = U^*$ a *hyperplane arrangement* $\mathcal{H}(X)$ is defined by taking the orthogonal hyperplane of each element of X. Conversely, given an arrangement of hyperplanes in a vector space V, let us choose for each hyperplane a nonzero vector in V^* orthogonal to it; let X be the list of such vectors. Since every element of X is determined up to scalar multiples, the matroid associated to X is well defined; in this way a Tutte polynomial is naturally associated to the hyperplane arrangement.

The importance of the Tutte polynomial in the theory of hyperplane arrangements is well known. Here we just recall some results that we generalize in the next sections.

To every sublist $A \subseteq X$ is associated the subspace $A^\perp$ of V that is the intersection of the corresponding hyperplanes of $\mathcal{H}(X)$; in other words, $A^\perp$ is the subspace of vectors that are orthogonal to every element of A. Let $\mathcal{L}(X)$ be the set of such subspaces, partially ordered by reverse inclusion, and having as minimal element **0** the whole space $V = \emptyset^\perp$. $\mathcal{L}(X)$ is called the *intersection poset* of the arrangement, and is "the most important combinatorial object associated to a hyperplane arrangement" (R. Stanley).

We also recall that to every finite poset $\mathcal{P}$ is associated a *Moebius function* $\mu : \mathcal{P} \times \mathcal{P} \to \mathbb{Z}$, recursively defined as follows:

$$\mu(L,M) = \begin{cases} 0 & \text{if } L > M \\ 1 & \text{if } L = M \\ -\sum_{L \leq N < M} \mu(L,N) & \text{if } L < M. \end{cases}$$

Notice that the poset $\mathcal{L}(X)$ is *ranked* by the dimension of the subspaces; then we define *characteristic polynomial* of the poset as

$$\chi(q) \doteq \sum_{L \in \mathcal{L}(X)} \mu(\mathbf{0}, L) q^{dim(L)}.$$

This is an important invariant of $\mathcal{H}(X)$. Indeed, let $\mathcal{M}_X$ be the complement in V of the union of the hyperplanes of $\mathcal{H}(X)$. Let $P(q)$ be *Poincaré polynomial* of $\mathcal{M}_X$, i.e. the polynomial having as coefficient of q^k the $k-$th Betti number of $\mathcal{M}_X$. Then if V is a complex vector space, by [19] we have the following theorem.

Theorem 5.1

$$P(q) = (-q)^n \chi(-1/q).$$

If on the other hand V is a real vector space, by [22] the number $Ch(X)$ of *chambers* (i.e., connected components of $\mathcal{M}_X$) is:

Theorem 5.2

$$Ch(X) = (-1)^n \chi(-1).$$

The Tutte polynomial $T_X(x, y)$ turns out to be a stronger invariant, in the following sense. Assume that $\underline{0} \notin X$; then

Theorem 5.3

$$(-1)^n T_X(1-q, 0) = \chi(q).$$

The proof of these theorems can be found for example in [6, Theorems 10.5, 2.34 and 2.33].

5.2 Toric arrangements and their generalizations

Let $\Gamma = \Lambda \times \Gamma_t$ be a finitely generated abelian group, and define $T_\Gamma \doteq Hom(\Gamma, \mathbb{C}^*)$. T_Γ has a natural structure of abelian linear algebraic group: indeed it is the direct product of a complex torus T_Λ of the same rank as Λ and of the finite group ${\Gamma_t}^*$ dual to Γ_t (and isomorphic to it).

Moreover Γ is identified with the group of characters of T_Γ: indeed given $\lambda \in \Lambda$ and $t \in T_\Gamma$ we can take any representative $\varphi_t \in Hom(\Gamma, \mathbb{C})$ of t and set $\lambda(t) \doteq e^{2\pi i \varphi_t(\lambda)}$. When this is not ambiguous we will denote T_Γ by T.

Let $X \subset \Lambda$ be a finite subset spanning a sublattice of Λ of finite index. The kernel of every character $\chi \in X$ is a (non-connected) hypersurface in T:

$$H_\chi \doteq \big\{t \in T | \chi(t) = 1\big\}.$$

The collection $\mathcal{T}(X) = \{H_\chi, \chi \in X\}$ is called the *generalized toric arrangement* defined by X on T.

We denote by $\mathcal{R}_X$ the complement of the arrangement:

$$\mathcal{R}_X \doteq T \setminus \bigcup_{\chi \in X} H_\chi$$

and by $\mathcal{C}_X$ the set of all the connected components of all the intersections of the hypersurfaces H_χ, ordered by reverse inclusion and having as minimal elements the connected components of T.

Since $rank(\Lambda) = dim(T)$, the maximal elements of $\mathcal{C}(X)$ are 0-dimensional, hence (since they are connected) they are points. We denote by $\mathcal{C}_0(X)$ the set of such layers, which we call the *points* of the arrangement.

Given $A \subseteq X$ let us define $H_A \doteq \bigcap_{\lambda \in A} H_\lambda$. Then we have:

Lemma 5.4 *$m(A)$ equals the number of connected components of H_A.*

In particular, when Γ is a lattice, T is a torus and $\mathcal{T}(X)$ is called the *toric arrangement* defined by X. Such arrangements have been studied for example in [14], [5], [15], [18]; see [6] for a complete reference. In particular, the complement $\mathcal{R}_X$ has been described topologically and geometrically. In this description the poset $\mathcal{C}(X)$ plays a major role, for many aspects analogous to that of the intersection poset for hyperplane arrangements (see [5], [18]).

We will now explain the importance in this framework of the polynomial $M_X(x, y)$ defined in Section 3.3.

5.3 Characteristic polynomial and Poincaré polynomial

Let μ be the Moebius function of $\mathcal{C}(X)$; notice that we have a natural rank function given by the dimension of the layers. For every $C \in \mathcal{C}(X)$, let T_C be the connected component of T that contains C. Then we define the characteristic polynomial of $\mathcal{C}(X)$:

$$\chi(q) \doteq \sum_{C \in \mathcal{C}(X)} \mu(T_C, C) q^{dim(C)}.$$

This polynomial is a specialization of the multiplicity Tutte polynomial:

Theorem 5.5

$$(-1)^n M_X(1-q, 0) = \chi(q)$$

Furthermore, by applying our results to a theorem proved in [5, Theor. 4.2] (or [6, 14.1.5]), we give a formula for the Poincare' polynomial $P(q)$ of $\mathcal{R}_X$:

Theorem 5.6

$$P(q) = q^n M_X \left(\frac{2q+1}{q}, 0 \right).$$

Therefore, by comparing Theorem 5.5 and Theorem 5.6, we get the following formula, which relates the combinatorics of $\mathcal{C}(X)$ with the topology of $\mathcal{R}_X$, and is the "toric" analogue of Theorem 4.1.

Corollary 5.7

$$P(q) = (-q)^n \chi \left(-\frac{q+1}{q} \right).$$

We recall that the *Euler characteristic* of a space can be defined as the evaluation at -1 of its Poincaré polynomial. Hence by Theorem 5.6 we have:

Corollary 5.8 $(-1)^n M_X(1, 0)$ *equals the Euler characteristic of* $\mathcal{R}_X$.

Example 5.9 *Take* $T = (\mathbb{C}^*)^2$ *with coordinates* (t, s) *and*

$$X = \{(2,0), (0,2), (1,1), (1,-1)\}$$

defining equations:

$$t^2 = 1, s^2 = 1, ts = 1, ts^{-1} = 1.$$

It is easily seen (see [17] for details) that this arrangement has six $1-$*dimensional layers and four* $0-$*dimensional layers, and that*

$$\chi(q) = q^2 - 6q + 8.$$

The polynomial $M_X(x, y)$ *is composed by the following summands:*

- $(x-1)^2$*, corresponding to the empty set;*
- $6(x-1)$*, corresponding to the 4 singletons, each giving contribution* $(x-1)$ *or* $2(x-1)$*;*
- 14*, corresponding to the 6 pairs: indeed, the basis* $X = \{(2,0), (0,2)\}$ *spans a sublattice of index* 4*, while the other bases span sublattices of index* 2*;*

- $8(y-1)$*, corresponding to the 4 triples, each contributing with* $2(y-1)$*;*
- $2(y-1)^2$*, corresponding to the whole set* X*.*

Hence

$$M_X(x,y) = x^2 + 2y^2 + 4x + 4y + 3.$$

Notice that

$$M_X(1-q,0) = q^2 - 6q + 8 = \chi(q)$$

as claimed in Theorem 5.5. Furthermore Theorem 5.6 (or Corollary 4.12) implies that

$$P(q) = 15q^2 + 8q + 1$$

and hence the Euler characteristic is $P(-1) = 8 = M_X(1,0)$*. Notice that this is the toric arrangement arising from the root system of type* C_2 *(see Section 7).*

5.4 Number of regions of the compact torus

In this section we consider the compact abelian group dual to Γ $\overline{T} \doteq Hom(\Gamma, \mathbb{S}^1)$, where we set $\mathbb{S}^1 \doteq \{z \in \mathbb{C} \mid |z| = 1\} \simeq \mathbb{R}/\mathbb{Z}$.

We assume for simplicity Γ to be a lattice; then $\overline{T}$ is a *compact torus*, i.e. it is isomorphic to $(\mathbb{S}^1)^n$, and in it every $\chi \in X$ defines a hypersurface $\overline{H_\chi} \doteq \left\{t \in \overline{T} | \chi(t) = 1\right\}$. We denote by $\overline{\mathcal{T}(X)}$ this arrangement; clearly its poset of layers is the same as for the arrangement $\mathcal{T}(X)$ defined in the complex torus T. We denote by $\overline{\mathcal{R}_X}$ the complement

$$\overline{\mathcal{R}_X} \doteq \overline{T} \setminus \bigcup_{\chi \in X} \overline{H_\chi}.$$

The compact toric arrangement $\overline{\mathcal{T}(X)}$ has been studied in [9]; in particular the number $R(X)$ of *region*s (i.e. of connected components) of $\overline{\mathcal{R}_X}$ is proved to be a specialization of the characteristic polynomial $\chi(q)$:

Theorem 5.10

$$R(X) = (-1)^n \chi(0).$$

By comparing this result with Theorem 5.5 we get the following

Corollary 5.11

$$R(X) = M_X(1,0)$$

6 Dahmen-Micchelli spaces

Until now we considered evaluations of $T_X(x,y)$ and $M_X(x,y)$ at $y = 0$ and $y = 1$. However, there is another remarkable specialization of the Tutte polynomial: $T_X(1,y)$, which is called the *polynomial of the external activity* of X. It is related with the corresponding specialization of $M_X(x,y)$ in a simple way:

Lemma 6.1

$$M_X(1,y) = \sum_{p \in \mathcal{C}_0(X)} T_{X_p}(1,y).$$

The previous lemma has an interesting consequence. In [4] to every finite set $X \subset V$ is associated a space $D(X)$ of functions $V \to \mathbb{C}$, and to every finite set $X \subset \Lambda$ is associated a space $DM(X)$ of functions $\Lambda \to \mathbb{C}$. Such spaces are defined as the solutions of a system, respectively of differential equations and of difference equations, in the following way.

For every $\lambda \in X$, let ∂_λ be the usual directional derivative $\partial_\lambda f(x) \doteq \partial f/\partial\lambda(x)$ and let ∇_λ be the difference operator $\nabla_\lambda f(x) \doteq f(x) - f(x-\lambda)$.

Then for every $A \subset X$ we define the differential operator $\partial_A \doteq \prod_{\lambda\in A} \partial_\lambda$ and the difference operator $\nabla_A \doteq \prod_{\lambda\in A} \nabla_\lambda$. We can now define define the *differentiable Dahmen-Micchelli space*

$$D(X) \doteq \{f : V \to \mathbb{C} \mid \partial_A f = 0 \,\forall A \text{ such that } r(X \setminus A) < n\}$$

and the *discrete Dahmen-Micchelli space*

$$DM(X) \doteq \{f : \Lambda \to \mathbb{C} \mid \nabla_A f = 0 \,\forall A \text{ such that } r(X \setminus A) < n\}\,.$$

The space $D(X)$ is a space of polynomials, which was introduced in order to study the *box spline*. This is a piecewise-polynomial function studied in Approximation Theory; its local pieces, together with their derivatives, span $D(X)$. On the other hand, $DM(X)$ is a space of quasipolynomials which arises in the study of the *partition function*. This is the function that counts in how many ways an element of Λ can be written as a linear combination with positive integer coefficients of elements of X. This function is piecewise-quasipolynomial, and its local pieces, together with their translates, span $DM(X)$. In the recent book [6] the spaces $D(X)$ and $DM(X)$ are shown to be deeply related respectively with the hyperlane arrangement and with the toric arrangement defined by X.

In order to compare these two spaces, we consider the elements of $D(X)$ as functions $\Lambda \to \mathbb{C}$ by restricting them to the lattice Λ. Since the elements of $DM(X)$ are polynomial functions, they are determined by their restriction. For every $p \in \mathcal{C}(X)^0$, let us define $\varphi_p : \Lambda \to \mathbb{C}$ as the map $\lambda \mapsto \lambda(p)$. (see Section 2.4.2). In [4] (see also [6, Formula 16.1]) the following result is proved.

Theorem 6.2

$$DM(X) = \bigoplus_{p\in\mathcal{C}_0(X)} \varphi_p D(X_p).$$

Since every $D(X_p)$ is defined by homogeneous differential equations, it is naturally graded, the degree of every element being just its degree as a polynomial. The Hilbert series of $D(X_p)$ is known to be $T_{X_p}(1, y)$; in other words, the coefficients of this polynomial equal the dimensions of the graded parts (see [2] or [6, Theorem 11.8]). Then, by the theorem above, also the space $DM(X)$ is graded, and by Lemma 6.1 we have:

Theorem 6.3 *$M_X(1, y)$ is the Hilbert series of $DM(X)$.*

By comparing this theorem with Proposition 2.1 we recover the following known result, which can be found for example in ([6, Chapter 13]) :

Corollary 6.4 *The dimension of $DM(X)$ equals the volume of the zonotope $\mathcal{Z}(X)$.*

7 The case of root systems

This section is devoted to describe a remarkable class of examples. We will assume standard notions about root systems, Lie algebras and algebraic groups, which are exposed for example in [13] and [12].

Let Φ be a root system, $\langle \Phi^\vee \rangle$ be the lattice spanned by the coroots, and Λ be its dual lattice (which is called the *cocharacters* lattice). Then we define as in Section 4.2 a torus $T = T_\Lambda$ having Λ as group of characters. In other words, if $\mathfrak{g}$ is the semisimple complex Lie algebra associated to Φ and $\mathfrak{h}$ is a Cartan subalgebra, T is defined as the quotient $T \doteq \mathfrak{h}/\langle \Phi^\vee \rangle$.

Each root α takes integer values on $\langle \Phi^\vee \rangle$, so it induces a character $e^\alpha : T \to \mathbb{C}/\mathbb{Z} \simeq \mathbb{C}^*$. Let X be the set of this characters; more precisely, since α and $-\alpha$ define the same hypersurface, we set

$$X \doteq \left\{ e^\alpha, \ \alpha \in \Phi^+ \right\}.$$

In this way to every root system Φ is associated a toric arrangement. These arrangements have been studied in [15]; we now show two applications to the present work. Let W be the (finite) Weyl group of Φ, and let $\widetilde{W}$ be the associated affine Weyl group. We denote by $s_0, \dots, s_n$ its generators, and by W_k the subgroup of $\widetilde{W}$ generated by all the elements s_i but s_k. Let $\Phi_k \subset \Phi$ be the root system of W_k, and denote by X_k the corresponding sublist of X. Then we have:

Corollary 7.1

$$M_X(1, y) = \sum_{k=0}^{n} \frac{|W|}{|W_k|} T_{X_k}(1, y).$$

Furthermore, in [15] the following theorem is proved. Let W be the Weyl group of Φ.

Theorem 7.2 *The Euler characteristic of $\mathcal{R}_X$ is equal to $(-1)^n |W|$.*

By comparing this statement with Corollary 5.8, we get the following

Corollary 7.3

$$M_X(1, 0) = |W|.$$

It would be interesting to have a more direct proof of this fact.

Acknowledgements

I am grateful to my advisor Corrado De Concini for many important suggestions. I also wish to thank Emanuele Delucchi, Olga Holtz, Mario Marietti, Claudio Procesi and Michèle Vergne for stimulating discussions and remarks. .

References

[1] FEDERICO ARDILA, ALEXANDER POSTNIKOV, *Combinatorics and geometry of power ideals*, Trans. Amer. Math. Soc., to appear.

[2] C. DE BOOR, N. DYN, A. RON, *On two polynomial spaces associated with a box spline*, Pacific J. Math., 147 (2): 249-267, 1991.

[3] H. H. CRAPO, *The Tutte polynomial*, Aequationes Math., 3: 211-229, 1969.

[4] W. Dahmen and C. A. Micchelli, *The number of solutions to linear Diophantine equations and multivariate splines*, Trans. Amer. Math. Soc., 308(2): 509-532, 1988.

[5] C. De Concini, C. Procesi, *On the geometry of toric arrangements*, Transformations Groups 10, N. 3-4, 2005.

[6] C. De Concini, C. Procesi, *Topics in hyperplane arrangements, polytopes and box-splines*, to appear, available on www.mat.uniroma1.it/people/procesi/dida.html.

[7] C. De Concini, C. Procesi, M. Vergne, *Partition function and generalized Dahmen-Micchelli spaces*, arXiv:math 0805.2907.

[8] C. De Concini, C. Procesi, M. Vergne, *Vector partition functions and index of transversally elliptic operators*, arXiv:0808.2545v1

[9] R. Ehrenborg, M. Readdy, M. Slone, *Affine and toric hyperplane arrangements*, arXiv:0810.0295v1 (math.CO), 2008.

[10] Olga Holtz, Amos Ron, *Zonotopal Algebra*, arXiv:0708.2632v2.

[11] Olga Holtz, Amos Ron, Zhiqiang Xu, *Hierarchical zonotopal spaces*, arXiv:0910.5543v2.

[12] J.E. Humphreys, *Linear Algebraic Groups*, Springer-Verlag, 1975.

[13] J.E. Humphreys, *Introduction to Lie Algebras and Representation theory*, Springer-Verlag, 3rd reprint, 1975.

[14] G. I. Lehrer, *The cohomology of the regular semisimple variety*, J. Algebra 199 (1998), no. 2, 666-689.

[15] L. Moci, *Combinatorics and topology of toric arrangements defined by root systems*, Rend. Lincei Mat. Appl. 19 (2008), 293-308.

[16] L. Moci, *Geometry and Combinatorics of toric arrangements*, Tesi di Dottorato (Ph. D. Thesis), Universitá di Roma Tre, March 2010.

[17] L. Moci, *A Tutte polynomial for toric arrangements*, arXiv:0911.4823 [math.CO].

[18] L. Moci, *Wonderful models for toric arrangements*, arXiv:0912.5461 [math.AG].

[19] P. Orlik, L. Solomon, *Combinatorics and topology of complements of hyperplanes*, Invent. Math. 56, no. 2 (1980), 167-189.

[20] G. C. Shephard, *Combinatorial properties of associated zonotopes*, Canad. J. Math., 26: 302-321, 1974.

[21] W. T. Tutte, *A contribution to the theory of chromatic polynomials*,Canadian J. Math., 6: 80-91, 1954.

[22] T. Zaslavsky, *Facing up to arrangements: face-count formulas for partitions of space by hyperplanes*, Mem. Amer. Math. Soc., 1(154): vii+102, 1975.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 301–312

Fully Packed Loop configurations in a triangle and Littlewood Richardson coefficients

Philippe Nadeau

Fakultät für Mathematik, Universität Wien, Nordbergstraße 15, A-1090 Wien, AUSTRIA.

Abstract. We are interested in Fully Packed Loops in a triangle (TFPLs), as introduced by Caselli at al. and studied by Thapper. We show that for Fully Packed Loops with a fixed link pattern (refined FPL), there exist linear recurrence relations with coefficients computed from TFPL configurations. We then give constraints and enumeration results for certain classes of TFPL configurations. For special boundary conditions, we show that TFPLs are counted by the famous Littlewood Richardson coefficients.

Résumé. Nous nous intéressons aux configurations de "Fully Packed Loops" dans un triangle (TFPL), introduites par Caselli et al. et étudiées par Thapper. Nous montrons que pour les Fully Packed Loops avec un couplage donné, il existe des relations de récurrence linéaires dont les coefficients sont calculés à partir de certains TFPLs. Nous donnons ensuite des contraintes et des résultats énumératifs pour certaines familles de TFPLs. Pour certaines conditions au bord, nous montrons que le nombre de TFPL est donné par les coefficients de Littlewood Richardson.

Keywords: Razumov Stroganov conjecture, Fully Packed Loop, Littlewood–Richardson coefficients

1 Introduction

The recently proved Razumov-Stroganov correspondence [RS04, CS10] states that the ground state components ψ_π of the so called *O(1) loop model* are equal to the refined Fully Packed Loop number A_π, where π is a link pattern (see Section 1.1 for definitions on FPLs). Although certain general expressions have been developed for the ψ_π's from which results could be obtained (see [ZJ] and references therein), explicit formulas for the A_π's are known only in certain very special cases of link patterns (cf. [ZJ06]).

The purpose of this article is to study the numbers A_π thanks to the decomposition found in [CKLN06] which involves the counting of FPLs in a triangle (TFPLs). More recently, the paper [Tha07] developed new ideas and conjectures concerning these TFPLs, and was the original motivation for the present paper. We will actually first prove a conjecture of [Tha07] about certain recurrence relations for the numbers A_π that involve coefficients computed from TFPLs. Then we will start the study of TFPL configurations themselves, gathering several of their properties, the most striking being Theorem 4.3 which shows that a certain subclass of TFPLs turns out to be enumerated by Littlewood Richardson coefficients.

This work is thus a starting point in the study of TFPLs. Our results will show that these are not only interesting by themselves, but are also a promising tool in order to obtain explicit recurrences or expressions for the refined FPL numbers A_π.

In the rest of this section we define FPL configurations, and notions related to words and partitions. In Section 2 FPLs in a triangle are defined, and we prove Theorem 2.5 about linear recurrence relations for the numbers A_π. In Section 3, we prove certain properties and constraints of TFPL numbers, giving in particular a very nice new proof of Theorem 3.1 from [CKLN06]. Finally, we prove Theorem 4.3 mentioned above in Section 4.

1.1 Fully Packed Loop configurations

We fix a positive integer n, and let G_n be the square grid with n^2 vertices; we impose also *periodic boundary conditions* on G_n, which means that we select every other external edge on the grid, starting by convention with the topmost on the left side, and we will number these $2n$ external edges counterclockwise. A *Fully Packed Loop (FPL)* configuration F of size n is defined as a subgraph of G_n such that each vertex of G_n is incident to two edges of F. An example of configuration is given on Figure 1 (left). We let A_n be the total number of FPL configurations on the grid G_n. It is well known that FPL configurations are in bijection with alternating sign matrices (cf. [Pro01] for instance), and thus we have the famous enumeration proved independently by Zeilberger [Zei96] and Kuperberg [Kup96]:

$$A_n = \prod_{i=0}^{n-1} \frac{(3i+1)!}{(n+i)!}$$

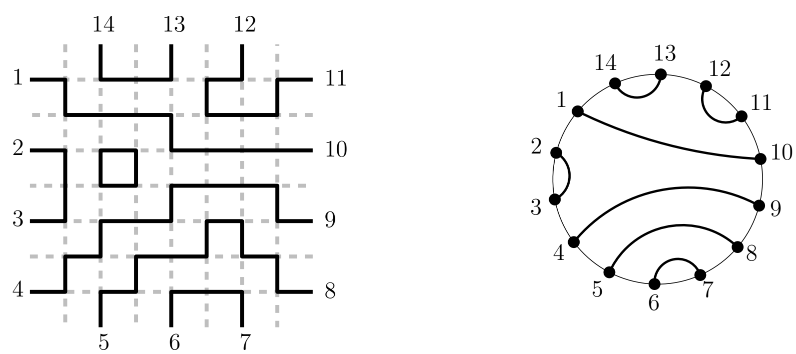


Fig. 1: A FPL configuration of size 7.

Define a *link pattern* π of size n as a matching on $\{1, \ldots, 2n\}$ of n *pairwise noncrossing* pairs $\{i, j\}$ between these $2n$ points, which means that there are no integers $i < j < k < \ell$ such that $\{i, k\}$ and $\{j, \ell\}$ are both in π. A FPL configuration on G_n naturally defines nonintersecting paths between its external edges, so we can define the link pattern $\pi(F)$ as the set of pairs $\{i, j\}$ where i, j label external edges which are the extremities of the same path in F. For instance, if F is the configuration of Figure 1, then $\pi(F)$ is the link pattern shown on its right, represented as a chord diagram.

Definition 1.1 ($\mathcal{A}_\pi$ **and** A_π) *Let π be a link pattern. The set $\mathcal{A}_\pi$ is defined as the set of all FPL configurations F of size n such that $\pi(F) = \pi$. We also let $A_\pi := |\mathcal{A}_\pi|$.*

Wieland's theorem: Given a link pattern π, consider the *rotated* link pattern $r(\pi)$ defined by $\{i, j\} \in r(\pi)$ if and only if $\{i-1, j-1\} \in \pi$, where indices are taken modulo $2n$. A beautiful result of Wieland [Wie00] states that $A_\pi = A_{r(\pi)}$, by giving a bijection between $\mathcal{A}_\pi$ and $\mathcal{A}_{r(\pi)}$.

Nested arches and $\mathbf{A}_\pi(m)$: Given a link pattern π on $\{1, \ldots, 2n\}$, and an integer $m \geq 0$, let us define $\pi \cup m$ as the link pattern on $\{1, \ldots, 2(n+m)\}$ given by the "nested" pairs $\{i, 2n+2m+1-i\}$ for $i = 1 \ldots m$, and the pairs $\{i+m, j+m\}$ for each $\{i, j\} \in \pi$. We will want to study the numbers $A_{\pi \cup m}$ as functions of m, so we introduce the notation $A_\pi(m) := A_{\pi \cup m}$.

1.2 Words, Ferrers diagrams, link patterns

We consider finite words on the alphabet with two letters 0 and 1, simply named *words*. For u a word, we let $|u|_0$ denote its number of zeros, $|u|_1$ its number of ones, and $|u| = |u|_0 + |u|_1$ its total length.

Proposition 1.2 *Given nonnegative integers k, ℓ, there is a bijection between words σ such that $|\sigma|_0 = k$ and $|\sigma|_1 = \ell$, and Ferrers diagrams fitting in the rectangle with k rows and ℓ columns.*

Proof: This is very standard. Given such a word $\sigma = \sigma_1 \cdots \sigma_{k+\ell}$, construct a path on the square lattice by drawing a North step when $\sigma_i = 0$ and an East step when $\sigma_i = 1$, for i from 1 to $k+\ell$. Then complete the picture by drawing a line up from the starting point, and a line left of the ending point; the resulting region enclosed in the wanted Ferrers diagram; see Figure 2 for an example. □

$\sigma = 0101011110$

$|\sigma| = 10, |\sigma|_0 = 4, |\sigma|_1 = 6$

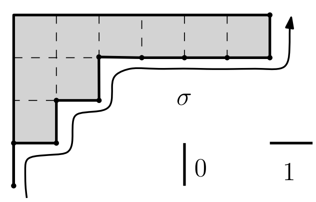


Fig. 2: Bijection between words and Ferrers diagrams.

Since we do not want to introduce too much notation, we use the bijection of Proposition 1.2 to *identify words and their corresponding Ferrers diagrams* in the rest of the article. The *conjugate* σ^* of $\sigma = \sigma_1 \cdots \sigma_n$ is the word of length n defined by $\sigma_i^* := 1 - \sigma_{n+1-i}$. Clearly we have $(\sigma^*)^* = \sigma$. The *degree* of σ is the number of indices $i < j$ such that $(\sigma_i, \sigma_j) = (1, 0)$, and is noted $d(\sigma)$; it is the number of boxes in the Ferrers diagram representation. For instance we have $d(\sigma) = 9$ for the example of Figure 2.

Suppose that σ, τ are words that verify $|\sigma|_0 = |\tau|_0$ and $|\sigma|_1 = |\tau|_1$, so that they form Ferrers diagrams included in a common rectangle by Proposition 1.2. We define $\sigma \leq \tau$ if σ is included in τ in the diagram representation: this is equivalent to $\sigma_{\leq i} \leq \tau_{\leq i}$ for all indices i, where $\sigma_{\leq i} = \sum_{j \leq i} \sigma_j$. If $\sigma \leq \tau$, we define the *skew shape* τ/σ as the set of boxes that are in τ but not in σ; if there are no two boxes in the same column, then τ/σ is a *horizontal strip*, and we write $\sigma \to \tau$. We define a *semistandard Young tableau* of shape σ, and length $N \geq 0$ to be a sequence $(\sigma_i)_{i=0 \ldots N}$ of words such that $\sigma_0 = \mathbf{0} \to \sigma_1 \ldots \to \sigma_N = \sigma$, where $\mathbf{0}$ is the empty partition. This is equivalent to the standard definition, i.e. a filling of the boxes of the diagram σ by positive integers not bigger than N, nondecreasing across each row from left to right and increasing down each column.

Suppose that u is a box in the diagram σ, which is in the kth row from the top and ℓth column from the left. The *content* $c(u)$ of u is defined as $\ell - k$, while its *hook-length* $h(u)$ is defined as the number of boxes in σ which are below u and in the same column, or right of u and in the same row (u itself being counted just once); define also $H_\sigma = \prod h(u)$ where the product is over all cells u of the diagram σ. We have then the *hook content formula*, which states that the number of semistandard Young tableaux of shape σ and length $N \geq 0$ is given by the following polynomial in N with leading term $\frac{1}{H_\sigma} N^{d(\sigma)}$:

$$\mathrm{SSYT}(\sigma, N) := \frac{1}{H_\sigma} \prod_{u \in \sigma} (N + c(u)) \tag{1}$$

Link patterns and the set $\mathcal{D}_n$: A link pattern π on $\{1, \ldots, 2n\}$ can also be considered as a word of length $2n$, where for each pair $\{i, j\}$ in π we set $\pi_i = 0$ and $\pi_j = 1$. Such words π form the following subset of $\{0, 1\}^{2n}$:

Definition 1.3 **($\mathcal{D}_n$)** *We denote by $\mathcal{D}_n$ the set of words σ of length $2n$, such that $|\sigma|_0 = |\sigma|_1 = n$, and each prefix u of σ verifies $|u|_0 \geq |u|_1$.*

These are known as Dyck words, and counted by the Catalan number $|\mathcal{D}_n| = C_n := \frac{1}{n+1}\binom{2n}{n}$. Note that $(\mathcal{D}_n, \leq)$ is a poset, with smallest element $\mathbf{0}_n := 0^n 1^n$ and greatest element $\mathbf{1}_n := (01)^n$. We will identify link patterns with words in $\mathcal{D}_n$.

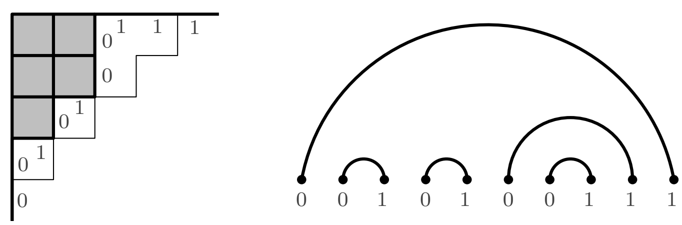


Fig. 3: The word $0010100111 \in \mathcal{D}_5$ as a diagram and a link pattern.

2 FPL in a triangle and linear recurrence relations

In all this section n will be a fixed positive integer.

2.1 FPL configurations in a triangle

We will here recall briefly the triangle arising in [CKLN06, Tha07], and refer to these works for more detail; we also advise the reader to look at Figure 4 while reading the definitions. We define the triangle $\mathcal{T}^n$ as the subset of $\mathbb{Z}^2$ consisting of the points of coordinates (x, y) which verify $x \geq y \geq 0$ and $x + y \leq 4n - 2$, with $2n$ external edges below all vertices $(2i, 0)$ for $i = 0 \ldots 2n - 1$, and horizontal edges between (i, i) and $(i+1, i)$, and between $(4n-2-i-1, i)$ and $(4n-2-i, i)$ for $i = 0, \ldots, 2n-2$, see left of Figure 4, where the edges in bold are the forced edges just described.

We consider the triangle with some extra conditions given by σ, τ words in $\mathcal{D}_n$: if $\sigma = \sigma_1 \ldots \sigma_{2n}$, we add a vertical edge below $(i-1, i-1)$ for each i such that $\sigma_i = 0$, while if $\tau = \tau_1 \ldots \tau_{2n}$, we add a vertical edge below $(2n-2+i, 2n-i)$ for each i such that $\tau_i = 1$. Note that σ and τ have to be interpreted differently than in [Tha07].

Definition 2.1 *A* FPL configuration f in a triangle (TFPL) *with boundary conditions σ, π, τ in $\mathcal{D}_n$ is a graph on $\mathcal{T}^n$, where vertical edges on the left and right boundary are given by σ and τ as above. All vertices (except on the left and right boundaries) are imposed to be of degree 2, and we have furthermore (1) the $2n$ bottom external edges must be linked by paths in $\mathcal{T}^n$ according to the link pattern π, and (2) the paths starting on the left boundary must end on the right boundary; cf Figure 4 for an example. The set of these TFPLs is denoted $\mathcal{T}^\pi_{\sigma,\tau}$, and we define $t^\pi_{\sigma,\tau}$ as the cardinality $|\mathcal{T}^\pi_{\sigma,\tau}|$.*

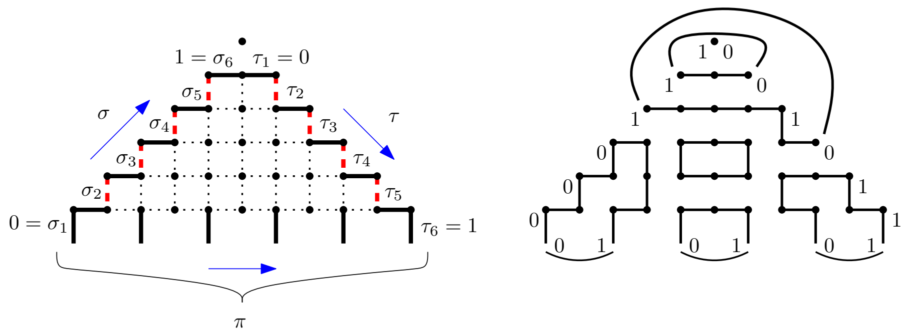


Fig. 4: Boundary conditions for FPL in a triangle.

2.2 Linear recurrences for refined FPL numbers

The link between FPLs and TFPLs is given by the following formula from [CKLN06]: for $m \geq 0$,

$$A_\pi(m) = \sum_{\sigma,\tau \in \mathcal{D}_n} \mathrm{SSYT}(\sigma, n) \cdot t^\pi_{\sigma,\tau} \cdot \mathrm{SSYT}(\tau^*, m-2n+1), \tag{2}$$

Following Thapper [Tha07], we now consider endomorphisms of $\mathbb{C}\mathcal{D}_n$, the vector space of formal complex linear combinations of elements of $\mathcal{D}_n$. We will write such endomorphisms g as matrices in the canonical basis $\mathcal{D}_n$, so that, if $\sigma, \tau \in \mathcal{D}_n$, we denote by $g_{\sigma\tau}$ the coefficient of σ in the expansion of $g(\tau)$. Then we define $\mathbf{b}$ by $\mathbf{b}_{\sigma\tau} = 1$ if $\sigma \to \tau$, and $\mathbf{b}_{\sigma\tau} = 0$ otherwise. We define $\widetilde{\mathbf{b}}$ by $\widetilde{\mathbf{b}}_{\sigma\tau} = 1$ if $\tau^* \to \sigma^*$ and $\widetilde{\mathbf{b}}_{\sigma\tau} = 1$ otherwise. Given $\pi \in \mathcal{D}_n$, we also let $(\mathbf{t}^\pi)_{\sigma\tau} = t^\pi_{\sigma,\tau}$. By definition of semistandard Young Tableaux, we have $\mathrm{SSYT}(\sigma, n) = (\mathbf{b}^n)_{\mathbf{0}_n\sigma}$ and $\mathrm{SSYT}(\tau^*, m-2n+1) = (\widetilde{\mathbf{b}}^{m-2n+1})_{\tau\mathbf{0}_n}$. So we can rewrite Equation (2) as

$$A_\pi(m) = \left(\mathbf{b}^n \mathbf{t}^\pi \widetilde{\mathbf{b}}^{m-2n+1}\right)_{\mathbf{0}_n\mathbf{0}_n} \tag{3}$$

We have then the following Proposition conjectured by Thapper [Tha07, Conjecture 3.4]:

Theorem 2.2

$$\mathbf{b}\mathbf{t}^\pi = \mathbf{t}^\pi \widetilde{\mathbf{b}} \quad \textit{for all} \quad \pi \in \mathcal{D}_n. \tag{4}$$

Proof (Sketch): As shown by Thapper, the coefficients on the left and right side enumerate some configurations in "extended" triangles. By studying Wieland's rotation (cf. Section 1.1), it is possible to show that this can be applied in these extended triangles, and that it indeed exchanges bijectively left and right extended triangles. Note that one has to apply either H_0 or H_1 in Wieland's original notation in [Wie00], and not the composition $H_0 \circ H_1$: this *shifts* the link pattern π, and one has to check that the boundary conditions σ and τ are indeed preserved. □

Now we can apply the commutation relation (4) repeatedly in Equation (3), and obtain $A_\pi(m) = \left(\mathbf{b}^{m-n+1}\mathbf{t}^\pi\right)_{\mathbf{0}_n\mathbf{0}_n}$ which can be expanded as $\sum_{\sigma\in\mathcal{D}_n}\mathrm{SSYT}(\sigma, m-n+1)\cdot t^\pi_{\sigma,\mathbf{0}_n}$; this involves only TFPLs with $\tau = \mathbf{0}_n$, so if we introduce $\mathbf{t}$ as $(\mathbf{t})_{\sigma\pi} = t^\pi_{\sigma,\mathbf{0}_n}$ we get :

Proposition 2.3 *For all integers $m \geq 0$, we have $A_\pi(m) = \left(\mathbf{b}^{m-n+1}\mathbf{t}\right)_{\mathbf{0}_n\pi}$.*

We can now use the beautiful idea of Thapper: by Theorem 3.1, the coefficients $t^\pi_{\sigma,\mathbf{0}_n}$ of $\mathbf{t}$ are integers, equal to 0 unless $\sigma \leq \pi$, and such that $t^\pi_{\pi,\mathbf{0}_n} = 1$. This means that, if we give the basis $\mathcal{D}_n$ a linear order extending $\leq$, then the matrix of $\mathbf{t}$ is upper triangular with ones on its diagonal. It is thus invertible, with its inverse $\mathbf{t}^{-1}$ being also triangular with ones on its diagonal, and with integer entries. We can thus define:

Definition 2.4 *We define the matrix $\mathbf{c}$ by $\mathbf{c} := \mathbf{t}^{-1}\mathbf{b}\mathbf{t}$.*

We can now state the main result of this section, conjectured by Thapper [Tha07, Proposition 3.5]:

Theorem 2.5 *For any $\pi \in \mathcal{D}_n$, we have the polynomial identity:*

$$A_\pi(m) = \sum_{\alpha\in\mathcal{D}_n} \mathbf{c}_{\alpha\pi} A_\alpha(m-1)$$

Proof: By Proposition 2.3 and the definition of $\mathbf{c}$, we get for any m

$$A_\pi(m) = \left(\mathbf{b}^{m-n+1}\mathbf{t}\right)_{\mathbf{0}_n\pi} = \left(\mathbf{b}^{m-n}\mathbf{t}\mathbf{c}\right)_{\mathbf{0}_n\pi} = \sum_{\alpha\in\mathcal{D}_n}\left(\mathbf{b}^{m-n}\mathbf{t}\right)_{\mathbf{0}_n\alpha}\mathbf{c}_{\alpha\pi},$$

from which the result follows, again by Proposition 2.3. □

We remark that the coefficients $c_{\alpha\pi}$ are not the unique integers verifying Theorem 2.5. But first, we have a uniform definition for them. Second, there is evidence that they are "good" coefficients, based on data communicated to the author by J. Thapper: these numbers are quite small (they are between -1 and 2 for $n = 5$, while the supremum of $\mathbf{t}$ exceeds 80000), and we conjecture that they verify $c_{\alpha\pi} = c_{\alpha^*\pi^*}$, that $c_{\alpha\pi}$ only depends on the skew shape π/α, and many other properties. It seems that there is hope that these coefficients have a direct combinatorial characterization.

3 Some properties of TFPL configurations

In this Section we will prove certain enumerative questions related to TFPL configurations. In particular we give a new proof of the following theorem, which was essential in Section 2.2:

Theorem 3.1 *Let σ, π, τ be in $\mathcal{D}_n$. Then $t^\pi_{\sigma,\tau} = 0$ unless $\sigma \leq \pi$. Moreover, if $\sigma = \pi$, then $t^\pi_{\pi,\mathbf{0}_n} = 1$ and $t^\pi_{\pi,\tau} = 0$ for $\tau \neq \mathbf{0}_n$.*

It was proved first in [CKLN06, Section 7] in a very technical way, while here our proof (see Section 3.2) is much shorter and illuminating.

3.1 Oriented TFPL configurations

The vertices of $\mathcal{T}_n$ can be partitioned in lines: for $i \in \{1, \ldots, 2n\}$, we define E_i as the vertices of $\mathcal{T}^n$ such that $x+y = 2i-2$, and for $i \in \{1, \ldots, 2n-1\}$, we define O_i as the vertices of $\mathcal{T}^n$ such that $x+y = 2i-1$. The case $n = 3$ is given on Figure 5. Now let us suppose we have boundary configurations σ, τ, π on the triangle $\mathcal{T}_n$. We first define an orientation for all edges around the triangle as follows. On the left boundary, we orient edges to the right and upwards; on the right boundary, we orient them to the right and downwards; for the $2n$ vertical external edges on the bottom, we orient the one attached to $(2i-2, 0)$ upwards if $\pi_i = 0$, and downwards if $\pi_i = 1$, for $i \in \{1, \ldots, 2n\}$. Now given a TFPL configuration f in $\mathcal{T}^{\pi}_{\sigma,\tau}$, we now orient all remaining edges so that *each vertex of degree* 2 *have one incoming edge and one outgoing edge*. This condition determines clearly the orientation of edges in a path of f joining external edges, and by convention we orient the closed paths of f clockwise. In this way we associate to each configuration $f \in \mathcal{T}^{\pi}_{\sigma,\tau}$ an *oriented configuration* that we will denote by $\mathrm{or}(f)$.

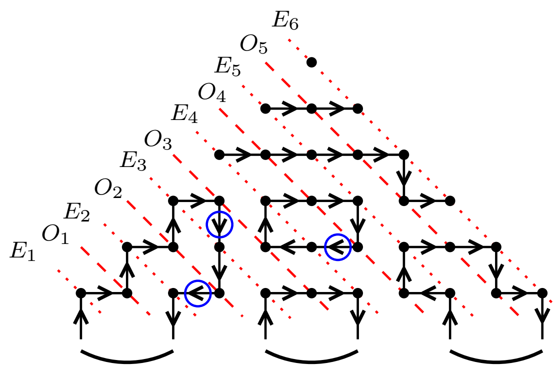


Fig. 5: Lines E_i and O_i.

3.2 Proof of Theorem 3.1

Definition 3.2 ($\mathcal{N}_i(f)$ **and** $N_i(f)$) *Let σ, τ, π be in $\mathcal{D}_n$, f be a configuration in $\mathcal{T}^{\pi}_{\sigma,\tau}$, and i be an integer in $\{1, \ldots, 2n-1\}$. We define $\mathcal{N}_i(f)$ as the set of oriented edges in $\mathrm{or}(f)$ which are directed from a vertex in O_i to a vertex in E_i. We also define $N_i(f) = |\mathcal{N}_i(f)|$, and $N_0(f) = 0$ by convention.*

These oriented edges are circled in the example of Figure 5, and we get $N_i(f) = 0, 1, 1, 1, 0$ for $i = 1, 2, 3, 4, 5$ respectively. We can now state the key lemma:

Lemma 3.3 *Let σ, τ, π be in $\mathcal{D}_n$, and f a configuration in $\mathcal{T}^{\pi}_{\sigma,\tau}$. Then*

$$N_i(f) - N_{i-1}(f) = \pi_i - \sigma_i, \quad \textit{for } i = 1, \ldots, 2n-1. \tag{5}$$

Proof: We consider the oriented configuration $\mathrm{or}(f)$. The i vertices of E_i have one incoming edge, except $(i-1, i-1)$ when $\sigma_i = 1$. If this incoming edge comes from O_i it is an element of $\mathcal{N}_i(f)$; let $X_i(f)$ be the other incoming edges, and $x_i(f) := |X_i(f)|$. We have then

$$N_i(f) + x_i(f) + \sigma_i = i. \tag{6}$$

Similarly, consider the $i-1$ vertices on the line O_{i-1}: each of them has exactly one outgoing edge, and if this edge goes to the line E_{i-1} it is by definition in $\mathcal{N}_{i-1}(f)$. We form the set $Y_i(f)$ with the other outgoing edges of O_{i-1}, and let $y_i(f) := |Y_i(f)|$. We obtain here

$$N_{i-1}(f) + y_i(f) = i - 1. \tag{7}$$

Now the sets $Y_i(f)$ and $X_i(f)$ coincide except in the case $\pi_i = 0$, where there is an external edge incoming in $(2i-2, 0) \in E_i$ (by definition of the orientation) and therefore belongs to X_i and not to Y_i. Thus $x_i(f) = y_i(f) + (1 - \pi_i)$ and by injecting this in Equations (6) and (7) we deduce Equation (5). □

We can now give the proof of the first half of Theorem 3.1. If we sum the relations (5) for i going from 1 to j, then for any $j \in \{1, \ldots, 2n\}$ we obtain $\pi_{\leq j} - \sigma_{\leq j} = N_j(f)$. Since this is nonnegative, this proves that $\sigma \leq \pi$ (cf. Section 1.2), and we are done. The second part of Theorem 3.1 is much easier, see the end of Section 7 in [CKLN06].

3.3 Common prefixes and suffixes

We just showed that TFPLs exist only when $\sigma \leq \pi$ (and $\tau \leq \pi$ by symmetry), and that in case of equality $\sigma = \pi$ there is just one configuration, when $\tau = \mathbf{0}_n$. It is natural to ask what happens when σ is smaller than π but "close" to it, and one possible answer is the following:

Theorem 3.4 *Let $\pi, \sigma, \tau \in \mathcal{D}_n$, and suppose that there exist words u, σ', π', v such that $\sigma = u\sigma' v$ and $\pi = u\pi' v$ (concatenation of words). Let $a = |u|_0 + |v|_0$ and $b = |u|_1 + |v|_1$. Then $t^{\pi}_{\sigma,\tau} = 0$ unless τ is of the form $\tau = 0^a \tau' 1^b$.*

The proof is quite technical and will be omitted here. It involves a slight variant of de Gier's lemma on *fixed edges* [dG05, Lemma 8], in which we make use of the oriented TFPL configurations of Section 3.1.

There is one special case emerging naturally in the proof, which is when $\pi' = 1^{n-b}0^{n-a}$; note that this means that π/σ is a *rotated diagram*, i.e. a skew shape which is the (translated of) a Ferrers diagram after a half turn. In this case, each vertex of $\mathcal{T}_n$ can be shown to be incident to at least one fixed edge, and another observation of de Gier can be used to show that the enumeration of $\mathcal{T}^{\pi}_{\sigma,\tau}$ is then reduced to a tiling problem, whose solution in our case can be written under the form of a single determinant of size $min(n-a, n-b)$. So if π/σ is a row or a column of cells, we get a single binomial coefficient.

3.4 Extremal TFPL configurations

We recall that $d(\sigma)$ is the number of boxes in the Ferrers diagram of σ.

Proposition 3.5 *One has $t^{\pi}_{\sigma,\tau} = 0$ unless $d(\sigma) + d(\tau) \leq d(\pi)$. Furthermore, for every $\pi \in \mathcal{D}_n$ we have*

$$\frac{1}{H_\pi} = \sum_{\substack{\sigma,\tau \in \mathcal{D}_n \\ d(\sigma)+d(\tau)=d(\pi)}} t^{\pi}_{\sigma,\tau} \cdot \frac{1}{2^{d(\sigma)} H_\sigma} \cdot \frac{1}{2^{d(\tau)} H_\tau}. \tag{8}$$

We reproduce the argument of [Tha07, Lemma 3.7] which is the first part of the proposition.

Proof: As Equation (2) shows, $A_\pi(m)$ is *polynomial* in m, and using Theorem 3.1 and (1), it is easy to deduce as in [CKLN06] that it is a polynom with leading term $\frac{1}{H_\pi} m^{d(\pi)}$. Now using relation (4) and assuming m is an even integer, we can get from (3) that $A_\pi(m) = \left(\mathbf{b}^{m/2}\mathbf{t}^\pi\widetilde{\mathbf{b}}^{m/2-n+1}\right)_{\mathbf{0}_n\mathbf{0}_n}$, i.e.

$$\sum_{\sigma,\tau\in\mathcal{D}_n} \mathrm{SSYT}(\sigma, m/2)\cdot t^\pi_{\sigma,\tau}\cdot \mathrm{SSYT}(\tau^*, m/2-n+1).$$

This is also polynomial in m, and thus the coefficients of degree $> d(\pi)$ must vanish, which implies the first part of the proposition. The second part follows by taking the coefficient of degree $d(\pi)$ in this last expression, which is necessarily equal to $\frac{1}{H_\pi}m^{d(\pi)}$. □

We will call *extremal* the TFPL configurations verifying $d(\sigma)+d(\tau)=d(\pi)$.

4 TFPL and Littlewood Richardson coefficients

In this section we will show that the coefficients $t^\pi_{\sigma,\tau}$ when $d(\sigma)+d(\tau)=d(\pi)$ are given by the Littlewood Richardson coefficients.

4.1 Littlewood Richardson coefficients and puzzles

We refer to [Sta99] for background on symmetric functions. Let $x=(x_1,x_2,\ldots)$ be commuting indeterminates, and let $\Lambda(x)$ be the ring of symmetric functions in x. Schur functions $s_\lambda(x)$ (λ a Ferrers diagram) form a basis of $\Lambda(x)$, and the *Littlewood–Richardson (LR) coefficients* $c^\nu_{\lambda,\mu}$ are defined as the coefficients in the expansion of their products $s_\mu(x)s_\nu(x)=\sum_\lambda c^\lambda_{\mu,\nu}s_\lambda(x)$. The LR coefficient $c^\nu_{\lambda,\mu}$ is 0 unless $\lambda\geq\mu,\nu$ and $d(\mu)+d(\nu)=d(\lambda)$. Schur functions can be defined combinatorially in terms of semistandard Young tableaux, and in this case it is clear that, under the specialization $x_i=1$ for $i=1\ldots N$ and $x_i=0$ otherwise, $s_\lambda(x)$ is equal to $\mathrm{SSYT}(\lambda,N)$.

If one introduces $s_\lambda(x,y)$ as the Schur function in variables $x_1,x_2,\ldots,y_1,y_2,\ldots$ then it is shown in [Sta99, p.341] that $s_\lambda(x,y)=\sum_{\mu,\nu}c^\lambda_{\mu,\nu}s_\mu(x)s_\nu(y)$. By specializing at $x_i=y_i=1$ for $i=1\ldots m$ and $x_i=y_i=0$ for $i>m$, we get a polynomial identity in m which in top degree can be written as:

$$\frac{1}{H_\lambda}=\sum_{\mu,\nu}c^\lambda_{\mu,\nu}\cdot\frac{1}{2^{d(\mu)}H_\mu}\cdot\frac{1}{2^{d(\nu)}H_\nu}. \tag{9}$$

The LR coefficients are easily seen to be nonnegative integers by character theory [Sta99, p.355]; many combinatorial descriptions of them are also known, the most famous being the original Littlewood Richardson rule [LR34]. We will here use the (slightly adatpted) Knutson Tao puzzles [KTW04, KT03]:

Definition 4.1 (Knutson Tao puzzle) *Let n be an integer, and σ,π,τ words in $\mathcal{D}_n$. Consider a triangle with edge size $2n$ on the regular triangular lattice, where unit edges on left, bottom and right side are labelled by σ,π,τ respectively. A* Knutson-Tao (KT) puzzle *with boundary σ,π,τ is a labeling of each internal edge of the triangle with* 0,1 *or* 2*, such that the labeling induced on each of the $(2n)^2$ unit triangles is composed either of three* 0*, or of three* 1*, or has* 0, 1, 2 *in counterclockwise order.*

The exhaustive list of all authorized labelings of triangles is given on the left of Figure 6, and on the right we have an example of a puzzle with boundaries $\sigma=00011011$, $\pi=00110101$, $\tau=00011011$. It turns out that KT puzzles give a combinatorial interpretation for LR coefficients.

Theorem 4.2 ([KTW04, KT03]) *KT-puzzles with boundary σ,π,τ are counted by $c^\pi_{\sigma,\tau}$.*

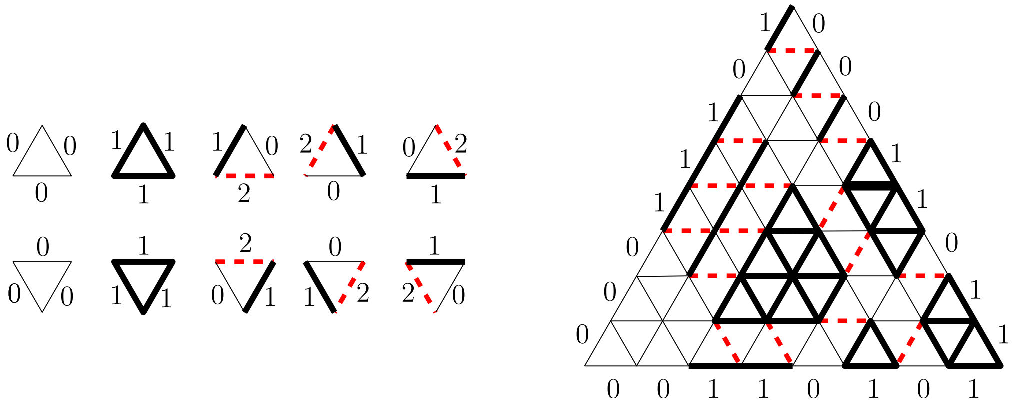


Fig. 6: Authorized triangles in a KT puzzle, and an example.

4.2 The enumeration of extremal TFPL configurations

We can finally state our final result:

Theorem 4.3 *Given σ, π, τ such that $d(\sigma) + d(\tau) = d(\pi)$, we have $t^{\pi}_{\sigma,\tau} = c^{\pi}_{\sigma,\tau}$.*

The proof consists in a bijective correspondence Φ from KT-puzzles with boundary σ, π, τ to TFPLs in $\mathcal{T}^{\pi}_{\sigma,\tau}$. The definition is local: each piece of a puzzle is transformed into a small part of a TFPL configuration. In fact, we will define directly a bijection to oriented configurations (defined in 3.1). The rules are described on Figure 7: non horizontal edges of unit triangles give rise to vertices in $\mathcal{T}_n$, while the horizontal ones are sent on lines $y = i + 1/2$. After every triangle of a puzzle P has been tranformed (see Figure 8, left), delete the original puzzle, and rescale the graph obtained so that vertices lie on a square grid. To finish, remove the superfluous horizontal edges that appear along the left boundary, double the length of the bottom vertical edges:the resulting graph on $\mathcal{T}_n$ is by definition $\Phi(P)$: see Figure 8 again.

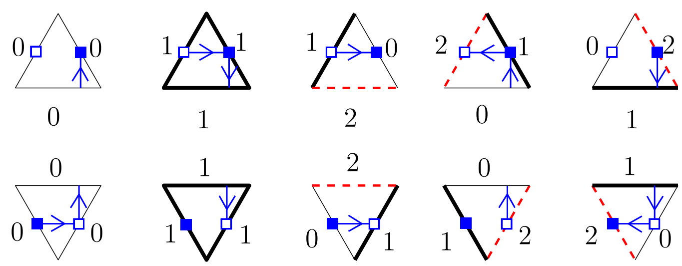


Fig. 7: The local transformations of the bijection Φ.

Lemma 4.4 *For any puzzle P with boundary σ, π, τ, $\Phi(P)$ is an (oriented) TFPL configuration in $\mathcal{T}^{\pi}_{\sigma,\tau}$.*

Proof: It is easily seen (albeit a bit tedious) to check by inspection of Figure 7 that the edges created on the left and right boundaries of $\Phi(P)$ correspond indeed to σ and τ, and that the bottom external edges are also present, all of them with their correct orientation. It is also the case, once again by inspection, that the graph $\Phi(P)$ is such that each of its vertices has one incoming edge and one outgoing.

Paths starting from the left side end up on the right side: indeed, the only other possibility is that such a path p ends on the bottom side (the left side is not possible because of conflicting orientations); but this case is easily dismissed, because in the region of $\mathcal{T}_n$ above p, there would remain less incoming edges (on the left boundary) than outgoing edges (on the right boundary), which is absurd.

Finally, one needs to check that the paths connecting the bottom external edges follow the link pattern π, and this is more subtle. We already checked that the orientation of these external edges is correct; we must also show that the paths go globally "from left to right", that is they should not connect two external edges such that the left one is directed downwards and the right one upwards. Now such a bad path would necessarily possess a subpath consisting of an up step followed by one or more steps to the left, followed by one downstep; but a quick look at the rules of Figure 7 reveals that a step to the left is either preceded by a down step, or followed by an up step, and thus bad paths cannot appear in $\Phi(P)$. A similar reasoning to the one for paths between the left and right boundaries then shows that paths between bottom external edges follow the link pattern π; this finally proves that $\Phi(P)$ is in $\mathcal{T}^{\pi}_{\sigma,\tau}$. $\square$

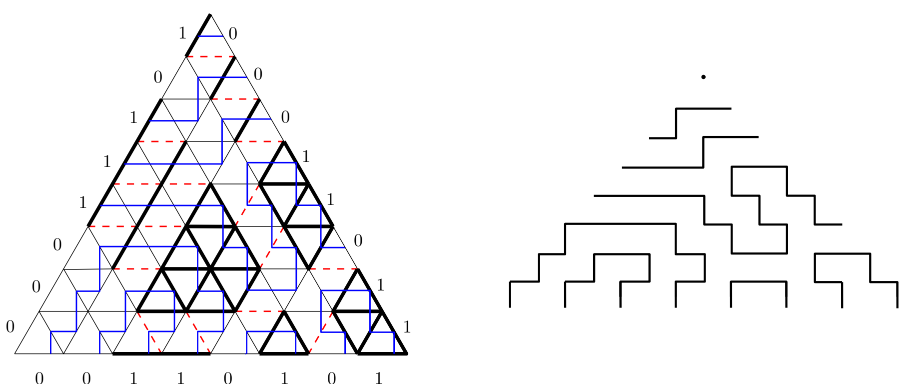


Fig. 8: Example of the bijection Φ.

Proof of Theorem 4.3: The previous lemma showed that Φ is well defined. It is also clear that Φ is injective, because the ten configurations of oriented edges on Figure 7 are all different, and thus from a puzzle $\Phi(P)$ one can reconstruct the labeling of all edges, i.e. the puzzle P. Note the importance of orienting configurations here, because without them some of the local configurations become identified. The injectivity implies by Theorem 4.2 that $t^{\pi}_{\sigma,\tau} \leq c^{\pi}_{\sigma,\tau}$. Now comparing Equations (8) and (9) tells us that for a fixed π, $\sum_{\sigma,\tau} c^{\pi}_{\sigma,\tau} X_{\sigma\tau} = \sum_{\sigma,\tau} t^{\pi}_{\sigma,\tau} X_{\sigma\tau}$ for certain *positive coefficients* $X_{\sigma\tau}$, the sum being over σ, τ such that $d(\sigma) + d(\tau) = d(\pi)$. Together with the injectivity of Φ, this proves that $t^{\pi}_{\sigma,\tau} = c^{\pi}_{\sigma,\tau}$ and Φ is in fact bijective, completing the proof of the theorem. $\square$

Acknowledgements

The author would like to thank Johan Thapper for sharing his data on TFPLs.

References

[CKLN06] F. Caselli, C. Krattenthaler, B. Lass, and P. Nadeau. On the number of fully packed loop configurations with a fixed associated matching. *Electron. J. Combin.*, 11(2):Research Paper 16, 43 pp. (electronic), 2004/06.

[CS10] Luigi Cantini and Andrea Sportiello, *Proof of the Razumov–Stroganov conjecture*, `arXiv:1003.3376v1`.

[dG05] Jan de Gier. Loops, matchings and alternating-sign matrices. *Discrete Math.*, 298(1-3):365–388, 2005.

[KT03] Allen Knutson and Terence Tao. Puzzles and (equivariant) cohomology of Grassmannians. *Duke Math. J.*, 119(2):221–260, 2003.

[KTW04] Allen Knutson, Terence Tao, and Christopher Woodward. The honeycomb model of $\mathrm{GL}_n(\mathbb{C})$ tensor products. II. Puzzles determine facets of the Littlewood-Richardson cone. *J. Amer. Math. Soc.*, 17(1):19–48 (electronic), 2004.

[Kup96] Greg Kuperberg. Another proof of the alternating-sign matrix conjecture. *Internat. Math. Res. Notices*, (3):139–150, 1996.

[LR34] Dudley E. Littlewood and Archibald Read Richardson. Group characters and algebra. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character (The Royal Society)*, 233:99–141, 1934.

[Pro01] James Propp. The many faces of alternating-sign matrices. In *Discrete models: combinatorics, computation, and geometry (Paris, 2001)*, Discrete Math. Theor. Comput. Sci. Proc., AA, pages 043–058 (electronic). Maison Inform. Math. Discrèt. (MIMD), Paris, 2001.

[RS04] A. V. Razumov and Yu. G. Stroganov. Combinatorial nature of the ground-state vector of the $\mathrm{O}(1)$ loop model. *Teoret. Mat. Fiz.*, 138(3):395–400, 2004.

[Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[Tha07] Johan Thapper. Refined counting of fully packed loop configurations. *Sém. Lothar. Combin.*, 56:Art. B56e, 27 pp. (electronic), 2006/07.

[Wie00] Benjamin Wieland. A large dihedral symmetry of the set of alternating sign matrices. *Electron. J. Combin.*, 7:Research Paper 37, 13 pp. (electronic), 2000.

[Zei96] Doron Zeilberger. Proof of the alternating sign matrix conjecture. *Electron. J. Combin.*, 3(2):Research Paper 13, approx. 84 pp. (electronic), 1996. The Foata Festschrift.

[ZJ] Paul Zinn-Justin. Six-Vertex, Loop and Tiling models: Integrability and Combinatorics. Habilitation thesis, arXiv:0901.0665.

[ZJ06] Paul Zinn-Justin. Proof of the Razumov–Stroganov conjecture for some infinite families of link patterns. *Electron. J. Combin.*, 13(1):Research Paper 110, 15 pp, 2006.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 313–324

The Homology of the Real Complement of a k-parabolic Subspace Arrangement

Christopher Severs[1] and Jacob A. White[2]

[1]*Mathematical Sciences Research Institute, Berkeley, CA*
[2]*School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ*

Abstract. The k-parabolic subspace arrangement, introduced by Barcelo, Severs and White, is a generalization of the well known k-equal arrangements of type-A and type-B. In this paper we use the discrete Morse theory of Forman to study the homology of the complements of k-parabolic subspace arrangements. In doing so, we recover some known results of Björner et al. and provide a combinatorial interpretation of the Betti numbers for any k-parabolic subspace arrangement. The paper provides results for any k-parabolic subspace arrangement, however we also include an extended example of our methods applied to the k-equal arrangements of type-A and type-B. In these cases, we obtain new formulas for the Betti numbers.

Résumé. L'arrangement k-parabolique, introduit par Barcelo, Severs et White, est une généralisation des arrangements, k-éguax de type A et de type B. Dans cet article, nous utilisons la théorie de Morse discrète proposée par Forman pour étudier l'homologie des compléments d'arrangements k-paraboliques. Ce faisant, nous retrouvons les résultats connus de Bjorner et al. mais aussi nous fournissons une interprétation combinatoire des nombres de Betti pour des arrangements k-paraboliques. Ce papier fournit alors des résultats pour n'importe quel arrangement k-parabolique, cependant nous y présentons un exemple étendu de nos méthodes appliquées aux arrangements k-éguax de type A et de type B. Pour ce cas, on obtient de nouvelles formules pour les nombres de Betti.

Keywords: Subspace Arrangements, Discrete Morse Theory, Coxeter Groups

1 Introduction

Recall that the real (essentialized) k-equal subspace arrangement, $\mathcal{A}_{n,k}$ is defined to be the collection of subspaces given by

$$x_{i_1} = x_{i_2} = \cdots x_{i_k} \text{ and subject to the relation } \sum_{i=1}^{n+1} x_i = 0.$$

This arrangement was originally studied by Björner and Lovász in connection with linear decision trees and the k-equal problem [3]. In particular, they showed how the Betti numbers could be used to give lower bounds on the number of leaves in a decision tree which solves the k-equal problem. The cohomology of the complement of this subspace arrangement was studied by Björner and Welker in [6]. The Betti numbers have applications in the study of linear decision trees and the k-equal problem. Recall that the complement of a subspace arrangement $\mathcal{A}$ is $M_{\mathcal{A}} = \mathbb{R}^n - \cup_{X \in \mathcal{A}} X$. In order to compute the cohomology group of the complement they make use of the well known Goresky-MacPherson formula for subspace arrangements [8].

Proposition 1.1 (Goresky-MacPherson Formula) *Let $\mathcal{A}$ be subspace arrangement with complement $M_{\mathcal{A}}$, and let $\mathcal{L}_{\mathcal{A}}$ be the intersection lattice of $\mathcal{A}$. Then*

$$\widetilde{H}^i(M_{\mathcal{A}}) \cong \bigoplus_{x \in \mathcal{L}^{\geq \hat{0}}} \widetilde{H}_{codim(x)-i-2}(\Delta(\hat{0}, x)), \tag{1}$$

where $\Delta(\hat{0}, x)$ is the order complex of the interval between $\hat{0}$ and x.

By considering the homology of the intersection lattice of $\mathcal{A}_{n,k}$, Björner and Welker were able to give a description of the cohomology of the complement, $M_{\mathcal{A}_{n,k}}$. Later, Björner and Wachs [5] used lexicographic shellability to study the intersection lattice $\mathcal{L}_{\mathcal{A}_{n,k}}$.

This method of studying subspace arrangements via the Goresky-MacPherson formula and shellability of the intersection lattice was continued in the work of Björner and Sagan [4]. In this work they define type-B and type-D analogues of the k-equal subspace arrangement. They then show that the intersection lattice in the case of type-B is shellable and prove results about the cohomology of the complement of the type-B k-equal arrangement using the Goresky-MacPherson formula.

The cohomology of complement of the type-D k-equal arrangement was further studied by Feichtner and Kozlov. In [7], they prove that the intersection lattice of the type-D k-equal subspace arrangement is shellable for large values of k and again use the Goresky-MacPherson formula to calculate the cohomology groups in those cases. Their approach uses a generalization of lexicographic shellability due to Kozlov [10]. For smaller values of k, Feichtner and Kozlov use more sophisticated techniques from algebraic topology to approach the problem.

In this paper, we revisit the study of the complement of the k-equal arrangements of type-A,B and D. via a generalization known as k-parabolic subspace arrangements. Introduced by Barcelo, Severs and White in [2], the k-parabolic arrangement is a real subspace arrangement that may be defined for any finite real reflection group, W. The subspaces in the arrangement are those which are fixed under the action of irreducible parabolic subgroups of W that have rank $k-1$. More information on these arrangements is provided in Section 2.

Our study of the homology of the complement of the k-parabolic subspace arrangement is not done via the Goresky-MacPherson formula and the intersection lattice, but rather by using the discrete Morse theory of Forman. First we construct a polyhedral complex $\Delta_k(W)$ that is homotopy equivalent to the complement of the k-parabolic arrangement. Recall that the poset of all cosets of parabolic subgroups under inclusion is isomorphic to the face lattice of the Coxeter arrangement corresponding to W. We obtain the face poset for $\Delta_k(W)$ by removing the upper order ideals corresponding to the irreducible parabolic subgroups of rank $k-1$. We then create an acyclic matching on this poset and use discrete Morse theory to calculate the homology groups and Betti numbers for $\Delta_k(W)$, and hence the complement.

The remainder of the paper is divided into three sections. In the first section we recall some definitions and theorems that we will need concerning k-parabolic subspace arrangements and discrete Morse theory. In the next section we provide an extended example of our methods using the type-A and type-B k-equal subspace arrangements. These arrangements should be familiar to most readers and give a more intuitive look at our matching and study of the homology of the complement. Moreover, we obtain new formulas for the Betti numbers of these arrangements, along with a new combinatorial interpretation for them. In the final section we present our matching in fullest generality as well, including some periodicity results for the homology, and a combinatorial interpretation of the Betti numbers. We also mention some open questions.

2 Background

The material on k-parabolic arrangements in this section is taken from [2]. The definitions and theorems concerning discrete Morse theory are primarily from Kozlov's *Combinatorial Algebraic Topology* [9], but may also be found in many other papers and books. This section contains no new material.

2.1 k-parabolic subspace arrangements

We start with the definition of the k-parabolic arrangement $\mathcal{W}_{n,k}$. Choose a finite real reflection group W of rank n with simple reflections S. A *parabolic* subgroup is a subgroup of the form $w\langle I\rangle w^{-1}$, for some $I \subseteq S, w \in W$. Consider $\mathcal{P}(W)$, the lattice of *all* parabolic subgroups of W. It was shown by Barcelo and Ihrig [1] that the intersection lattice of the Coxeter arrangement associated to W, $\mathcal{L}(\mathcal{H}(W))$, is isomorphic to $\mathcal{P}(W)$. The isomorphism is established by sending a parabolic subgroup to the set of points in $\mathbb{R}^n$ that it fixes, and sending an intersection of hyperplanes in the Coxeter arrangement to the parabolic subgroup of W that fixes the points in the intersection. We will use this correspondence to define a k-parabolic subspace arrangement associated to W.

Definition 2.1 (k-parabolic subspace arrangement) *Let W be a finite real reflection group of rank n and let $\mathcal{P}_{n,k}(W)$ be the lattice of all irreducible parabolic subgroups of W of rank $k-1$. The k-parabolic subspace arrangement, $\mathcal{W}_{n,k}$ is the collection of subspaces*

$$\{Fix(G)|G \in \mathcal{P}_{n,k}(W)\}$$

As with the k-equal arrangement, the k-parabolic arrangement can be embedded in its corresponding Coxeter arrangement. This allows us to think of the subspaces in $\mathcal{W}_{n,k}$ as intersections of hyperplanes of the Coxeter arrangement. Also, $\mathcal{L}(\mathcal{W}_{n,k})$ is a subposet of $\mathcal{L}(\mathcal{H}(W)) = \mathcal{L}(\mathcal{W}_{n,2})$.

If we let $W = A_n$ then it is easy to see that we recover the k-equal subspace arrangement from the introduction. Furthermore, if we let $W = B_n$, we recover the $\mathcal{B}_{n,k,k-1}$ subspace arrangements of Björner and Sagan. We note however when $W = D_n$, we do not recover the $\mathcal{D}_{n,k}$ arrangements of Björner and Sagan except in the special case $k = 3$.

The natural embedding of $\mathcal{W}_{n,k}$ into the Coxeter arrangement allows us to use a construction, due to Orlik [11], to create a cell complex homotopy equivalent to the complement. To be more precise, there is a order-reversing correspondence between the face poset of the Coxeter arrangement and the dual zonotope, known as the W-Permutahedron. Under this correspondence, removing a subcomplex of faces of $\mathcal{W}_{n,k}$ is homotopy equivalent to a polyhedral subcomplex of the Permutahedron. In the case of the complement $M_{\mathcal{W}_{n,k}}$, we apply this fact to the subcomplex generated by all faces contained in a subspace of $\mathcal{W}_{n,k}$. We shall call the resulting cell complex $\Delta_k(W)$.

Lemma 2.2 *There exists a polyhedral complex, $\Delta_k(W)$, such that $\Delta_k(W)$ is homotopy equivalent to $M_{\mathcal{W}_{n,k}}$. Moreover, the face poset of $\Delta_k(W)$ is a subposet of the poset of cosets of parabolic subgroups of W, ordered by inclusion. It is obtained by removing all upper order ideals generated by any uG, where G is an irreducible parabolic subgroup of rank $k-1$.*

2.2 Discrete Morse theory

Robin Forman's discrete Morse theory provides a means to calculate the homology groups of a simplicial or CW complex by studying the face poset of the complex. Informally, the goal is to find an acyclic matching on the face poset of the complex. Then the main results of discrete Morse theory state that the original complex is homotopy equivalent to a complex that has one cell of dimension i for each unmatched element in the poset on level $i-1$. We will see that this is especially useful if there are unmatched elements on only one level, or if there is a gap between the levels which share unmatched elements. In these cases, we are able to easily calculate the homology groups of the complex.

We begin by presenting the definition of an acyclic matching that we will use.

Definition 2.3 ([9], Definition 11.1) *Let P be a poset.*

1. *A partial matching in P is a partial matching in the underlying graph of the Hasse diagram of P, i.e., it is a subset $M \subseteq P \times P$ such that*
   - *$(a,b) \in M$ implies $b \succ a$ (b covers a);*
   - *each $a \in P$ belongs to at most one element in M.*

   When $(a,b) \in M$ we write $a = d(b)$ and $b = u(a)$.

2. *A partial matching on P is called acyclic if there does not exist a cycle*

$$b_1 \succ d(b_1) \prec b_2 \succ d(b_2) \prec \cdots \prec b_n \succ d(b_n) \prec b_1$$

   with $n > 2$ and all $b_i \in P$ being distinct.

The second condition may be thought of in the following way. Consider the Hasse diagram of P as a directed graph, with all of the edges oriented downwards, from larger to smaller. Now, if $(a,b) \in M$, we change the orientation of the edge connecting a and b. The matching is called acyclic if the directed graph obtained in the above manner is acyclic. Oftentimes the Cluster Lemma, or Patchwork Theorem, is used to construct acyclic matchings.

Lemma 2.4 ([9], Theorem 11.10) *Assume that $\varphi : P \to Q$ is an order-preserving map, and assume that we have acyclic matchings on subposets $\varphi^{-1}(q)$ for all $q \in Q$. Then the union of these matchings is itself an acyclic matching on P.*

The proof that our matching for the k-parabolic arrangements is acyclic will rely on Lemma 2.4. In application, one tries to break a poset P into 'smaller' pieces, and place an acyclic matching on these pieces. In our case, we will give the matching, and then show that it restricts to fibers, to help simplify the proof that the matching is acyclic.

Given an acyclic matching M on a poset P, we say the elements in $P \setminus M$ are critical. The fundamental theorem of discrete Morse theory states that the critical elements correspond to the cells or simplices of a new complex that the original complex is homotopy equivalent to.

Theorem 2.5 ([9], Theorem 11.13) *Let Δ be a polyhedral complex, and let M be an acyclic matching on $\mathcal{F}(\Delta) \setminus \{\hat{0}\}$. Let c_i denote the number of critical i-dimensional cells of Δ.*

(a) *If the critical cells form a subcomplex Δ_c of Δ, then there exists a sequence of cellular collapses leading from Δ to Δ_c.*

(b) *In general, the space Δ is homotopy equivalent to Δ_c, where Δ_c is a CW complex with c_i cells in dimension i.*

(c) *There is a natural indexing of cells of Δ_c with the critical cells of Δ such that for any two cells σ and τ of Δ_c satisfying dim $\sigma =$ dim $\tau + 1$, the incidence number $[\tau : \sigma]$ is given by*

$$[\tau : \sigma] = \sum_c w(c).$$

Here the sum is taken over all alternating paths c connecting σ with τ, i.e., over all sequences $c = (\sigma, a_1, u(a_1), \ldots, a_t, u(a_t), \tau)$ such that $\sigma \succ a_1$, $u(a_t) \succ \tau$, and $u(a_i) \succ a_{i+1}$, for $i = 1, ..., a_{t-1}$. For such an alternating path, the quantity $w(c)$ is defined by

$$w(c) := (-1)^t [a_1 : \sigma][\tau : u(a_t)] \prod_{i=1}^{t} [a_i : u(a_i)] \prod_{i=1}^{t-1} [a_{i+1} : u(a_i)]$$

where the incidence numbers in the right-hand side are taken in the complex Δ.

3 The type-A and type-B k-equal arrangements

In this section we give two examples of our usage of discrete Morse theory on k-parabolic subspace arrangements. The examples we have chosen, the type-A k-equal arrangement ($\mathcal{A}_{n,k}$) and the type-B k-equal arrangement ($\mathcal{B}_{n,k,k-1}$) have been studied and should be familiar to many readers. We have chosen to present these two examples first because we believe they give the most intuitive look at our matching, and because our matching can be used to obtain new results regarding these arrangements.

3.1 Acyclic matching and homology results for the k-equal arrangement

We start with the type-A k-equal arrangement. This arrangement, described in the introduction, is embedded in the Coxeter arrangement, $\mathcal{H}(A_n)$. It is well known that the face lattice $\mathcal{F}(\mathcal{H}(A_n))$ may be thought of as the poset of all set compositions of $[n+1]$ with reverse refinement as the partial order.

We need to obtain a combinatorial description of the face lattice of $\Delta_k(A_n)$. Since the face poset of the Permutahedron is dual to $\mathcal{F}(\mathcal{H}(A_n))$, first we reverse the partial order. Then we consider faces of the Permutahedron whose corresponding set compositions have a block of size k and remove the upper order ideal of these elements. This will leave us with a subposet of $\mathcal{F}(\mathcal{H}(A_n))$ in which all elements have blocks of size at most $k-1$. In order to construct a matching on the face poset of the complement, $\mathcal{F}(\Delta_k(W))$ we need the following definition.

Definition 3.1 *Given two sets $S, T \in [n+1]$ with $S \cap T = \emptyset$, we say that there is a descent from set S to set T if $\max(S) > \min(T)$. Otherwise we say there is an ascent from S to T.*

We are now ready to construct a matching M on $\mathcal{F}(\Delta_k(W))$. The matching is given by the following algorithm: Given an element $(B_1, B_2, \ldots, B_t)$ we consider pairs of adjacent blocks B_i and B_{i+1}. We start with $i = 1$.

1. If B_i is not a singleton, we match
$$(B_1, \ldots, B_i, B_{i+1}, \ldots, B_t)$$
with
$$(B_1, \ldots, B_{i-1}, \{\min(B_i)\}, B_i \setminus \{\min(B_i)\}, \ldots, B_t).$$
2. If there is a descent from B_i to B_{i+1}, we set $i = i + 1$ and start over at step one.
3. If $|B_{i+1}| = k - 1$, then we set $i = i + 2$ and start over at step one.
4. We match
$$(B_1, \ldots, B_i, B_{i+1}, \ldots, B_t)$$
with the element
$$(B_1, \ldots, B_i \cup B_{i+1}, \ldots, B_t).$$

Note that it is only possible to match elements that differ by the addition or removal of a singleton and hence are on adjacent levels of the poset. The algorithm finishes when a match is found or we reach $i = t$. In the latter case, we have identified a critical element.

Proposition 3.2 *The matching M described above is acyclic.*

The result follows from the general case, Proposition 4.1. The elements that are unmatched have a series of singletons with a descent between each adjacent pair, followed by an ascent to a size $k-1$ block, followed by a series of singletons with a descent between each adjacent pair, followed by an ascent to a size $k-1$ block, etc. An example of the matching along with some critical elements is shown in Figure 1. In the example, there are three elements on the same level of the poset. The first is matched with an element above by merging two blocks. The second is matched with an element below by splitting a block, and the third is a critical element. In the example, $n = 8$ and $k = 4$.

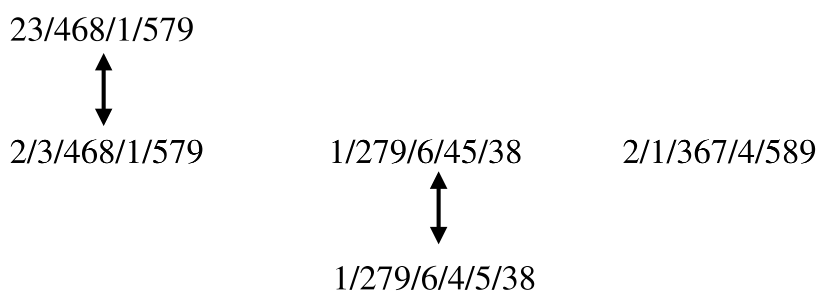


Fig. 1: A matching between elements in $\Delta_4(A_8)$

Also note that critical elements may only occur on levels that are a multiple of $k-2$. By Theorem 2.5, we can already conclude that $H_i(M_{\mathcal{A}_{n,k}})$ is trivial when i is not a multiple of $k-2$. In the case where $k > 3$ we also know that the non-trivial homology groups are free. We see this in the following way. Suppose there are $c_{j(k-2)}$ critical elements on level $j(k-2)$. These correspond to cells of dimension $j(k-2)$

in a CW complex that is homotopy equivalent to $M_{\mathcal{A}_{n,k}}$. Furthermore, there are no critical elements on levels $j(k-2)-1$ or $j(k-2)+1$ and hence no cells of these dimensions. Thus, we have that the chain groups $C_{j(k-2)-1}$ and $C_{j(k-2)+1}$ are trivial and $C_{j(k-2)}$ is free abelian of rank $c_{j(k-2)}$. The boundary maps $\partial_{j(k-2)+1}$ and $\partial_{j(k-2)}$ can only be the trivial map, which implies $H_{j(k-2)}(M_{\mathcal{A}_{n,k}}) \cong C_{j(k-2)}$.

It remains to calculate the number of critical cells at each level. Let j be an integer such that $0 \leq j \leq n/k$. Then the number of unmatched cells in dimension $j(k-2)$ is given by:

$$\sum_{\substack{i_0+\cdots+i_j=n \\ i_m \geq k, \forall 1 \leq m \leq j}} \binom{n}{i_0, \ldots, i_j} \prod_{m=1}^{j} \binom{i_m - 1}{k-1}$$

where the sum is over all integer compositions of n into $j+1$ parts, such that each part, with the exception of the first part, has size at least k. In all other dimensions there are no critical cells. The formula comes from the following: consider a composition of $[n]$ into $j+1$ parts whose sizes are given by $i_0, \ldots, i_k$. For each block, besides the first one, take $k-1$ elements that are not the minimum of that part. Make this a block, and place all other elements of that block as singletons in descreasing order. Finally partition the first block into singletons and append them to the end of the composition in decreasing order. Clearly this gives all set compositions that meet our criteria for not being matched.

Combining the results above, Theorem 2.5 and Lemma 2.2, we have the following.

Theorem 3.3 *The homology groups $H_i(M_{\mathcal{A}_{n,k}})$ are non-trivial only when $i = j(k-2)$, for $j \leq \lfloor \frac{n}{k} \rfloor$. Furthermore, $H_{j(k-2)}(M_{\mathcal{A}_{n,k}})$ is free abelian of rank*

$$\sum_{\substack{i_0+\cdots+i_j=n \\ i_m \geq k, \forall 1 \leq m \leq j}} \binom{n}{i_0, \ldots, i_j} \prod_{m=1}^{j} \binom{i_m - 1}{k-1},$$

where the sum is over all integer compositions of n into $j+1$ parts, such that each part, with the exception of the first part, has size at least k.

Note that the above formula is new, and simpler than previous formulas obtained. The case $j = 1$ specializes to a formula previously known by Björner and Welker [6]. The case where $k = 3$ is the most difficult case. We will discuss this case for general W later.

3.2 Acyclic matching and homology results for $\mathcal{B}_{n,k}$

We now turn to the type-B k-equal arrangement, $\mathcal{B}_{n,k,h}$. The arrangement $\mathcal{B}_{n,k,h}$ has subspaces given by

$$\pm x_{i_1} = \cdots = \pm x_{i_k} \text{ as well as } x_{j_1} = \cdots = x_{j_h} = 0$$

This arrangement is embedded in the type-B Coxeter arrangement, $\mathcal{H}(B_n)$, and the face lattice $\mathcal{F}(\mathcal{H}(B_n))$ has a description in terms of set compositions of $\{0, 1, \ldots, n, \bar{1}, \ldots, \bar{n}\}$. For technical reasons, we will use the linear order $\bar{n} < \cdots < \bar{1} < 0 < 1 < \cdots < n$.

A type-B set composition consists of a set composition of $[0, n]$ such that 0 is in the *first* block (henceforth called the zero block). The numbers in the 0 block are all unbarred, and in the non-zero blocks we may replace i with $\bar{i}$. The order is reverse refinement, and when blocks are merged the bars do not change, except if a block is merged with the zero block. In this latter case all elements become unbarred. We again say there is a descent from B_i to B_{i+1} if $\max(B_i) > \min(B_{i+1})$.

As in the type-A case, we obtain the face poset, $\mathcal{F}(\Delta_k(B_n))$, by removing all type-$B$ set compositions that have blocks, including the zero block, of size k or greater. Thus we only study type-B set compositions with blocks of size at most $k-1$. Recall that we also reverse the partial order, so now the partial order is refinement.

The matching algorithm that we give is sort of dual to the one in the type-A case. Instead we start at the last block and work our way towards the zero block. Given an element $(B_0, B_1, B_2, \ldots, B_t)$ we consider pairs of adjacent blocks B_i and B_{i+1}. We start with $i = t$.

1. If B_i is not a singleton, we match
$$(B_0, \ldots, B_i, B_{i+1}, \ldots, B_t)$$
with
$$(B_0, \ldots, B_{i-1}, \{B_i \setminus \{\max(B_i)\}, \{\max(B_i)\}, \ldots, B_t).$$
2. If there is a descent from B_{i-1} to B_i, we set $i = i-1$ and start over again at step one.
3. If $|B_{i-1}| = k-1$, then we set $i = i-2$ and start over again at step one.
4. We match
$$(B_0, \ldots, B_{i-1}, B_i, \ldots, B_t)$$
with the element
$$(B_0, \ldots, B_{i-1} \cup B_i, \ldots, B_t).$$

Again the above algorithm gives an acyclic matching on $\mathcal{F}(\Delta_k(B_n))$. Critical cells have the following properties:

- all blocks are singletons or have size $k-1$,
- every block of size $k-1$ is followed by a singleton, with an ascent between them,
- every pair of adjacent singletons forms a descent.

Again, as in the type-A case, when $k > 3$ we have non-trivial free abelian homology groups only in period $k-2$. We also note here that our restriction of $h = k-1$ may be removed and the matching will work with any h. Removing this restriction takes us out of the class of k-parabolic subspace arrangements and covers all of the $\mathcal{B}_{n,k,h}$ arrangements defined by Björner and Sagan.

The only modification to the matching algorithm is that when considering whether we may merge a singleton and the zero block we check to see if the zero block is of size $< h$, rather than $< k-1$. The condition for being a critical cell involves checking if the 0-block has size 1 or h, rather than $k-1$.

For $h = k$ or $h = 2$, we run into the same issues as the $k = 3$ case, namely that we end up with critical elements on adjacent levels of the poset. In the case where $h \neq 2, k$ it is easy to see that the non-trivial homology groups are free abelian and appear in period $t(k-2)$ and $t(k-2) + h - 1$, recovering the periodicity results obtained by Björner and Sagan [4]. Moreover, by using a counting argument similar to the type-A case, we obtain a new formula for the Betti numbers. Using the same reasoning as in the type-A case, we obtain the following result.

Theorem 3.4 *The homology groups $H_i(M_{\mathcal{B}_{n,k}})$ are non-trivial only when $i = j(k-2)$, for $j \leq \lfloor \frac{n}{k} \rfloor$. Furthermore, $H_{j(k-2)}(M_{\mathcal{B}_{n,k}})$ is free abelian of rank*

$$\sum_{\substack{i_0+\cdots+i_j=n \\ i_m \geq k, \forall 1 \leq m \leq j}} \binom{n}{i_0,\ldots,i_j} \prod_{m=1}^{j} 2^{i_m} \binom{i_m-1}{k-1} + \sum_{\substack{i_0+\cdots+i_j=n \\ i_1 \geq k-1 \\ i_m \geq k, \forall 2 \leq m \leq j}} \binom{n}{i_0,\ldots,i_j} \binom{i_1-1}{k-2} \prod_{m=2}^{j} 2^{i_m} \binom{i_m-1}{k-1},$$

where the sums are over all integer compositions of n into $j+1$ parts, such that each part, with the exception of the first two parts, has size at least k. The first part has size at least k in the first sum, and at least $k-1$ in the second sum.

We also mention that there is a modification of the matching algorithm for the $\mathcal{D}_{n,k}$-arrangement as defined by Björner and Sagan. However, the modification requires making additional rules, and thus is not related to our matching for the general case. We do not give it here, but it will appear in the full version of the paper.

4 A matching for any k-parabolic subspace arrangement

We now present an algorithm to produce an acyclic matching for the complement of any k-parabolic subspace arrangement. First, we fix a finite real reflection group W with corresponding root system Φ. Let Δ be a simple system in Φ and S the set of generators corresponding to roots in Δ. Given $I \subseteq \Delta$, we denote by W_I the subgroup of W generated by the elements of I. Finally, fix a linear order $\{s_1, \ldots, s_n\}$ on S.

We again consider the Coxeter arrangement $\mathcal{H}(W)$, and the corresponding complex $\Delta_k(W)$ from Lemma 2.2. We will define an acyclic matching on $\mathcal{F}(\Delta_k(W))$.

Recall that the length ℓ of an element w of W is the length of any reduced expression for w in terms of the generators. Given a coset vW_I, there is a unique element $u \in vW_I$ of minimal length. We let $D(w) = \{s \in S : \ell(ws) < \ell(w)\}$ denote the descents of w. Given a set I and a reflection $s \in S$, we let $P(I,s) = \{J : J \subseteq I \cup \{s\}, W_J \in \mathcal{P}_{n,k}(W)\}$. We linearly order $P(I,s)$ by $J \leq K$ if $J \subseteq K$ or $\min(J \setminus K) < \min(K \setminus J)$. Finally, let $I^{\perp} = \{t \in S : st = ts, \forall s \in I\}$.

Now, consider a coset uW_I, where u is the element of minimal length in uW_I. The matching algorithm is as follows:

Let $L = S$, $X = \emptyset$.
While $L \neq \emptyset$
    Let $s = \min L$
    If $s \in D(u)$
        Set $L = L - s$
    Else If $s \in I$
        Set $I' = I - s$, **and Return** $uW_{I'}$
    Else If $P(I,s) \neq \emptyset$
        Set $L = (\min(P(I,s)) - s)^{\perp}$ **And** $X = X + s$
    Else
        Set $I' = I + s$, **Return** $uW_{I'}$
End While

Return uW_I

Given a coset uW_I, we will refer to the output of the algorithm as $M(uW_I)$. We will let $X(uW_I)$ be the set of elements in X when the algorithm terminates.

If one takes the natural linear order $s_i < s_j$ for all $i < j$ on A_n, this algorithm specializes to the one made earlier for the type-A case. If one takes the linear order $s_j < s_i$ for all $i < j$ on B_n, this algorithm specializes to the one made earlier for the type-B case.

Proposition 4.1 *The algorithm above gives an acyclic matching M.*

The proof involves using Lemma 2.4 twice. The first time involves the partial order on W that is dual to the right weak order. That is, we take the transitive closure of the cover relations $u \succ us$ for all $u \in W, s \in S \setminus D(u)$. We will denote the resulting poset by W^*. The second application of Lemma 2.4 will use the Boolean poset on S.

Proof: We only sketch the ideas of the proof. The full version of the paper will have complete details. It is not hard to see that the map φ which sends a coset uW_I to its minimal coset representative is an order-perserving map between $\mathcal{F}(\Delta_k(W))$ and W^*. From properties of minimal coset representatives, one can deduce that for any coset uW_I, $\varphi(uW_I) = \varphi(M(uW_I))$. Thus, by the Patchwork Lemma, it suffices to show that the algorithm gives an acyclic matching on the fibers of φ.

So fix an element $w \in W$. Then we attempt to apply 2.4 a second time. For each fiber $\varphi^{-1}(w)$, we consider the map ψ_w to the boolean lattice on S given by sending uW_I to $X(uW_I)$. A careful study of the algorithm shows that this map is order perserving, and that for any coset $uW_I \in \varphi^{-1}(w)$, $\psi(uW_I) = \psi(M(uW_I))$. Thus, we only have to show that the algorithm gives an acyclic matching on the fibers of ψ_w. It is not hard to see that it is a matching. We give an example of the argument for acyclicity. $\square$

Figure 2 is an example that the matching is acyclic on the fibers of ψ_w. We consider A_6, with $k = 3$, w the identity permutation, and the set $X = \{s_1\}$. For type-A, the minimum length coset representative for uW_I is obtained by refining the set composition so that elements in each block are in increasing order. Every directed cycle has to have alternating edges in the matching. Consider the edge $1/23/4/5/67 - 1/23/45/67$. This is an edge in the matching in the fiber $\psi_w^{-1}(X)$. Consider trying to make a directed cycle with this edge. Let the next vertex in the directed cycle be $1/23/45/6/7$. The next edge in the 'cycle' must be from the matching. However, $1/23/45/6/7$ is matched to $1/23/4/5/6/7$, and as we see in figure 2, this edge is pointed the wrong way. We prove acyclicity by showing that this situation always happens on the fibers of ψ_w. In general, a matched edge is always of the form $uW_I \to uW_{I+s}$, where u is the coset representative of minimal length, and $s \in S$. Let $t \in I$, and consider uW_{I+s-t}, and assume t was chosen so that $X(uW_{I+s-t}) = X(uW_I)$. Then the algorithm will match uW_{I+s-t} with uW_{I-t}, resulting in a picture similar to our figure. Thus, we cannot get a directed cycle in the matching when restricted to $\psi_w^{-1}(X)$, for any $w \in W, X \subseteq S$.

We would like to understand the structure of unmatched elements. A simple example will show us that some linear orders on S are more useful than others. For instance, take A_5 and $k = 4$, with linear order $s_3 < s_2 < s_4 < s_1 < s_5$, $k = 4$. Consider the set composition $1/23/456$, which corresponds to the coset $A_{\{s_2,s_4,s_5\}}$. Under the algorithm, this coset is unmatched, yet it is of dimension 3, which is not even. Hence a simple proof of a periodicity condition requires being more specific when picking a linear order for S, as the periodicity results followed easily from the linear order we considered previously for A_n.

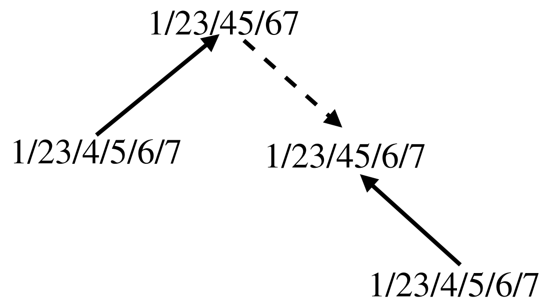


Fig. 2: An example of acyclicity - solid lines are matched edges, dotted lines are not

We restrict ourselves to the case where W is irreducible for a moment. In this case, consider the dynkin diagram D for W, and let P be a maximum length path in D. Consider a linear order of S such that adjacent vertices of P are adjacent in the order, and the only vertex of $D - P$ is at the end of the linear order, if it exists. A very careful analysis of the algorithm on this linear order reveals that the unmatched cells occur in dimensions that are a multiple of $k - 2$.

Now given any finite real reflection group W, we fix a linear order on each connected component of its Dynkin diagram in a way similar to above. Then we take a linear extension of these orders, and we see again that unmatched cells occur dimensions that are a multiple of $k - 2$.

So combining these observations with Theorem 2.5 and Lemma 2.2, we obtain:

Theorem 4.2 *The homology groups $H_i(M_{\mathcal{W}_{n,k}})$ are non-trivial only when $i = j(k-2)$, for $j \leq \lfloor \frac{n}{k} \rfloor$. Furthermore, $H_{j(k-2)}(M_{\mathcal{W}_{n,k}})$ is free abelian and has rank given by the number of unmatched elements of rank $j(k-2)$ in our matching M.*

Of course, the case $k = 3$ is challenging. Much like in the type-A and type-B case, there are unmatched cells on every level. However, an involution can be used to show that the summation formula in Theorem 2.5, part c, are all zero. Hence the boundary map is the zero map, and the critical cells still index a basis for the homology groups.

We remark that the number of critical cells of dimension 0 is 1. The only unmatched cell is $w_0W_\emptyset$, where w_0 is the element of maximum length in W. We also note that for exceptional groups, and arbitrary $k > 3$, we can compute the Betti numbers without studying the matching, except possibly the case when $W = E_8, k = 4$. In all other cases the peridocity conditions, combined with use of group theory to compute the number of cells of $\Delta_k(W)$ of dimension i, can be used to determine the Betti numbers. A table of these numbers will be included in the full version.

We close with a few open questions. First, similar arguments give a basis for the cohomology. It would be nice to understand the cohomology ring structure in terms of this basis, although this is challenging. Secondly, we note that discrete Morse theory allows us to obtain a minimal cell complex with the same homotopy type as $\mathcal{W}_{n,k}$ for $k > 3$. However, with current methods it is complicated to understand the attachment maps of this minimal complex. We do know that, for $k = 3$, and large n, the resulting complex is not homotopy equivalent to a wedge of spheres, but we currently do not have more information than that. Also, it is still an open problem if the intersection lattice of $\mathcal{W}_{n,k}$ is shellable or not. Even though our approach avoids this question, it would be nice if there was another proof of our results, using lexicographic shellability.

Finally, our results have some application to linear decision trees. However, this application no longer has the same simplicity as the k-equal problem, and our results do not add anything new to the theory of linear decision trees. For length considerations, we mention more regarding linear decision trees in the full paper.

References

[1] Hélène Barcelo and Edwin Ihrig, *Lattices of parabolic subgroups in connection with hyperplane arrangements*, J. Algebraic Combin. **9** (1999), no. 1, 5–24. MR MR1676736 (2000g:52023)

[2] Hélène Barcelo, Christopher Severs, and Jacob A. White, *k-parabolic subspace arrangements*, (2009).

[3] Anders Björner and László Lovász, *Linear decision trees, subspace arrangements and Möbius functions*, J. Amer. Math. Soc. **7** (1994), no. 3, 677–706. MR MR1243770 (95e:52024)

[4] Anders Björner and Bruce E. Sagan, *Subspace arrangements of type B_n and D_n*, J. Algebraic Combin. **5** (1996), no. 4, 291–314. MR MR1406454 (97g:52028)

[5] Anders Björner and Michelle L. Wachs, *Shellable nonpure complexes and posets. I*, Trans. Amer. Math. Soc. **348** (1996), no. 4, 1299–1327. MR MR1333388 (96i:06008)

[6] Anders Björner and Volkmar Welker, *The homology of "k-equal" manifolds and related partition lattices*, Adv. Math. **110** (1995), no. 2, 277–313. MR MR1317619 (95m:52029)

[7] Eva Maria Feichtner and Dmitry N. Kozlov, *On subspace arrangements of type D*, Discrete Math. **210** (2000), no. 1-3, 27–54, Formal power series and algebraic combinatorics (Minneapolis, MN, 1996). MR MR1731606 (2001k:52039)

[8] Mark Goresky and Robert MacPherson, *Stratified Morse theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 14, Springer-Verlag, Berlin, 1988. MR MR932724 (90d:57039)

[9] Dmitry Kozlov, *Combinatorial algebraic topology*, Algorithms and Computation in Mathematics, vol. 21, Springer, Berlin, 2008. MR MR2361455 (2008j:55001)

[10] Dmitry N. Kozlov, *General lexicographic shellability and orbit arrangements*, Ann. Comb. **1** (1997), no. 1, 67–90.

[11] Peter Orlik, *Complements of subspace arrangements*, J. Algebraic Geom. **1** (1992), no. 1, 147–156.

Part III
Posters – Affiches

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 325–335

Compositions and samples of geometric random variables with constrained multiplicities

Margaret Archibald[1] and Arnold Knopfmacher[2†] and Toufik Mansour[3]

[1] *Laboratory of Foundational Aspects of Computer Science, Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7701, South Africa*
margaret.archibald@uct.ac.za
[2] *The John Knopfmacher Centre for Applicable Analysis and Number Theory, Department of Mathematics, University of the Witwatersrand, P. O. Wits, 2050, Johannesburg, South Africa*
arnold.knopfmacher@wits.ac.za
[3] *Department of Mathematics, University of Haifa, 31905 Haifa, Israel*
toufik@math.haifa.ac.il

received 15 Nov 2009, revised 27th July 2010, accepted tomorrow.

Abstract. We investigate the probability that a random composition (ordered partition) of the positive integer n has no parts occurring exactly j times, where j belongs to a specified finite *'forbidden set'* A of multiplicities. This probability is also studied in the related case of samples $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_n)$ of independent, identically distributed random variables with a geometric distribution.

Résumé. Nous examinons la probabilité qu'une composition faite au hasard (une partition ordonnée) du nombre entier positif n n'a pas de partie qui arrivent exactement j fois, où j appartient à une série interdite, finie et spécifié A de multiplicités. Cette probabilité est aussi étudiée dans le cas des suites $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_n)$ de variables aléatoires identiquement distribués et indépendants avec une distribution géométrique.

Keywords: compositions, generating functions, geometric random variable, Mellin transform, Poisson transform, multiplicity

1 Introduction

In this paper we derive generating functions for random compositions (ordered partitions) of a positive integer n in which no parts occur exactly j times, where j belongs to a specified finite *'forbidden set'* A of multiplicities. For notational convenience we shall refer to such compositions as being 'A*-avoiding*'. We go on to find the probabilities that compositions and samples of geometric random variables are A-avoiding.

†This material is based upon work supported by the National Research Foundation under grant number 2053740

As a simple example of a forbidden set, we may wish to consider a sample where none of the n elements occur exactly a times. In this case $A = \{a\}$. Another example is when a letter can occur only a times or more (or not at all), i.e., $A = \{1, 2, \ldots, a-1\}$, for $a \geq 2$. Note that we do not allow 0 in the forbidden set.

Previously in [6, 12], geometric samples with the multiplicity constraint that certain values must occur at least once were studied. These were called 'gap-free' and 'complete' samples. A gap-free sample has elements whose values form an interval, namely if elements 2 and 6 are in the sample, then so are 3, 4 and 5. A complete sample is gap-free with minimal element 1.

In this paper we drop the 'interval' restriction, hence no value 0 in our forbidden sets. Here we are more interested in the number of times the elements do occur than in the values of the elements. However, in Section 2, the idea of forbidden sets is generalised even further when we allocate each value a different forbidden set. For example, one could provide the restriction that the value 2 is not allowed to occur once, but that the number of times that 5 can occur is anything except $2, 3$ or 6 times. We denote the forbidden set for the value i by A_i, so in this case, we have $A_2 = \{1\}$ and $A_5 = \{2, 3, 6\}$.

The paper begins with a discussion on compositions (Section 2), where explicit generating functions are derived for A-avoiding compositions and particular forbidden sets are highlighted. In Section 3, the link between compositions and samples of geometric random variables is explained. Section 4 is devoted to geometric samples, and Theorem 2 gives the probability that a geometric sample is A-avoiding, along with some further examples of specific forbidden sets. Finally in Section 5, we state the result for compositions - i.e., the probability that a random composition of n is A-avoiding. Some of the longer proofs, in particular, the proof of Theorem 2 in Section 4 will be detailed in the full version of this paper.

2 Compositions

In this section we investigate the generating function for the number of A-avoiding compositions of n, that is the number of compositions of n such that each part does not appear exactly j times, where $j \in A$. We then go on to generalise this by allowing a different forbidden set for each value, as described in the introduction.

Let $C_{A,d}(x; m)$ be the generating function for the number of A-avoiding compositions of n with exactly m parts from the set $[d] = \{1, 2, \ldots, d\}$. If σ is any A-avoiding composition with m parts in $[d]$, then σ contains the part d exactly j times with $j \notin A$ and $0 \leq j \leq m$. Deleting the parts that equal to d from σ we get an A-avoiding composition σ' of $m - j$ parts in $[d-1]$. Thus, rewriting the above rule in terms of generating functions we get that

$$C_{A,d}(x; m) = \sum_{\substack{j=0 \\ j \notin A}}^{m} \binom{m}{j} x^{dj} C_{A,d-1}(x; m-j),$$

which is equivalent to

$$\frac{C_{A,d}(x; m)}{m!} = \sum_{\substack{j=0 \\ j \notin A}}^{m} \frac{x^{dj}}{j!} \frac{C_{A,d-1}(x; m-j)}{(m-j)!}. \tag{1}$$

We denote the exponential generating function for the sequence $C_{A,d}(x;m)$ by $C_{A,d}(x,y)$, that is,

$$C_{A,d}(x,y)=\sum_{m\geq 0}C_{A,d}(x;m)\frac{y^m}{m!}.$$

Therefore, the recurrence in (1) can be written as

$$C_{A,d}(x,y)=C_{A,d-1}(x,y)\left(e^{x^d y}-\sum_{j\in A}\frac{x^{dj}y^j}{j!}\right),$$

which implies that

$$C_{A,d}(x,y)=\prod_{k=1}^{d}\left(e^{x^k y}-\sum_{j\in A}\frac{x^{kj}y^j}{j!}\right),$$

for all $d\geq 1$. Hence, we can state the following result.

Proposition 1 *The generating function $C_A(x,y)=\sum_{m\geq 0}C_A(x;m)\frac{y^m}{m!}$ is given by*

$$C_A(x,y)=\prod_{k\geq 1}\left(e^{x^k y}-\sum_{j\in A}\frac{x^{kj}y^j}{j!}\right),$$

where $C_A(x;m)$ is the generating function for the number of A-avoiding compositions of n with exactly m parts in $\mathbb{N}$.

Let $C_A(n,m)$ be the number of A-avoiding compositions of n with m parts and $C_A(n)=\sum_{m\geq 1}C_A(n,m)$ be the number of A-avoiding compositions of n.

Corollary 1 *The generating function $C_A(x)=\sum_{n\geq 0}C_A(n)x^n$ is given by*

$$C_A(x)=\int_0^\infty e^{-y}\prod_{k\geq 1}\left(e^{x^k y}-\sum_{j\in A}\frac{x^{kj}y^j}{j!}\right)dy.$$

Proof: We use the fact that $\int_0^\infty e^{-y}y^m dy=m!$. Then

$$\int_0^\infty e^{-y}C_A(x,y)dy=\sum_{n\geq 0}x^n\sum_{m\geq 0}\frac{C_A(n,m)}{m!}\int_0^\infty y^m e^{-y}dy=\sum_{n\geq 0}C_A(n)x^n.$$

□

Example 1 *Let $A_i=\{1\}$ for all i, then the above proposition gives that*

$$C_{\{1\}}(x,y)=\prod_{k\geq 1}(e^{x^k y}-x^k y).$$

and Corollary 1 gives

$$C_{\{1\}}(x)=\int_0^\infty e^{-y}\prod_{k\geq 1}\left(e^{x^k y}-x^k y\right)dy.$$

Similar techniques as before show the following general result.

Proposition 2 *The generating function $D_{A_1,A_2,\dots}(x,y)=\sum_{m\geq 0} D_{A_1,A_2,\dots}(x;m)\frac{y^m}{m!}$ is given by*

$$D_{A_1,A_2,\dots}(x,y)=\prod_{k\geq 1}\left(e^{x^k y}-\sum_{j\in A_k}\frac{x^{kj}y^j}{j!}\right),$$

where $D_{A_1,A_2,\dots}(x;m)$ is the generating function for the number of compositions σ of n with exactly m parts in $\mathbb{N}$ such that if σ contains the part i exactly d_i times, then $d_i\notin A_i$. Furthermore,

$$D_{A_1,A_2,\dots}(x)=\int_0^\infty e^{-y}\prod_{k\geq 1}\left(e^{x^k y}-\sum_{j\in A_k}\frac{x^{kj}y^j}{j!}\right)dy.$$

Example 2 *For instance, let $A_1=\{1\}$ and $A_i=\emptyset$ for $i\geq 2$, then the above proposition gives that*

$$F(x,y)=D_{\{1\},\emptyset,\emptyset,\dots}(x,y)=(e^{xy}-xy)e^{\frac{x^2y}{1-x}}.$$

If we expand $F(x,y)$ as a power series at $x=y=0$, then we obtain that

$$F(x,y)=\sum_{j\geq 0}\frac{x^jy^j}{j!(1-x)^j}-xy\sum_{j\geq 0}\frac{x^{2j}y^j}{j!(1-x)^j},$$

which implies that

$$D_{\{1\},\emptyset,\emptyset,\dots}(x;m)=\frac{x^m}{(1-x)^m}-m\frac{x^{2m-1}}{(1-x)^{m-1}}.$$

Summing over all $m\geq 0$, we get that the ordinary generating function for the number of compositions σ of n such that the number occurrence of the part 1 in σ does not equal 1 is given by

$$\frac{1-x}{1-2x}-\frac{x(1-x)^2}{(1-x-x^2)^2}.$$

Note that it is not hard to generalize the above enumeration to obtain that the ordinary generating function for the number of compositions σ of n such that the number occurrence of the part 1 in σ does not equal ℓ is given by

$$\frac{1-x}{1-2x}-\ell!\frac{x^\ell(1-x)^{\ell+1}}{(1-x-x^2)^{\ell+1}}.$$

Example 3 *For instance, let $A_1=A_2=\{1\}$ and $A_i=\emptyset$ for $i\geq 3$, then the above proposition gives that*

$$G(x,y)=D_{\{1\},\{1\},\emptyset,\emptyset,\dots}(x,y)=(e^{xy}-xy)(e^{x^2y}-x^2y)e^{\frac{x^3y}{1-x}}.$$

If we expand $G(x,y)$ as a power series at $x=y=0$, then we find that

$$\begin{aligned}&D_{\{1\},\{1\},\emptyset,\emptyset,\dots}(x;m)\\&=\frac{x}{(1-x)^m}-m\frac{x^{m+1}(1-x+x^2)^{m-1}}{(1-x)^{m-1}}-m\frac{x^{2m-1}}{(1-x)^{m-1}}-m(m-1)\frac{x^{3m-3}}{(1-x)^{m-2}}.\end{aligned}$$

Summing over all $m \geq 0$, we get that the ordinary generating function for the number of compositions σ of n such that the number occurrence of the part i, $i = 1, 2$, in σ does not equal 1 is given by

$$\frac{1-x}{1-2x} - \frac{x(1-x)^2}{(1-x-x^2)^2} - \frac{x^2(1-x)^2}{(1-2x+x^2-x^3)^2} + \frac{2x^3(1-x)^3}{(1-x-x^3)^3}.$$

Theorem 1 *Fix $a \in \mathbb{N}$. Let $A_i = \{a\}$ for all $i = 1, 2, \ldots, \ell$ and $A_{\ell+i} = \emptyset$ for all $i \geq 1$. The ordinary generating function for the number of compositions π of n such that π does not contain part i exactly a times for all $i = 1, 2, \ldots, \ell$ is given by*

$$\sum_{m\geq 0} D_{A_1,A_2,\ldots}(x;m) = \frac{1-x}{1-2x} + \sum_{j=1}^{\ell} \sum_{1\leq i_1<i_2<\cdots<i_j\leq \ell} (-1)^j \frac{j!x^{a\sum_{k=1}^{j} i_k}}{\left(1-\frac{x}{1-x}+\sum_{k=1}^{j} x^{a i_k}\right)^{j+1}},$$

The proof of this result will be given in the full version of this paper. From the theorem we can deduce the following result.

Corollary 2 *The ordinary generating function for the number of $\{a\}$-avoiding compositions of n is given by*

$$\frac{1-x}{1-2x} + \sum_{j\geq 1} \sum_{B\subseteq\mathbb{N},\,|B|=j} (-1)^j \frac{(aj)!(x^a/a!)^{\sum_{b\in B} b}}{\left(1-\frac{x}{1-x}+\sum_{b\in B} x^b\right)^{aj+1}}.$$

Even in this simple case of $A = \{a\}$ it does not seem easy to find asymptotic estimates for the coefficients from the generating functions appearing in either Corollary 1 or Corollary 2. Instead we will exploit the correspondence between compositions and geometric random variables of parameter $p = 1/2$, as detailed in the next section.

3 Reduction of compositions to geometric samples

In order to derive asymptotic estimates, it will be convenient to adopt a probabilistic viewpoint. That is, rather than think of the proportion of A-avoiding compositions we will equip the set of all compositions of n with the uniform probability measure and will be interested in the probability that a randomly chosen composition of n is A-avoiding. In that setting, compositions of n are closely related to the special case for geometric random variables when $p = 1/2$, as shown in [7, 8] and again in this section.

The starting point for reducing compositions to samples of geometric random variables is the following representation of compositions of n (see e.g., [2]). Consider sequences of n black and white dots subject to the following constraints

(i) the last dot is always black

(ii) each of the remaining $n-1$ dots is black or white.

Then there is a 1-1 correspondence between all such sequences and compositions of n. Namely, part sizes in a composition correspond to "waiting times" for occurrences of black dots. For example, the sequence

$$\underbrace{\bullet}_{1}\ \underbrace{\circ\ \circ\ \bullet}_{3}\ \underbrace{\circ\ \bullet}_{2}\ \underbrace{\bullet}_{1}\ \underbrace{\bullet}_{1}\ \underbrace{\circ\ \bullet}_{2}\ \underbrace{\circ\ \bullet}_{2}$$

represents the composition of 12 into parts $(1,3,2,1,1,2,2)$. As discussed e.g. in [7, 8] this leads to the following representation of random compositions. Let $p = 1/2$ and define

$$\tau = \tau_n = \inf\{k \geq 1 : \; \Gamma_1 + \Gamma_2 + \cdots + \Gamma_k \geq n\}.$$

Then a randomly chosen composition κ of n has distribution given by

$$\kappa = (\Gamma_1, \Gamma_2, \ldots, \Gamma_{\tau-1}, n - \sum_{j=1}^{\tau-1} \Gamma_j) := (\tilde{\Gamma}_1, \tilde{\Gamma}_2, \ldots, \tilde{\Gamma}_\tau).$$

Furthermore, τ has known distribution, namely,

$$\tau \stackrel{d}{=} 1 + \text{Bin}(n-1, \frac{1}{2}),$$

where $\text{Bin}(m,p)$ denotes a binomial random variable with parameters m and p and $\stackrel{d}{=}$ stands for equality in distribution. Hence, τ is heavily concentrated around its mean. Specifically, since $\text{var}(\tau) = \text{var}(\text{Bin}(n-1, 1/2)) = (n-1)/4$, for every $t > 0$ we have (see [1, Section A.1])

$$\mathbb{P}(|\tau - \mathsf{E}\tau| \geq t) \leq 2\exp\{-\frac{2t^2}{n-1}\}.$$

In particular, for $t_n \sim \sqrt{cn \ln n}$,

$$\mathbb{P}(|\tau - \mathsf{E}\tau| \geq t_n) = O\left(\frac{1}{n^{2c}}\right),$$

for any $c > 0$.

Let $\mathbb{P}(\kappa \in \mathcal{C})$ be the probability that a random composition is A-avoiding. We proceed by series of refinements exactly as in [6]. Set m_n^- to be

$$m_n^- = \left\lfloor \frac{n+1}{2} - t_n \right\rfloor.$$

As shown in [6], with overwhelming probability, κ is A-avoiding if and only if the first m_n^- of its parts are A-avoiding. In [6] the property considered is “complete” rather than “A-avoiding”, but the arguments remain unchanged.

Ultimately we obtain, exactly as in [6],

$$\mathbb{P}(\kappa \in \mathcal{C}) = \mathbb{P}((\Gamma_1, \ldots, \Gamma_{m_n^-}) \in \mathcal{C}) + O\left(\frac{\ln^{3/2} n}{\sqrt{n}}\right),$$

thereby reducing the problem to samples of geometric random variables.

4 Geometric random variables

Following the discussion in Section 3 above it is natural to start the investigation for the probability that a composition is A-avoiding with samples of geometric random variables with arbitrary parameter p, where $0 < p < 1$. There is now an extensive literature on the combinatorics of geometric random variables and its applications in Computer Science which includes [3, 5, 6, 11, 12, 13, 14].

Let $\Gamma = (\Gamma_1, \Gamma_2, \ldots, \Gamma_n)$ be a sample of independent identically distributed (i.i.d.) geometric random variables with parameter p, that is, $\mathbb{P}(\Gamma_i = k) = pq^{k-1}$, with $p + q = 1$, where $k = 1, 2, \ldots$ and $i = 1, 2, \ldots, n$. We shall restrict the multiplicity of elements in a sample of length n by prohibiting any occurrences of exactly j entries of a given size, for j a natural number belonging to a specified finite set of excluded numbers A, the forbidden set. We also call such a random sample of n geometric variables A-avoiding.

The method used in [6] can be applied to the problem described above. We start with a recursion for the probabilities that depends on the set A and then use Poissonisation and Mellin transforms followed by de-Poissonisation to obtain our asymptotic estimates.

Using this approach, the following main result for geometric random variables will be proved in the full version of this paper. We define $\chi_k := \frac{2k\pi i}{\ln(1/q)}$.

Theorem 2 *Let A be any finite set of positive integers. The probability p_n that a geometric sample of length n has no letter appearing with multiplicity j, for any $j \in A$ is (asymptotically as $n \to \infty$)*

$$p_n = 1 - \frac{T^*(0)}{\ln(1/q)} - \delta\big(\log_{1/q}(n/q)\big) + O(n^{-1}),$$

with

$$T^*(0) = \sum_{j \in A} p^j \sum_{n \geq 0} p_n q^n \frac{1}{n+j} \binom{n+j}{j} \tag{2}$$

and

$$\delta(x) = \frac{1}{\ln(1/q)} \sum_{k \neq 0} T^*(\chi_k) e^{-2k\pi i x}$$

where

$$T^*(\chi_k) = \sum_{j \in A} \frac{p^j}{j!} \sum_{n \geq 0} p_n \frac{q^n}{n!} \Gamma(n + j + \chi_k), \quad \textit{for} \quad k \in \mathbb{Z} \backslash \{0\}. \tag{3}$$

Here $\delta(x)$ is a periodic function of x with period 1, mean 0 and small amplitude.

The corresponding result for compositions of n is given in Section 5.

4.1 Examples of finite forbidden sets A

In the sections above we mentioned a few specific examples that would satisfy this definition of the forbidden set. Here we simplify the $T^*(0)$ and $T^*(\chi_k)$ formulae from Theorem 2 for a few specific cases. The simplest case for A is a singleton set consisting of one value a. If $A = \{a\}$, then

$$T^*(0) = p^a \sum_{n \geq 0} p_n q^n \frac{1}{n+a} \binom{n+a}{a} \quad \text{and} \quad T^*(\chi_k) = \frac{p^a}{a!} \sum_{n \geq 0} p_n \frac{q^n}{n!} \Gamma(n + a + \chi_k).$$

If we consider the case where $A = \{1, \dots, a-1\}$, then

$$T^*(0) = \sum_{j=1}^{a-1} p^j \sum_{n\geq 0} p_n q^n \frac{1}{n+j}\binom{n+j}{j} \quad \text{and} \quad T^*(\chi_k) = \sum_{j=1}^{a-1} \frac{p^j}{j!} \sum_{n\geq 0} p_n \frac{q^n}{n!} \Gamma(n+j+\chi_k).$$

In particular if we want the probability that no element occurs exactly once (all elements must occur at least twice if they occur at all), we have a main term for p_n of

$$1 - \frac{p}{\ln(1/q)} \sum_{n\geq 0} p_n q^n.$$

This main term is plotted as a function of p in Figure 1.

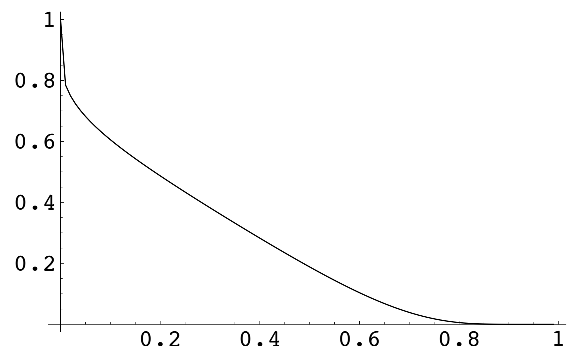


Fig. 1: *Plot of the non-oscillating limit term for* p_n *for* $0 \leq q \leq 1$.

The corresponding picture for the probability that no element occurs exactly twice is given in Figure 2.

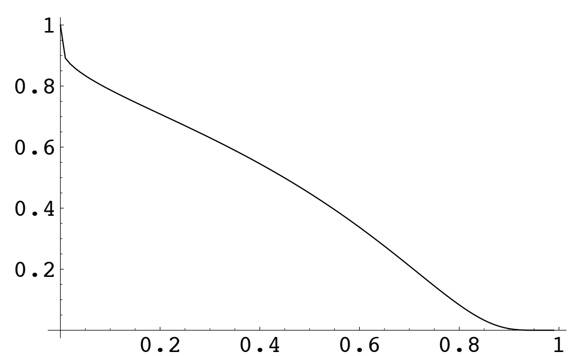


Fig. 2: *Plot of the non-oscillating limit term for* p_n *for* $0 \leq q \leq 1$.

In spite of what the Figures 1 and 2 tend to suggest for q near 1, the main term here is strictly greater than zero for every $0 < p < 1$ as

$$\begin{aligned}T^*(0) &= p^a \sum_{n\geq 0} p_n q^n \frac{1}{n+a}\binom{n+a}{a} \\ &\leq p^a \sum_{n\geq 0} q^n \frac{1}{n+a}\binom{n+a}{a} = p^a \frac{(1-q)^{-a}}{a} \\ &\leq 1 < \ln(1/q).\end{aligned}$$

We observe also that the sequences (p_n) in this section do not have a limit, but exhibit small oscillations where both the period and amplitude of the oscillations depend on p. Such oscillations are almost ubiquitous in problems solved using Mellin transform techniques. For example, Figures 3 and 4 (Section 5) show these oscillations in the case that no element occurs exactly once (twice) when $p = 1/2$.

5 Compositions revisited

From Section 3, we conclude that probabilities for compositions can be reduced to probabilities for samples of geometric random variables. This result together with the special case $p = q = \frac{1}{2}$ in Theorem 2 leads to the following corollary.

Corollary 3 *Let A be any finite set of positive integers. The probability p_n that a composition of n has no part appearing with multiplicity j, for any $j \in A$ is (asymptotically as $n \to \infty$)*

$$p_n = 1 - \frac{T^*(0)}{\ln 2} - \delta(\log_2 n) + O\left(\frac{\ln^{3/2} n}{\sqrt{n}}\right),$$

with

$$T^*(0) = \sum_{j\in A}\left(\frac{1}{2}\right)^j \sum_{n\geq 0} p_n \left(\frac{1}{2}\right)^n \frac{1}{n+j}\binom{n+j}{j} \tag{4}$$

and

$$\delta(x) = \frac{1}{\ln 2}\sum_{k\neq 0} T^*(\chi_k) e^{-2k\pi i x}$$

where $\chi_k = \frac{2k\pi i}{\ln 2}$ *and*

$$T^*(\chi_k) = \sum_{j\in A}\frac{1}{j!}\left(\frac{1}{2}\right)^j \sum_{n\geq 0}\frac{p_n}{n!}\left(\frac{1}{2}\right)^n \Gamma(n+j+\chi_k), \quad \textit{for} \quad k \in \mathbb{Z}\backslash\{0\}. \tag{5}$$

As in Theorem 2, $\delta(x)$ is a periodic function of x with period 1, mean 0 and small amplitude. In Figures 3 and 4 we plot the probabilities that no element occurs exactly once (twice) in compositions of n.

In particular, we see that the probabilities p_n that a composition is A-avoiding, do not converge to a limit as $n \to \infty$, but instead oscillate around the value $1 - \frac{T^*(0)}{\ln 2}$.

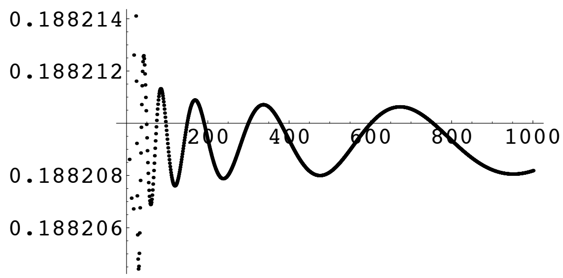


Fig. 3: *Plot of* p_n *for* $b = 1$ *and* $1 \le n \le 1000$.

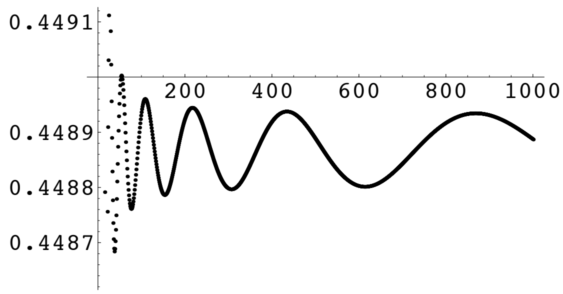


Fig. 4: *Plot of* p_n *for* $b = 2$ *and* $1 \le n \le 1000$.

Acknowledgements

We would like to thank Helmut Prodinger for helpful comments on finding asymptotics from recurrence relations.

References

[1] N. Alon and J. H. Spencer, The Probabilistic Method, Wiley, 2000.

[2] G. E. Andrews, The Theory of Partitions, Addison – Wesley, Reading, MA, 1976.

[3] L. Devroye, A limit theory for random skip lists, *Ann. Appl. Probab.* **2** (1992) 597–609.

[4] P. Flajolet, X. Gourdon and P. Dumas, Mellin Transforms and Asymptotics: Harmonic Sums, *Theoret. Comput. Sci.* **144** (1995) 3–58.

[5] P. Flajolet and G.N. Martin, Probabilistic Counting Algorithms for Data Base Applications, *J. Comput. Syst. Sci.* **31** (1985) 182–209.

[6] P. Hitczenko and A. Knopfmacher, Gap-free compositions and gap-free samples of geometric random variables, *Discrete Math.* **294** (2005) 225–239.

[7] P. Hitczenko and G. Louchard, Distinctness of compositions of an integer: a probabilistic analysis, *Random Struct. Alg.* **19** (2001) 407–437.

[8] P. Hitczenko and C. D. Savage, On the multiplicity of parts in a random composition of a large integer, *SIAM J. Discrete Math.* **18** (2004) 418–435.

[9] P. Jacquet and W. Szpankowski, Analytical de–Poissonization and its applications, *Theoret. Comput. Sci.* **201** (1998) 1–62.

[10] S. Janson and W. Szpankowski, Analysis of the asymmetric leader election algorithm, *Electr. J. Combin.* **4** (1997) #R17.

[11] P. Kirschenhofer and H. Prodinger, On the Analysis of Probabilistic Counting, *Lecture Notes in Mathematics* **1452** (1990) 117–120.

[12] G. Louchard and H. Prodinger, The number of gaps in sequences of geometrically distributed random variables, *Discrete Math.* **308** (2008) 1538–1562.

[13] H. Prodinger, Combinatorics of geometrically distributed random variables: Left-to-right maxima, *Discrete Math.* **153** (1996) 253–270.

[14] W. Pugh, Skip lists: a probabilistic alternative to balanced trees, *Comm. ACM* **33** (1990) 668–676.

[15] W. Szpankowski, Average Case Analysis of Algorithms on Sequences, John Wiley and Sons, New York, 2001.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 336–347

The spectrum of an asymmetric annihilation process

Arvind Ayyer[1] and Volker Strehl[2]

[1] *Institut de Physique Théorique, C. E. A. Saclay, 91191 Gif-sur-Yvette Cedex, France*
[2] *Department of Computer Science, Universität Erlangen-Nürnberg, Haberstrasse 2, D-91058 Erlangen, Germany*

Abstract. In recent work on nonequilibrium statistical physics, a certain Markovian exclusion model called an asymmetric annihilation process was studied by Ayyer and Mallick. In it they gave a precise conjecture for the eigenvalues (along with the multiplicities) of the transition matrix. They further conjectured that to each eigenvalue, there corresponds only one eigenvector. We prove the first of these conjectures by generalizing the original Markov matrix by introducing extra parameters, explicitly calculating its eigenvalues, and showing that the new matrix reduces to the original one by a suitable specialization. In addition, we outline a derivation of the partition function in the generalized model, which also reduces to the one obtained by Ayyer and Mallick in the original model.

Résumé. Dans un travail récent sur la physique statistique hors équilibre, un certain modèle d'exclusion Markovien appelé "processus d'annihilation asymétriques" a été étudié par Ayyer et Mallick. Dans ce document, ils ont donné une conjecture précise pour les valeurs propres (avec les multiplicités) de la matrice stochastique. Ils ont en outre supposé que, pour chaque valeur propre, correspond un seul vecteur propre. Nous prouvons la première de ces conjectures en généralisant la matrice originale de Markov par l'introduction de paramètres supplémentaires, calculant explicitement ses valeurs propres, et en montrant que la nouvelle matrice se réduit à l'originale par une spécialisation appropriée. En outre, nous présentons un calcul de la fonction de partition dans le modèle généralisé, ce qui réduit également à celle obtenue par Ayyer et Mallick dans le modèle original.

Keywords: Reaction diffusion process, non-equilibrium lattice model, transfer matrix Ansatz, partition function, characteristic polynomial, Hadamard transform.

1 Introduction

In the past few years, special stochastic models motivated by nonequilibrium statistical mechanics have motivated several combinatorial problems. The most widely studied problem among these has been the totally asymmetric simple exclusion process (TASEP). The model is defined on a one dimensional lattice of L sites, each site of which either contains a particle or not. Particles in the interior try to jump with rate 1 to the site to the right. The jump succeeds if that site is empty and fails if not. On the boundary, particles enter with rate α on the first site if it is empty and leave from the last site with rate β. This was first solved in 1993 by developing a new technique now called the *matrix product representation* [1].

It was initially studied in a combinatorial setting by Shapiro and Zeilberger in an almost forgotten paper [3] in 1982, but only after the steady state distribution of the model was explicitly presented in

[1], the problem gained widespread attention. One of the reasons for this interest was that the common denominator of the steady state probabilities for a system of size L was C_{L+1}, the $(L+1)$-th Catalan number. One of the first articles to explain this fact combinatorially was the one by Duchi and Schaeffer [4], who enlarged the space of configurations to one in bijection with bicolored Motzkin paths and showed that the steady state distribution was uniform on this space. The analogous construction for the partially asymmetric version of the model (PASEP) has been done in [5].

Further work has been on the relationship of the total and partially asymmetric exclusion processes to different kinds of tableaux by Corteel and Williams [6, 7, 8] (permutation tableaux, staircase tableaux) and by Viennot [9] (Catalan tableaux), to lattice paths [10], and to Askey-Wilson polynomials [7].

Just like the common denominator for the TASEP of size L was the Catalan number C_L (which has many combinatorial interpretations), the common denominator for the asymmetric annihilation process considered in [2] in a system of size L at $\alpha = 1/2, \beta = 1$ is $2^{\binom{L+1}{2}}$ which is the number of domino tilings of an Aztec diamond of size L as well as the number of 2-enumerated $L \times L$ alternating sign matrices. One can therefore hope to enlarge the configuration space as was done for the TASEP [4] to explain this phenomena.

The remainder of this extended abstract is organized as follows: In Sec. 2 we describe the model of the asymmetric annihilation process. In Sec. 3 we present some of the main results obtained by Ayyer and Mallick in [2]. Their work lead to a conjecture about the spectrum of this process. In Sec. 4 we prove this conjecture by appropriately extending the model and viewing it in a different basis obtained by a variant of the Hadamard transform. In the concluding section we outline the derivation of the partition function for the generalized model using the same transformation, an approach very different from the way Ayyer and Mallick obtained the partition function in the original model.

2 The model

Motivated by Glauber dynamics of the Ising model, Ayyer and Mallick [2] considered a non-equilibrium system on a finite lattice with L sites labelled from 1 to L. States of the system are encoded by bitvectors $\boldsymbol{b} = b_1 b_2 \dots b_L$ of length L, where $b_j \in \mathbb{B} = \{0, 1\}$, so that we have a total of 2^L states. These bit vectors may be represented numerically using the binary expansion $(\boldsymbol{b})_2 = b_L + b_{L-1}\, 2^1 + b_{L-2}\, 2^2 + \cdots + b_1\, 2^{L-1}$, which introduces a total order on $\mathbb{B}^L$, so that we shall write $\boldsymbol{b} < \boldsymbol{c}$ iff $(\boldsymbol{b})_2 < (\boldsymbol{c})_2$. All matrices and vectors are indexed w.r.t. this order.

The evolution rules of the system introduced in [2] can now be stated as rewrite rules for bit vectors:

- In the bulk we have *right shift* and *annihilation* given by

$$\begin{array}{lll} \text{right shift} & 10 \to 01 & \text{with rate } 1, \\ \text{annihilation} & 11 \to 00 & \text{with rate } \lambda, \end{array}$$

  and visualized (for $L = 8$) in Fig. 1.

Fig. 1: Right shift $00110\underline{10}1 \to 00110\underline{01}1$ and annihilation $00\underline{11}0101 \to 00\underline{00}0101$

- On the left boundary, particles enter by *left creation* in a way consistent with the bulk dynamics. A particle at site 1 may also be *left annihilated* (due to a virtual particle at site 0). Therefore, the first site evolves as

left creation	$0 \to 1$	with rate α,
left annihilation	$1 \to 0$	with rate $\alpha\lambda$,

as illustrated Fig. 2.

Fig. 2: Left creation $\underline{0}0110101 \to \underline{1}0110011$ and left annihilation $\underline{1}0110101 \to \underline{0}0110101$

- Particles can exit from the last site by *right annihilation* (with a virtual particle at site $L + 1$) according to

$$\text{right annihilation} \quad 1 \to 0 \quad \text{with rate } \beta,$$

as illustrated by Fig. 3.

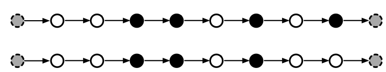


Fig. 3: Right annihilation $0011010\underline{1} \to 1011001\underline{0}$

Note that all transition rules except left creation are monotonically decreasing w.r.t. the natural order of bit vectors. Thus the transition matrix, as discussed in the next section, is not in triangular shape.

Following [2], we will take $\lambda = 1$ as that is the only case for which they derive explicit formulae.

3 Algebraic properties of the model

Is this section we present without proofs the main results as obtained by Ayyer and Mallick in [2]. First recall the general concept:

Definition 1 *A (continuous-time)* transition matrix *or* Markov matrix *or* stochastic matrix *is a square matrix of size equal to the cardinality of the configuration space whose (i, j)-th entry is given by the rate of the transition from configuration j to configuration i, when i is not equal to j. The (i, i)-th entry is then fixed by demanding that the entries in each column sum to zero.*

The Markov chain we defined in the previous section satisfies what [2] call the "transfer matrix Ansatz". The following general definition applies to any family of Markov processes defined by Markov matrices $\{M_L\}$ of increasing sizes (in most physical applications, L is the size of the system).

Definition 2 *A family M_L of Markov processes satisfies the* Transfer Matrix Ansatz *if there exist matrices $T_{L,L+1}$ for all sizes L such that*

$$(\text{TMA}) \quad M_{L+1}T_{L,L+1} = T_{L,L+1}M_L\,.$$

We also impose that this equality is nontrivial in the sense that $M_{L+1}T_{L,L+1} \neq 0$.

The rectangular transfer matrices $T_{L,L+1}$ can be interpreted as *semi-similarity transformations* connecting Markov matrices of different sizes.

The last condition is important because there is always a trivial solution whenever we are guaranteed a unique Perron-Frobenius eigenvector for all transition matrices M_L. If $|v_L\rangle$ is this eigenvector of M_L and $\langle 1_L| = (1, 1, \ldots, 1)$, the matrix $V_{L,L+1} = |v_{L+1}\rangle\langle 1_L|$ satisfies (TMA) since the Markov matrices satisfy the conditions $\langle 1_L|M_L = 0$ and $M_{L+1}|v_{L+1}\rangle = 0$.

The above definition leads immediately to a recursive computation of the steady state vector, which is the zero eigenvector. First we have

$$0 = T_{L,L+1}M_L|v_L\rangle = M_{L+1}T_{L,L+1}|v_L\rangle,$$

which, assuming $T_{L,L+1}|v_L\rangle \neq 0$, and taking into account the uniqueness of the steady state, allows us to define $|v_{L+1}\rangle$ so that

$$T_{L,L+1}|v_L\rangle = |v_{L+1}\rangle.$$

This is very analogous to the matrix product representation of [1] because the steady state probability of any configuration of length $L+1$ is expressed as a linear combination of those of length L. The transfer matrix Ansatz is a stronger requirement than the matrix product representation in the sense that not every system which admits the representation satisfies the Ansatz. For example, the only solution for (TMA) in the case of the TASEP is the trivial one.

For our system introduced above, the Markov matrices M_L are of size 2^L. As mentioned, the entries of these matrices are indexed w.r.t. the naturally ordered basis of binary vectors of length L. For convenience, here are the first three of these matrices:

$$M_1 = \begin{bmatrix} -\alpha & \alpha+\beta \\ \alpha & -\alpha-\beta \end{bmatrix}, \quad M_2 = \begin{bmatrix} \star & \beta & \alpha & 1 \\ 0 & \star & 1 & \alpha \\ \alpha & 0 & \star & \beta \\ 0 & \alpha & 0 & \star \end{bmatrix}, \quad M_3 = \begin{bmatrix} \star & \beta & 0 & 1 & \alpha & 0 & 1 & 0 \\ 0 & \star & 1 & 0 & 0 & \alpha & 0 & 1 \\ 0 & 0 & \star & \beta & 1 & 0 & \alpha & 0 \\ 0 & 0 & 0 & \star & 0 & 1 & 0 & \alpha \\ \alpha & 0 & 0 & 0 & \star & \beta & 0 & 1 \\ 0 & \alpha & 0 & 0 & 0 & \star & 1 & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 & \star & \beta \\ 0 & 0 & 0 & \alpha & 0 & 0 & 0 & \star \end{bmatrix}.$$

As for the diagonal elements $\star$, they have to be set such that the column sums vanish.

We now state without proof some important results on the Markov matrices of the system. These are proved in [2]. We first show that the Markov matrix itself satisfies a recursion of order one.

Theorem 1 *Let σ denote the matrix $\sigma = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and $\mathbb{1}_L$ denote the identity matrix of size 2^L. Then*

$$M_L = \left[\begin{array}{c|c} M_{L-1} - \alpha(\sigma \otimes \mathbb{1}_{L-2}) & \alpha\mathbb{1}_{L-1} + (\sigma \otimes \mathbb{1}_{L-2}) \\ \hline \alpha\mathbb{1}_{L-1} & M_{L-1} - \mathbb{1}_{L-1} - \alpha(\sigma \otimes \mathbb{1}_{L-2}) \end{array}\right],$$

where M_L is written as a 2×2 block matrix with each block made up of matrices of size 2^{L-1}.

The transfer matrices can also be explicitly constructed by a recursion of order one.

Theorem 2 *There exist transfer matrices for the model. If one writes the transfer matrix from size 2^{L-1} to size 2^L by a block decomposition of matrices of size $2^{L-1} \times 2^{L-1}$ as*

$$T_{L-1,L} = \left[\begin{array}{c} T_1^{(L-1)} \\ \hline T_2^{(L-1)} \end{array}\right], \text{ then the matrix } T_{L,L+1} \text{ can be written as } T_{L,L+1} = \left[\begin{array}{c} T_1^{(L)} \\ \hline T_2^{(L)} \end{array}\right], with$$

$$T_1^{(L)} = \left[\begin{array}{c|c} T_1^{(L-1)} + \alpha^{-1} T_2^{(L-1)} & 2T_2^{(L-1)} + \alpha^{-1} T_2^{(L-1)} \\ \hline (\sigma \otimes \mathbb{1}_{L-2}) T_2^{(L-1)} & \alpha^{-1} T_2^{(L-1)} \end{array}\right], \; T_2^{(L)} = \left[\begin{array}{c|c} 2T_2^{(L-1)} & T_2^{(L-1)}(\sigma \otimes \mathbb{1}_{L-2}) \\ \hline 0 & T_2^{(L-1)} \end{array}\right].$$

This, along with the initial condition

$$T_{1,2} = \begin{bmatrix} 1+\beta+\alpha\beta & \alpha+\beta+\alpha\beta \\ \alpha & 1 \\ \alpha+\alpha\beta & \alpha\beta \\ 0 & \alpha \end{bmatrix},$$

determines recursively a family of transfer matrices for the matrices M_L.

We can also use the transfer matrices to calculate properties of the steady state distribution of the Markov process. One quantity of interest is the so called normalization factor or partition function.

Definition 3 *Let the entries of the kernel $|\, v_L \,\rangle$ of M_L be normalized so that their sum is 1 and each entry written in rationally reduced form. Then the* partition function *Z_L for the system of size L is the least common multiple of the denominators of the entries of $|\, v_L \,\rangle$.*

Because of the way the transfer matrix has been constructed Z_L is the sum of the entries in v_L. For example, the system of size one has $|\, v_1 \,\rangle = \begin{bmatrix} \alpha+\beta \\ \alpha \end{bmatrix}$, whence $Z_1 = 2\alpha+\beta$.

Corollary 3 *The partition function of the system of size L is given by*

$$Z_L = 2^{\binom{L-1}{2}} (1+2\alpha)^{L-1} (1+\beta)^{L-1} (2\alpha+\beta).$$

4 Spectrum of the Markov matrices

In this section, we consider the eigenvalues of the Markov matrices M_L of the asymmetric annihilation process. The following result was stated as a conjecture in [2]. This will be a corollary of the main result (Theorem 8) of this article.

Theorem 4 *Let the polynomials $A_L(x)$ and $B_L(x)$ be defined as*

$$A_L(x) = \prod_{k=0}^{\lceil L/2 \rceil} (x+2k)^{\binom{L-1}{2k}}, \qquad B_L(x) = \prod_{k=0}^{\lfloor L/2 \rfloor} (x+2k+1)^{\binom{L-1}{2k+1}}.$$

Then the characteristic polynomial $P_L(x)$ of M_L is given by

$$P_L(x) = A_L(x) A_L(x+2\alpha+\beta) B_L(x+\beta) B_L(x+2\alpha),$$

and successive ratios of characteristic polynomials are given by

$$\frac{P_{L+1}(x)}{P_L(x)} = B_L(x+1)B_L(x+2\alpha+\beta+1)A_L(x+\beta+1)A_L(x+2\alpha+1).$$

This gives only $2L$ distinct eigenvalues out of a possible 2^L. There is therefore the question of diagonalizability of the Markov matrix. Ayyer and Mallick [2] further conjecture the following.

Conjecture 1 *The matrix M_L is maximally degenerate in the sense that it has exactly $2L$ eigenvectors.*

For $L \geq 1$ we regard $\mathbb{B}^L$ as the vector space of bitvectors of length L (over the binary field) . The usual scalar product of vectors $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{B}^L$ will be denoted by $\boldsymbol{b} \cdot \boldsymbol{c}$. We will take the set $V_L = \left\{ |\, \boldsymbol{b}\,\rangle\, ;\, \boldsymbol{b} \in \mathbb{B}^L \right\}$ as the standard basis of a 2^L-dimensional (real or complex) vector space, which we denote by $\mathcal{V}_L$. Indeed, we will consider $\mathcal{V}_L$ as a vector space over an extension over the real or complex field which contains all the variables that we introduce below. To be precise, we take $\mathcal{V}_L$ as a vector space over a field of rational functions which extends the real or complex field.

The following definitions of linear transformations, when considered as matrices, refer to this basis, if not stated otherwise. $\mathcal{V}_L$ is the L-th tensor power of the 2-dimensional space $\mathcal{V}_1$ in an obvious way.

The transformation σ of $\mathcal{V}_1$ is given by the matrix $\sigma = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and this extends naturally to transformations $\sigma^{\boldsymbol{b}}$ of $\mathcal{V}_L$ for $\boldsymbol{b} = b_1 b_2 \dots b_L \in \mathbb{B}^L$:

$$\sigma^{\boldsymbol{b}} = \sigma^{b_1 b_2 \dots b_L} = \sigma^{b_1} \otimes \sigma^{b_2} \otimes \cdots \otimes \sigma^{b_L}.$$

Definition 4 *For a vector $\boldsymbol{\alpha} = (\alpha_{\boldsymbol{b}})_{\boldsymbol{b} \in \mathbb{B}^L}$ of variables we define the transformation $\mathcal{A}_L(\boldsymbol{\alpha})$ of $\mathcal{V}_L$ as*

$$A_L(\boldsymbol{\alpha}) = \sum_{\boldsymbol{b} \in \mathbb{B}^L} \alpha_{\boldsymbol{b}}\, \sigma^{\boldsymbol{b}}.$$

A direct way to define these matrices is $\langle\, \boldsymbol{b} \,|\, A_L \,|\, \boldsymbol{c}\,\rangle = \alpha_{\boldsymbol{b} \oplus \boldsymbol{c}},\ (\boldsymbol{b}, \boldsymbol{c} \in \mathbb{B}^L)$, where $\oplus$ denotes the component wise mod-2-addition (exor) of bit vectors.

For $1 \leq j \leq L$ we define the involutive mappings

$$\phi_j : \mathbb{B}^L \to \mathbb{B}^L : b_1 \dots b_{j-1} b_j\, b_{j+1} \dots b_L \mapsto \phi_j \boldsymbol{b} = b_1 \dots b_{j-1}\, \overline{b_j}\, b_{j+1} \dots b_L$$

by complementing the j-th component, and involutions

$$\psi_j : \mathbb{B}^L \to \mathbb{B}^L : \boldsymbol{b} \mapsto \phi_j \phi_{j+1} \boldsymbol{b}$$

by complementing components indexed j and $j+1$, where ψ_L is the same as ϕ_L.

Definition 5 1. *For $1 \leq j \leq L$ we define the projection operators $\mathcal{P}_{L,j}$ acting on $\mathcal{V}_L$ by*

$$\mathcal{P}_{L,j} = \sum_{\boldsymbol{b} \in \mathbb{B}^L} |\, \boldsymbol{b}\,\rangle\langle\, \boldsymbol{b}\,| - |\, \psi_j^{b_j}(\boldsymbol{b})\,\rangle\langle\, \boldsymbol{b}\,|$$

2. *For a vector $\boldsymbol{b} = (\beta_1, \beta_2, \dots, \beta_L)$ of variables we put $\mathcal{B}_L(\boldsymbol{\beta}) = \sum_{1 \leq j \leq L} \beta_j\, \mathcal{P}_{L,j}$.*
   $B_L(\boldsymbol{\beta})$ denotes the matrix representing $\mathcal{B}_L(\boldsymbol{\beta})$ in the standard basis V_L.

Note that in the sum for $\mathcal{P}_{L,j}$ only summands for which $b_j = 1$, i.e., for which $\psi_j(\boldsymbol{b}) < \boldsymbol{b}$, occur. Indeed: this condition allows only for two situations to contribute:

$$b_j\, b_{j+1} = 10 \;\mapsto\; \overline{b_j}\,\overline{b_{j+1}} = 01 \qquad \text{(right shift)} \tag{1}$$

$$b_j\, b_{j+1} = 11 \;\mapsto\; \overline{b_j}\,\overline{b_{j+1}} = 00 \qquad \text{(annihilation)} \tag{2}$$

Thus these operators ψ_j encode the transitions of our model. Also note that by its very definition $B_L(\boldsymbol{\beta})$ is an upper triangular matrix.

Writing $\boldsymbol{\beta} = (\beta, \gamma, \delta)$ instead of $(\beta_1, \beta_2, \beta_3)$ we have for $L = 3$

$$B_3(\beta,\gamma,\delta) = \begin{bmatrix} 0 & -\delta & 0 & -\gamma & 0 & 0 & -\beta & 0 \\ 0 & \delta & -\gamma & 0 & 0 & 0 & 0 & -\beta \\ 0 & 0 & \gamma & -\delta & -\beta & 0 & 0 & 0 \\ 0 & 0 & 0 & \delta+\gamma & 0 & -\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta & -\delta & 0 & -\gamma \\ 0 & 0 & 0 & 0 & 0 & \delta+\beta & -\gamma & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma+\beta & -\delta \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \delta+\gamma+\beta \end{bmatrix}.$$

Our main concern is now with the transformation given by

$$\mathcal{M}_L(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \mathcal{A}_L(\boldsymbol{\alpha}) - \mathcal{B}_L(\boldsymbol{\beta}).$$

Before we can state the main result we have to introduce some more notation, But before doing so, we note that the corresponding matrix $M_L(\boldsymbol{\alpha}, \boldsymbol{\beta}) = A_L(\boldsymbol{\alpha}) - B_L(\boldsymbol{\beta})$ reduces to the matrix M_L above when properly specialized:

Lemma 5 *We have $A_L(\boldsymbol{\alpha}') - B_L(\boldsymbol{\beta}') = M_L$ for $\boldsymbol{\alpha}' = (\alpha'_{\boldsymbol{b}})_{\boldsymbol{b}\in\mathbb{B}^L}$ and $\boldsymbol{\beta}' = (\beta'_j)_{1\le j\le L}$ given by*

$$\alpha'_{\boldsymbol{b}} = \begin{cases} -\alpha & \textit{if } \boldsymbol{b} = 00\ldots00 \\ \alpha & \textit{if } \boldsymbol{b} = 10\ldots00 \\ 0 & \textit{otherwise} \end{cases} \quad \textit{and} \quad \beta' = \begin{cases} 1 & \textit{if } 1 \le j < L \\ \beta & \textit{if } j = L \end{cases}.$$

We will now consider the transformation $\mathcal{M}_L(\boldsymbol{\alpha}, \boldsymbol{\beta})$ in a different basis of $\mathcal{V}_L$. Let $H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ be the familiar Hadamard matrix and define H_L as its L-th tensor power, the matrix$L$ Hadamard transform of order L:

$$H_L = H^{\otimes L} = \frac{1}{2^{L/2}} \left[(-1)^{\boldsymbol{b}\cdot\boldsymbol{c}} \right]_{\boldsymbol{b},\boldsymbol{c}\in\mathbb{B}^L}.$$

The columns of this matrix, denoted by $|\, w^{\boldsymbol{b}} \,\rangle = H_L |\, \boldsymbol{b} \,\rangle$ for $\boldsymbol{b} \in \mathbb{B}^L$, form an orthonormal basis $W_L = \left\{ H_L |\, \boldsymbol{b} \,\rangle \,;\, \boldsymbol{b} \in \mathbb{B}^L \right\}$ of $\mathcal{V}_L$. The following assertion is easily checked:

Lemma 6 *The (pairwise commuting) transformations $\sigma^{\boldsymbol{c}}$ ($\boldsymbol{c} \in \mathbb{B}^L$) diagonalize in the W_L-basis. More precisely:*

$$\sigma^{\boldsymbol{c}} |\, w^{\boldsymbol{b}} \,\rangle = (-1)^{\boldsymbol{b}\cdot\boldsymbol{c}} |\, w^{\boldsymbol{b}} \,\rangle \quad (\boldsymbol{b}, \boldsymbol{c} \in \mathbb{B}^L).$$

Thus also the transformation $\mathcal{A}_L$ diagonalizes in the W_L-basis and its eigenvalues are given by

$$(H_L \cdot A_L \cdot H_L)|\, w^{\boldsymbol{b}}\,\rangle = \lambda_{\boldsymbol{b}}|\, w^{\boldsymbol{b}}\,\rangle$$

where $\lambda_{\boldsymbol{b}} = \sum_{c\in\mathbb{B}^L} \alpha_{\boldsymbol{c}}\,(-1)^{\boldsymbol{b}\cdot\boldsymbol{c}} = \sum_{c\in\mathbb{B}^L} \alpha_{\boldsymbol{c}}\,\langle\, \boldsymbol{b}\,|\, H\,|\,\boldsymbol{c}\,\rangle$.

The crucial observation is now the following: even though the transformation $\mathcal{A}_L$ diagonalizes in the W_L-basis, the transformation $\mathcal{B}_L$ doesn't, it is not even triangular in this basis. But it turns out that a slight modification of the W_L-basis will be suitable for at the same time diagonalizing $\mathcal{A}_L$ and bringing the $\mathcal{B}_L$ in (lower) triangular form. For that purpose we introduce the invertible linear transformation

$$\Delta : \mathbb{B}^L \to \mathbb{B}^L : \boldsymbol{b} = b_1 b_2 \dots b_L \mapsto \boldsymbol{b}^{\Delta} = \left[\sum_{1\le i\le L-j+1} b_i\right]_{1\le j\le L}$$

where the sum has to be taken in the binary field. As an example ($L = 3$):

$\boldsymbol{b}$	000	001	010	011	100	101	110	111
$\boldsymbol{b}^{\Delta}$	000	100	110	010	111	011	001	101

The basis $\widetilde{W}_L = \{|\, w^{\boldsymbol{b}^{\Delta}}\,\rangle\}$ is nothing but a rearrangement of the W_L-basis, hence the transformation $\mathcal{A}_L$ diagonalizes in this basis as well (with the corresponding eigenvalues). We will write $\widetilde{H}_L$ for the rearrangement of the Hadamard matrix in this new ordering of the elements of $\mathcal{B}^L$. The clue is now contained in the following proposition:

Proposition 7

$$\widetilde{H}_L \cdot B_L(\boldsymbol{\beta}) \cdot \widetilde{H}_L = B_L^{\mathtt{t}}(\boldsymbol{\beta}^{\mathtt{rev}})$$

where $\boldsymbol{\beta}^{\mathtt{rev}} = (\beta_L, \beta_{L-1}, \dots, \beta_1)$ *is the reverse of* $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_L)$, $\mathtt{t}$ *denoting transposition.*

We illustrate this proposition in the case $L = 3$ by displaying matrices $2^{3/2}\widetilde{H}_3$ (left) and $\widetilde{H}_3 \cdot B_3(\beta,\gamma,\delta)\cdot \widetilde{H}_3 = B_3(\delta,\gamma,\beta)^{\mathtt{t}}$ (right). Note that $\widetilde{H}_L$ is symmetric because Δ (as a matrix) is symmetric.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\beta & \beta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\gamma & \gamma & 0 & 0 & 0 & 0 & 0 \\ -\gamma & 0 & -\beta & \beta+\gamma & 0 & 0 & 0 & 0 \\ 0 & 0 & -\delta & 0 & \delta & 0 & 0 & 0 \\ 0 & 0 & 0 & -\delta & -\beta & \beta+\delta & 0 & 0 \\ -\delta & 0 & 0 & 0 & 0 & -\gamma & \gamma+\delta & 0 \\ 0 & -\delta & 0 & 0 & -\gamma & 0 & -\beta & \beta+\gamma+\delta \end{bmatrix}.$$

The proof of Proposition 7 will be given below. It leads to the main result by looking at the matrix representation of $\mathcal{M}_L(\boldsymbol{\alpha},\boldsymbol{\beta}) = \mathcal{A}_L - \mathcal{B}_L$ in the $\widetilde{W}_L$-basis, where it takes lower triangular form. Hence the eigenvalues, which are 2^L pairwise distinct linear polynomials in the α- and β-variables, can be read directly from the main diagonal. In contrast to Theorem 4, the ex-conjecture, all eigenvalues are simple.

Theorem 8

$$\det M_L(\boldsymbol{\alpha},\boldsymbol{\beta}) = \det\left[A_L(\boldsymbol{\alpha}) - B_L(\boldsymbol{\beta})\right] = \prod_{\boldsymbol{b}\in\mathbb{B}^L} \left(\lambda_{\boldsymbol{b}^\Delta} - \boldsymbol{\beta}^{\mathtt{rev}}\cdot\boldsymbol{b}\right)$$

Illustration of the Theorem for $L=3$ recalling $\boldsymbol{\beta} = (\beta,\gamma,\delta) = (\beta_1,\beta_2,\beta_3)$:

$\boldsymbol{b}$	$\boldsymbol{b}^\Delta$	λ_{b^Δ}	$(\delta,\gamma,\beta)\cdot\boldsymbol{b}$
000	000	$[++++++++]\cdot\boldsymbol{\alpha}$	0
001	100	$[++++----]\cdot\boldsymbol{\alpha}$	β
010	110	$[++----++]\cdot\boldsymbol{\alpha}$	γ
011	010	$[++--++--]\cdot\boldsymbol{\alpha}$	$\beta+\gamma$
100	111	$[+--+-++-]\cdot\boldsymbol{\alpha}$	δ
101	011	$[+--++--+]\cdot\boldsymbol{\alpha}$	$\beta+\delta$
110	001	$[+-+-+-+-]\cdot\boldsymbol{\alpha}$	$\gamma+\delta$
111	101	$[+-+--+-+]\cdot\boldsymbol{\alpha}$	$\beta+\gamma+\delta$

So, as an example, the line for $\boldsymbol{b}=101$ contributes the factor $\alpha_{000}-\alpha_{001}-\alpha_{010}-\alpha_{011}+\alpha_{100}-\alpha_{101}-\alpha_{110}-\alpha_{111}-\beta-\delta$ to the product.

To prepare for the proof of Proposition 7 we state without proof simple relations between the transformations ψ_j, ϕ_{L-j+1} and Δ:

Lemma 9 *For $\boldsymbol{b},\boldsymbol{c}\in\mathbb{B}^L$ and $1\le j\le L$ we have*

1. $(\psi_j\boldsymbol{b})^\Delta = \phi_{L-j+1}(\boldsymbol{b}^\Delta)$
2. $\boldsymbol{b}^\Delta\cdot\psi_j\boldsymbol{c} = \boldsymbol{b}^\Delta\cdot\boldsymbol{c} + b_{L-j+1}$

Fact 2. is a consequence of fact 1.

Proof of Proposition 7. The actions of the transformations $\mathcal{P}_{L,j}$, seen in the $\widetilde{W}_L$-basis, are given by:

$$\mathcal{P}_{L,j} : |\, w^{\boldsymbol{b}^\Delta}\,\rangle \mapsto \begin{cases} -|\, w^{(\psi_{L-j+1}\boldsymbol{b})^\Delta}\,\rangle & \text{if } b_{L-j+1}=0,\\ |\, w^{\boldsymbol{b}^\Delta}\,\rangle & \text{if } b_{L-j+1}=1.\end{cases}$$

To see this, we compute

$$\begin{aligned}\mathcal{P}_{L,j}|\, w^{\boldsymbol{b}^\Delta}\,\rangle &= \sum_{\boldsymbol{c}\in\mathbb{B}^L}\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\boldsymbol{c}\,\rangle \mathcal{P}_{L,j}|\,\boldsymbol{c}\,\rangle = \sum_{\boldsymbol{c}>\psi_j\boldsymbol{c}}\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\boldsymbol{c}\,\rangle\,(|\,\boldsymbol{c}\,\rangle - |\,\psi_j\boldsymbol{c}\,\rangle)\\ &= \sum_{\boldsymbol{c}>\psi_j\boldsymbol{c}}\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\boldsymbol{c}\,\rangle\,|\,\boldsymbol{c}\,\rangle - \sum_{\boldsymbol{c}<\psi_j\boldsymbol{c}}\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\psi_j\boldsymbol{c}\,\rangle|\,\boldsymbol{c}\,\rangle,\end{aligned}$$

using the involutive nature of ψ_j for the second sum. Now, using 2. from Lemma 9,

$$\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\psi_j\boldsymbol{c}\,\rangle = (-1)^{\boldsymbol{b}^\Delta\cdot\psi_j\boldsymbol{c}} = (-1)^{\boldsymbol{b}^\Delta\cdot\boldsymbol{c}+b_{L-j+1}} = (-1)^{b_{L-j+1}}\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\boldsymbol{c}\,\rangle$$

and thus

$$\mathcal{P}_{L,j}|\, w^{\boldsymbol{b}^\Delta}\,\rangle = \sum_{\boldsymbol{c}>\psi_j\boldsymbol{c}}\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\boldsymbol{c}\,\rangle\,|\,\boldsymbol{c}\,\rangle - (-1)^{b_{L-j+1}}\sum_{\boldsymbol{c}<\psi_j\boldsymbol{c}}\langle\, \boldsymbol{b}^\Delta \,|\, H\,|\,\boldsymbol{c}\,\rangle|\,\boldsymbol{c}\,\rangle.$$

The conclusion in the case $b_{L-j+1} = 1$ is now obvious.

As for the case $b_{L-j+1} = 1$, we see, using item 1. from Lemma 9, that

$$
\begin{aligned}
|\, w^{(\psi_{L-j+1}\boldsymbol{b})^\Delta} \rangle &= \sum_{\boldsymbol{c}\in\mathbb{B}^L} \langle\, (\psi_{L-j+1}\boldsymbol{b})^\Delta \,|\, H \,|\, \boldsymbol{c} \,\rangle \,|\, \boldsymbol{c} \,\rangle = \sum_{\boldsymbol{c}\in\mathbb{B}^L} (-1)^{(\psi_{L-j+1}\boldsymbol{b})^\Delta\cdot \boldsymbol{c}} \,|\, \boldsymbol{c} \,\rangle \\
&= \sum_{\boldsymbol{c}\in\mathbb{B}^L} (-1)^{\phi_j(\boldsymbol{b}^\Delta)\cdot \boldsymbol{c}} \,|\, \boldsymbol{c} \,\rangle = \sum_{\boldsymbol{c}\in\mathbb{B}^L} (-1)^{\boldsymbol{b}^\Delta\cdot \boldsymbol{c} + c_j} \,|\, \boldsymbol{c} \,\rangle \\
&= -\sum_{\boldsymbol{c}:c_j=1} \langle\, \boldsymbol{b}^\Delta \,|\, H \,|\, \boldsymbol{c} \,\rangle \,|\, \boldsymbol{c} \,\rangle + \sum_{\boldsymbol{c}:c_j=0} \langle\, \boldsymbol{b}^\Delta \,|\, H \,|\, \boldsymbol{c} \,\rangle \,|\, \boldsymbol{c} \,\rangle \;= -\mathcal{P}_{L,j} |\, w^{\boldsymbol{b}^\Delta} \,\rangle. \qquad \square
\end{aligned}
$$

Corollary 10 *If we consider the special case where $\alpha_{\boldsymbol{b}} = 0$ for all $\boldsymbol{b} \in \mathbb{B}^L$, except $\alpha_{00\ldots 0} = \alpha_0$ and $\alpha_{10\ldots 00} = \alpha_1$, and where $\beta_1 = \ldots = \beta_{L-1} = 1$ and $\beta_L = \beta$, then the determinant of the Theorem simplifies to the product $\Pi_1 \cdot \Pi_2 \cdot \Pi_3 \cdot \Pi_4$ of the following four terms:*

$$
\Pi_1 = \prod_{0\le 2k<L} (\alpha_0 + \alpha_1 - 2k)^{\binom{L-1}{2k}} \qquad \Pi_2 = \prod_{0\le 2k-1<L} (\alpha_0 + \alpha_1 - \beta - 2k + 1)^{\binom{L-1}{2k-1}}
$$

$$
\Pi_3 = \prod_{0\le 2k-1<L} (\alpha_0 - \alpha_1 - 2k + 1)^{\binom{L-1}{2k-1}} \qquad \Pi_4 = \prod_{0\le 2k<L} (\alpha_0 - \alpha_1 - \beta - 2k)^{\binom{L-1}{2k}}
$$

For the proof note that each $\boldsymbol{b} \in \mathbb{B}^L$ we get as the contribution from $\mathcal{A}_L$

$$
\begin{aligned}
\lambda_{\boldsymbol{b}^\Delta} \cdot \boldsymbol{\alpha} &= \sum_{\boldsymbol{c}\in\mathbb{B}^L} \alpha_{\boldsymbol{c}} \langle\, \boldsymbol{b}^\Delta \,|\, H \,|\, \boldsymbol{c} \,\rangle \\
&= \alpha_0 \,\langle \boldsymbol{b}^\Delta \,|\, H \,|\, 00\ldots 00\rangle + \alpha_1 \,\langle \boldsymbol{b}^\Delta \,|\, H \,|\, 10\ldots 00\rangle \\
&= \alpha_0 (-1)^{\boldsymbol{b}\cdot\Delta\cdot 00\ldots 00} + \alpha_1 (-1)^{\boldsymbol{b}\cdot\Delta\cdot 10\ldots 00} = \alpha_0 + (-1)^{\|\boldsymbol{b}\|}\, \alpha_1
\end{aligned}
$$

because $\Delta \cdot 00\ldots 00 = 11\ldots 11$ and then $\boldsymbol{b} \cdot 11\ldots 11 \equiv \|\boldsymbol{b}\| \bmod 2$, where $\|\boldsymbol{b}\|$ denotes the Hamming weight of $\boldsymbol{b}$ and where we have used the fact that Δ is a symmetric matrix. Thus the 2^L eigenvalues are

$$
\alpha_0 + (-1)^{\|\boldsymbol{b}\|}\, \alpha_1 - \boldsymbol{\beta}^{\mathtt{rev}} \cdot \boldsymbol{b} \quad (\boldsymbol{b} \in \mathbb{B}^L).
$$

Now there are four cases to consider:

1. $\|\boldsymbol{b}\|$ is even and $b_1 = 0$: this gives eigenvalues $\alpha_0 + \alpha_1 - \|b_2 b_3 \ldots b_L\|$ and since b_1 does not contribute to $\|\boldsymbol{b}\|$ the vector $b_2 b_3 \ldots b_L$ must have even weight $2k$. There are $\binom{L-1}{2k}$ possibilities which account for Π_1.

2. $\|\boldsymbol{b}\|$ is even and $b_1 = 1$: this gives eigenvalues $\alpha_0 + \alpha_1 - \beta - \|b_2 b_3 \ldots b_L\|$ and since b_1 does contribute to $\|\boldsymbol{b}\|$ the vector $b_2 b_3 \ldots b_L$ must have odd weight $2k-1$. There are $\binom{L-1}{2k-1}$ possibilities which account for Π_2.

3. $\|\boldsymbol{b}\|$ is odd and $b_1 = 0$: this gives eigenvalues $\alpha_0 - \alpha_1 - \|b_2 b_3 \ldots b_L\|$ and since b_1 does not contribute to $\|\boldsymbol{b}\|$ the vector $b_2 b_3 \ldots b_L$ must have odd weight $2k-1$. There are $\binom{L-1}{2k-1}$ possibilities which account for Π_3.

4. $\|\boldsymbol{b}\|$ is odd and $b_1 = 1$: this gives eigenvalues $\alpha_0 - \alpha_1 - \beta - \|b_2 b_3 \dots b_L\|$ and since b_1 does contribute to $\|\boldsymbol{b}\|$ the vector $b_2 b_3 \dots b_L$ must have even weight $2k$. There are $\binom{L-1}{2k}$ possibilities which account for Π_4.

Corollary 11 *Setting now $\alpha_0 = -\alpha$ and $\alpha_1 = \alpha$, i.e., specializing as in Lemma 5, gives for* $\det M_L$ *a product* $\Pi_1' \cdot \Pi_2' \cdot \Pi_3' \cdot \Pi_4'$ *of the following four terms:*

$$\Pi_1' = \prod_{0 \le 2k < L} (-2k)^{\binom{L-1}{2k}} \qquad \Pi_2 = \prod_{0 \le 2k-1 < L} (-\beta - 2k + 1)^{\binom{L-1}{2k-1}}$$

$$\Pi_3 = \prod_{0 \le 2k-1 < L} (-2\alpha - 2k + 1)^{\binom{L-1}{2k-1}} \qquad \Pi_4 = \prod_{0 \le 2k < L} (-2\alpha_1 - \beta - 2k)^{\binom{L-1}{2k}}$$

which are precisely the $2\,L$ distinct eigenvalues of the original Conjecture.

5 Concluding remarks

We have been able to solve Ayyer and Mallick's conjecture about the eigenvalues of the asymmetric annihilation process by embedding it into a more general model and using an orthogonal transform which triagonalizes the transition matrix. In sharp contrast to the original problem, the general situation with parameters $\alpha_{\boldsymbol{b}}$ and b_j (which may be given a "physical" interpretation using $\langle\, \boldsymbol{b} \,|\, A_L \,|\, \boldsymbol{c} \,\rangle = \alpha_{\boldsymbol{b} \oplus \boldsymbol{c}}$ and (1),(2)) is easier to handle because it is not degenerate: all "symbolic" eigenvalues are simple. Our proof does not seem to explain the maximum amount of degeneracy, as stated in Conjecture 1.

On the other hand, we mention that the result of Corollary 3 about the partition function can be extended to the more general model. Again, in contrast to the inductive approach of Ayyer and Mallick in [2], as outlined in Sec. 3, we can solve this problem directly by transforming it orthogonally into the basis where it shows its triangular structure.

We start by remarking that the columns sums of the extended model are constant $\sum \alpha_c$, though not zero. This implies that $\langle\, 1_L \,|$ is the unique left eigenvector with eigenvalue $\overline{\alpha} = \sum \alpha_c$ of $M_L(\alpha, \beta)$. The right eigenvector $|\, \boldsymbol{x} \,\rangle$ with the same eigenvalue corresponds to the steady state distribution of the original problem. Then $|\, \boldsymbol{y} \,\rangle = \widetilde{H}_L |\, \boldsymbol{x} \,\rangle$ satisfies $\widetilde{M}_L(\boldsymbol{\alpha}, \boldsymbol{\beta}) |\, \boldsymbol{y} \,\rangle = \overline{\alpha} |\, \boldsymbol{y} \,\rangle$, where $\widetilde{M}_L = \widetilde{H}_L \cdot M_L(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cdot \widetilde{H}_L$ is the matrix seen in the $\widetilde{W}_L$-basis. This triangular system for $\boldsymbol{y}$ is written explicitly as

$$(\lambda^*_{b^\Delta} + \boldsymbol{\beta}^{\mathtt{rev}} \cdot \boldsymbol{b})\; y_{\boldsymbol{b}} = \sum_{j \,:\, b_j = 1} \beta_{L-j+1} y_{\psi_j \boldsymbol{b}} \quad (\boldsymbol{b} \in \mathbb{B}^L)$$

where

$$\lambda^*_{\boldsymbol{b}} = \overline{\alpha} - \lambda_{\boldsymbol{b}} = 2 \sum_{\boldsymbol{c} \,:\, \boldsymbol{b} \cdot \boldsymbol{c} = 1} \alpha_{\boldsymbol{c}} \quad (\boldsymbol{b} \in \mathbb{B}^L).$$

Note that the sum on the right only contains terms $y_{\boldsymbol{c}}$ where $\boldsymbol{c} = \psi_j \boldsymbol{b} < \boldsymbol{b}$. For $\boldsymbol{b} = 00 \dots 0$ the equation is void, so we may put $y_{00\dots 0} = 1$. Since the polynomials $\lambda^*_{b^\Delta} + \boldsymbol{\beta}^{\mathtt{rev}} \cdot \boldsymbol{b}$ are mutually coprime, this shows by induction that the denominator of the rational normal form of $y_{\boldsymbol{b}}$ is the product of all polynomials $\lambda^*_{c^\Delta} + \boldsymbol{\beta}^{\mathtt{rev}} \cdot \boldsymbol{c}$, where $\boldsymbol{c}$ runs over the binary vectors that can be obtained from $\boldsymbol{b}$ by successive application of decreasing ψ_j-transformations. Consequently, the product of linear polynomials

$$Z(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{\boldsymbol{0} \ne \boldsymbol{b} \in \mathbb{B}^L} (\lambda^*_{b^\Delta} + \boldsymbol{\beta}^{\mathtt{rev}} \cdot \boldsymbol{b})$$

is the least common multiple of the denominators of the $y_{\boldsymbol{b}}$. This property is invariant under the Hadamard transform, so it applies also to the coefficients of $|\,\boldsymbol{x}\,\rangle = \widetilde{H}_L|\,\boldsymbol{y}\,\rangle$. But

$$\langle \mathbf{1}_L|\boldsymbol{x}\rangle = \langle \mathbf{1}_L|\widetilde{H}_L|\boldsymbol{y}\rangle = 2^{L/2}\,\langle 100\ldots00|\boldsymbol{y}\rangle = 2^{L/2}\,y_{00\ldots0} = 2^{L/2},$$

so $2^{-L/2}\,|\,\boldsymbol{x}\,\rangle$ is already normalized and can be seen as the "symbolic" stationary distribution in the generalized model. What we have shown is:

Theorem 12 *$Z(\boldsymbol{\alpha},\boldsymbol{\beta})$ is the partition function related to $M_L(\boldsymbol{\alpha},\boldsymbol{\beta})$.*

We conclude by remarking that the specialization as in Lemma 5 and Corollary 11 brings us back to Corollary 3. This is not completely obvious, since the expression in Corollary 3 has only $\binom{L+1}{2}$ factors, whereas in Theorem 12 there are 2^L-1 factors. What happens is that upon specialization the requirements for least common multiples and greatest common divisors change. Taking this into account one finds that from the general expression for $Z(\boldsymbol{\alpha},\boldsymbol{\beta})$ only the $\binom{L+1}{2}$ terms where $\boldsymbol{b}\in\mathbb{B}^L$ with $\|\boldsymbol{b}\|=1$ or 2 contribute – and this is precisely the statement of Corollary 3.

References

[1] B. Derrida, M. R. Evans, V. Hakim, V. Pasquier, Exact solution of a 1D asymmetric exclusion model using a matrix formulation, *J. Phys. A: Math. Gen.* **26** (1993), 1493–1517.

[2] A. Ayyer and K. Mallick, Exact results for an asymmetric annihilation process with open boundaries, *J. Phys. A: Math. Gen.* **43** (2010) 045033, 22pp.

[3] L. Shapiro and D. Zeilberger, A Markov chain occurring in Enzyme Kinetics, *J. Math. Biology* **15** (1982) 351–357.

[4] E. Duchi and G. Schaeffer, A combinatorial approach to jumping particles: The parallel TASEP, *Random Structures and Algorithms*, **33** no. 4 (2008),434–451.

[5] Richard Brak, S. Corteel, J. Essam R. Parviainen and A. Rechnitzer, A Combinatorial Derivation of the PASEP Stationary State, *Electron. J. Combin*. **13** no. 1 (2006), 108, 23 pp.

[6] S. Corteel and L. Williams, Tableaux combinatorics for the asymmetric exclusion process, *Advances in Applied Mathematics*, **39** no. 3 (2007), 293–310.

[7] S. Corteel and L. Williams, A Markov chain on permutations which projects to the PASEP, *International Mathematics Research Notices*, **2007** (2007) rnm055, 27 pp.

[8] S. Corteel and L. Williams, Staircase tableaux, the asymmetric exclusion process and Askey-Wilson polynomials, to appear in *Proc. Nat. Acad. Sci. USA*, `arXiv:0910.1858`.

[9] G. X. Viennot, Canopy of binary trees, Catalan tableaux and the asymmetric exclusion process, *Proceedings of FPSAC 2007*, Tianjin, China, preprint, `arXiv:0905.3081`.

[10] S. Corteel, M. Josuat-Verges, T. Prellberg and M. Rubey, Matrix Ansatz, lattice paths and rook placements, *Proceedings of FPSAC2009*, DMTCS proc. AK, **104** (2009), 313–324, `arXiv:0811.4606`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 348–359

Weakly directed self-avoiding walks

Axel Bacher[1] and Mireille Bousquet-Mélou[2]

[1] *LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence, France*
[2] *CNRS, LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence, France*

Abstract. We define a new family of self-avoiding walks (SAW) on the square lattice, called *weakly directed walks*. These walks have a simple characterization in terms of the irreducible bridges that compose them. We determine their generating function. This series has a complex singularity structure and in particular, is not D-finite. The growth constant is approximately 2.54 and is thus larger than that of all natural families of SAW enumerated so far (but smaller than that of general SAW, which is about 2.64). We also prove that the end-to-end distance of weakly directed walks grows linearly. Finally, we study a diagonal variant of this model.

Résumé. Nous définissons une nouvelle famille de chemins auto-évitants (CAE) sur le réseau carré, appelés *chemins faiblement dirigés*. Ces chemins ont une caractérisation simple en termes des ponts irréductibles qui les composent. Nous déterminons leur série génératrice. Cette série a une structure de singularités complexe et n'est en particulier pas D-finie. La constante de croissance est environ 2,54, ce qui est supérieur à toutes les familles naturelles de SAW étudiées jusqu'à présent, mais inférieur aux CAE généraux (dont la constante est environ 2,64). Nous prouvons également que la distance moyenne entre les extrémités du chemin croît linéairement. Enfin, nous étudions une variante diagonale du modèle.

Keywords: Enumeration – Self-avoiding walks

1 Introduction

A lattice walk is *self-avoiding* if it never visits twice the same vertex (Fig. 1). Self-avoiding walks (SAW) have attracted interest for decades, first in statistical physics, where they are considered as polymer models, and then in combinatorics and in probability theory [20]. However, their properties remain poorly understood in low dimension, despite the existence of remarkable conjectures. See [20] for dimension 5 and above.

On two-dimensional lattices, it is strongly believed that the number c_n of n-step SAW and the average end-to-end distance D_n of these walks satisfy

$$c_n \sim \alpha\mu^n n^\gamma \quad \text{and} \quad D_n \sim \kappa n^\nu \tag{1}$$

where $\gamma = 11/32$ and $\nu = 3/4$. Several independent, but so far not completely rigorous methods predict these values, like numerical studies [12, 24], comparisons with other models [6, 21], probabilistic arguments involving SLE processes [19], enumeration of SAW on random planar lattices [10]... The *growth constant* (or *connective constant*) μ is lattice-dependent, and believed to be $\sqrt{2+\sqrt{2}}$ for the honeycomb lattice, and another bi-quadratic number (approximately 2.64) for the square lattice [16].

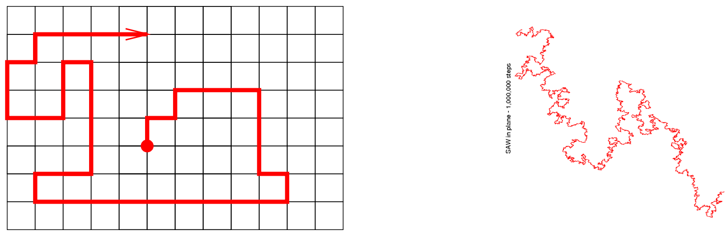


Fig. 1: A self-avoiding walk on the square lattice, and a (quasi-)random SAW of length 1,000,000, constructed by Kennedy using a pivot algorithm [17].

Given the difficulty of the problem, the study of *restricted* classes of SAW is natural, and probably as old as the interest in SAW itself. The rule of this game is to design new classes of SAW that have both:

– a natural description (to be conceptually pleasant),

– some structure (so that the walks can be counted, and their asymptotic properties determined).

The two simplest classes of SAW on the square lattice probably consist of *directed* and *partially directed* walks: a walk is directed if it involves at most two types of steps (for instance North and East), and partially directed if it involves at most three types of steps. (Partially directed walks play a prominent role in the definition of our *weakly directed* walks.) Among other solved classes, let us cite spiral SAW [22, 13] and prudent walks [3, 8, 7]. Each time such a new class is solved, one compares its properties to (1): have we reached with this class a large growth constant? Is the end-to-end distance of the walks sub-linear?

At the moment, the largest growth constant (about 2.48) is obtained with prudent SAW. However, this is beaten by certain classes whose description involves a (small) integer k, like SAW confined to a strip of height k [1, 26], or SAW consisting of *irreducible bridges* of length at most k [15, 18]. The structure of these walks is rather poor, which makes them little attractive from a combinatorial viewpoint. In the former case, they are described by a transfer matrix (the size of which increases exponentially with the height of the strip); in the latter case, the structure is even simpler, since these walks are just arbitrary *sequences* of irreducible bridges of bounded length. In both cases, improvements on the growth constant much rely on progresses in the computer power. Regarding asymptotic properties, almost all solved classes of SAW exhibit a linear end-to-end distance, with the exception of spiral walks. But there are very few such walks [13], as their growth constant is 1.

With the *weakly directed walks* of this paper, we reach a growth constant of about 2.54. These walks are defined in the next section. Their generating function is given in Section 5, after some preliminary results on partially directed *bridges* (Sections 3 and 4). This series turns out to be much more complicated that the generating functions of directed and partially directed walks, which are rational: we prove that it has a natural boundary in the complex plane, and in particular is not D-finite (that is, it does not satisfy any linear differential equation with polynomial coefficients). However, we are able to derive from this series certain asymptotic properties of weakly directed walks, like their growth constant and average end-to-end distance (which we find, unfortunately, to grow linearly with the length). Finally, we perform in Section 6 a similar study for a diagonal variant of weakly directed walks.

Due to space constraints, most proofs are only sketched or even omitted in this abstract. Details will appear in the complete version of the paper.

2 Weakly directed walks: definition

Let us denote by N, E, S and W the four square lattice steps. All walks in this paper are self-avoiding, so that this precision will often be omitted. For any subset $\mathcal{S}$ of $\{\mathsf{N},\mathsf{E},\mathsf{S},\mathsf{W}\}$, we say that a (self-avoiding) walk is an $\mathcal{S}$-*walk* if all its steps lie in $\mathcal{S}$. We say that a SAW is *directed* if it involves at most two types of steps, and *partially directed* if it involves at most three types of steps. The definition of weakly directed walks stems for the following simple observations:

(i) between two visits to any given horizontal line, a NE-walk only takes E steps,

(ii) between two visits to any given horizontal line, a NEW-walk only takes E and W steps.

Conversely, a walk satisfies (i) if and only if it is either a NE-walk or, symmetrically, a SE-walk. Similarly, a walk satisfies (ii) if and only if it is either a NEW-walk or, symmetrically, a SEW-walk. Conditions (i) and (ii) thus respectively characterize (up to symmetry) NE-walks and NEW-walks.

Definition 1 *A walk is* weakly directed *if, between two visits to any given horizontal line, the walk is partially directed (that is, avoids at least one of the steps* N, E, S, W*).*

Examples are shown in Fig. 2.

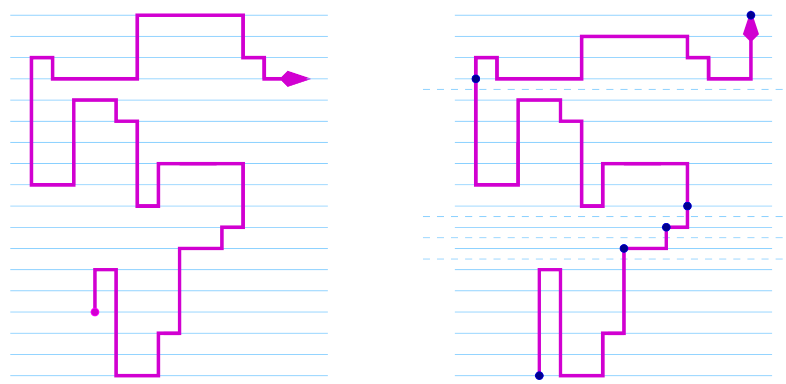


Fig. 2: Two weakly directed walks. The second one is a bridge, formed of 5 irreducible bridges. Observe that these irreducible bridges are partially directed.

We will primarily focus on the enumeration of weakly directed *bridges*. As we shall see, this does not affect the growth constant. A self-avoiding walk starting at v_0 and ending at v_n is a *bridge* if all its vertices $v \neq v_n$ satisfy $h(v_0) \le h(v) < h(v_n)$, where $h(v)$, the *height* of v, is its ordinate. Concatenating two bridges always gives a bridge. Conversely, every bridge can be uniquely factored into a sequence of *irreducible* bridges (bridges that cannot be written as the product of two non-empty bridges). This factorization is obtained by cutting the walk above each horizontal line of height $n + 1/2$ (with $n \in \mathbb{Z}$) that the walk intersects only once (Fig. 2, right). It is known that the growth constant of bridges is the same as for general self-avoiding walks [20]. Generally speaking, the fact that bridges can be freely concatenated makes them useful objects in the study of self-avoiding walks [14, 15, 18, 19, 20].

The following result shows that the enumeration of weakly directed bridges boils down to the enumeration of (irreducible) partially directed bridges.

Proposition 2 *A bridge is weakly directed if and only if each of its irreducible bridges is partially directed (that is, avoids at least one of the steps* N, E, S, W*).*

We discuss in Section 6 a variant of weakly directed walks, where we constrain the walk to be partially directed between two visits to the same *diagonal* line (Fig. 3). The notion of bridges is adapted accordingly, by defining the *height* of a vertex as the sum of its coordinates. We will refer to this model as the diagonal model, and to the original one as the horizontal model. There is, however, no simple counterpart of Proposition 2: a (diagonal) bridge whose irreducible bridges are partially directed is always weakly directed, but the converse is not true, as can be seen in Fig. 3. Thus bridges with partially directed irreducible bridges form a proper subclass of weakly directed bridges. We will enumerate this subclass, and study its asymptotic properties.

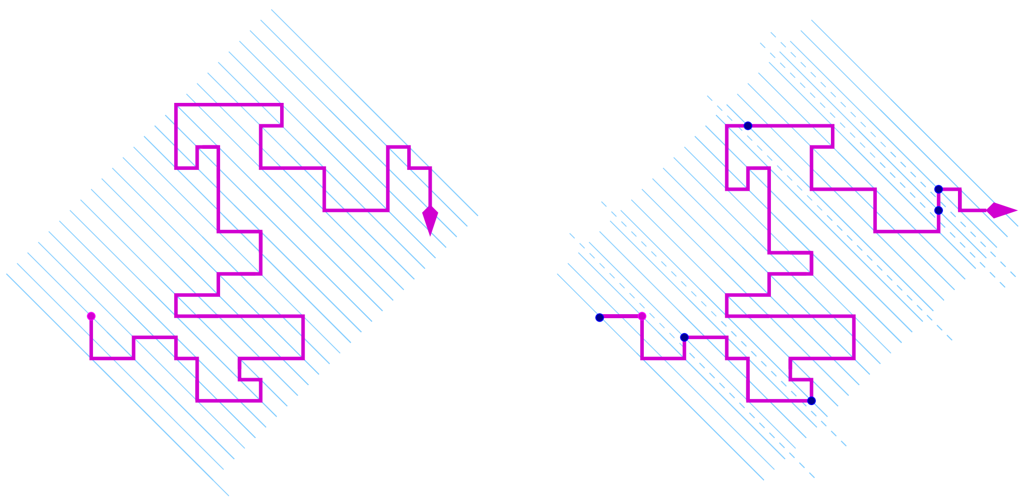


Fig. 3: Two weakly directed walks in the diagonal model. The second one is a bridge, factored into 6 irreducible bridges. Observe that the third irreducible bridge is *not* partially directed.

3 Partially directed bridges: a step-by-step approach

Let us equip the square lattice $\mathbb{Z}^2$ with its standard coordinate system. With each model (horizontal or diagonal) is associated a notion of *height*: the height of a vertex v, denoted $h(v)$, is its ordinate in the horizontal model, while in the diagonal model, it is the sum of its coordinates. Recall that a walk, starting at v_0 and ending at v_n, is a bridge if all its vertices $v \neq v_n$ satisfy $h(v_0) \leq h(v) < h(v_n)$. If the weaker inequality $h(v_0) \leq h(v) \leq h(v_n)$ holds for all v, we say the walk is a *pseudo-bridge*. Note that nonempty bridges are obtained by adding a step of height 1 to a pseudo-bridge (a N step in the horizontal model, a N or E step in the diagonal model). It is thus equivalent to count bridges or pseudo-bridges.

By Proposition 2, the enumeration of weakly directed bridges boils down to the enumeration of (irreducible) partially directed bridges. In this section and the following one, we address the enumeration of these building blocks, first in a rather systematic way based on a step-by-step construction, then in a more combinatorial way based on *heaps of cycles*.

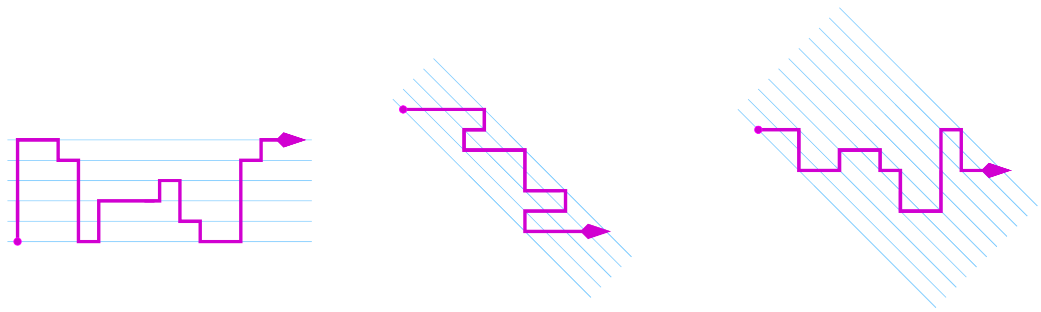


Fig. 4: A NES-pseudo-bridge in the horizontal model. (b) An ESW-pseudo-bridge in the diagonal model. (c) A NES-pseudo-bridge in the diagonal model.

As partially directed walks are defined by the avoidance of (at least) one step, there are four kinds of these. Hence, in principle, we should count, for each model (horizontal and diagonal), four families of partially directed bridges. However, in the horizontal model, there exists no ESW-bridge, and every NEW-walk is a pseudo-bridge. The latter class of walks is very easy to count. Moreover, a symmetry transforms NES-bridges into NSW-bridges, so that there is really one class of bridges that we need to count. In the diagonal model, we need to count ESW-bridges (which are equivalent to NSW-bridges by a diagonal symmetry) and NES-bridges (which are equivalent to NEW-bridges). Finally, to avoid certain ambiguities, we need to count ES-bridges, but this has already been done in [5].

From now on, the starting point of our walks is always at height 0. The height of a walk is then defined to be the maximal height reached by its vertices.

3.1 NES*-bridges in the horizontal model*

Proposition 3 *Let $k \geq 0$. In the horizontal model, the length generating function of* NES*-pseudo-bridges of height k is*

$$B^{(k)}(t) = \frac{t^k}{G_k(t)},$$

where $G_k(t)$ is the sequence of polynomials defined by

$$G_{-1} = 1, \quad G_0 = 1 - t, \quad \text{and for } k \geq 0, \quad G_{k+1} = (1 - t + t^2 + t^3)G_k - t^2 G_{k-1}.$$

Equivalently,

$$\sum_{k \geq 0} \frac{v^k t^k}{B^{(k)}(t)} = \sum_{k \geq 0} v^k G_k = \frac{1 - t - t^2 v}{1 - (1 - t + t^2 + t^3)v + t^2 v^2},$$

or

$$B^{(k)}(t) = \frac{U - \bar{U}}{\big((1-t)U - t\big)U^k - \big((1-t)\bar{U} - t\big)\bar{U}^k},$$

where

$$U = \frac{1 - t + t^2 + t^3 - \sqrt{(1 - t^4)\,(1 - 2t - t^2)}}{2t}$$

is a root of $tu^2 - \left(1 - t + t^2 + t^3\right) u + t = 0$ and $\bar{U} := 1/U$ is the other root of this polynomial.

Proof: Let $\mathcal{T}$ be the set of NES-walks that end with an E step, and in which each vertex v satisfies $0 \leq h(v) \leq k$. Let $\mathcal{T}_i$ be the subset of $\mathcal{T}$ consisting of walks that end at height i. Let $T_i(t) \equiv T_i$ be the length generating function of $\mathcal{T}_i$, and define the bivariate generating function

$$T(t;u) \equiv T(u) = \sum_{i=0}^{k} T_i(t)u^i.$$

This series counts walks of $\mathcal{T}$ by their length and the height of their endpoint. Pseudo-bridges walks of height k containing at least one E step are obtained by adding a sequence of N steps of appropriate length to a walk of $\mathcal{T}$, and this gives

$$B^{(k)}(t) = t^k + \sum_{i=0}^{k} T_i(t)t^{k-i} = t^k\left(1 + T(1/t)\right) \tag{2}$$

(the term t^k accounts for the walk formed of k consecutive N steps.)

A step-by-step construction of the walks of $\mathcal{T}$ yields the following lemma.

Lemma 4 *Let $\bar{u} = 1/u$. The series $T(t;u)$, denoted $T(u)$ for short, satisfies the following equation:*

$$\left(1 - \frac{ut^2}{1-tu} - \frac{t}{1-t\bar{u}}\right) T(u) = t\,\frac{1-(tu)^{k+1}}{1-tu} - t\,\frac{(tu)^{k+1}}{1-tu}T(1/t) - \frac{t^2\bar{u}}{1-t\bar{u}}T(t).$$

The equation of Lemma 4 is easily solved using the *kernel method* (see e.g. [2, 4, 23]). The *kernel* of the equation is the coefficient of $T(u)$, namely

$$1 - \frac{ut^2}{1-tu} - \frac{t}{1-t\bar{u}}.$$

It vanishes when $u = U$ and $u = \bar{U} := 1/U$, where U is defined in the lemma. Since $T(u)$ is a polynomial in u, the series $T(U)$ and $T(\bar{U})$ are well-defined. Replacing u by U or $\bar{U}$ in the functional equation cancels the left-hand side, and hence the right-hand side. One thus obtains two linear equations between $T(t)$ and $T(1/t)$, which involve the series U. Solving them gives in particular the value of $T(1/t)$, and thus of $B^{(k)}(t)$ (thanks to (2)). This provides the second expression of $B^{(k)}(t)$ given in the lemma. The other results easily follow, using elementary arguments about linear recurrence relations and rational generating functions. ∎

3.2 ESW*-bridges in the diagonal model*

Proposition 5 *Let $k \geq 0$. In the diagonal model, the length generating function of* ESW*-pseudo-bridges of height k is*

$$B_1^{(k)}(t) = \frac{t^k}{G_k(t)},$$

where $G_k(t)$ is the sequence of polynomials defined by

$$G_0 = 1, \quad G_1 = 1 - t^2 \quad \textit{and for } k \geq 1, \quad G_{k+1} = (1+t^2)G_k - t^2(2-t^2)G_{k-1}.$$

The length generating function of NES*-pseudo-bridges of height k is*

$$B_2^{(k)}(t) = (2-t^2)^k B_1^{(k)}(t).$$

Finally, the length generating function of ES*-pseudo-bridges of height k is*

$$B_0^{(k)}(t) = \frac{t^k}{F_k(t)},$$

where the sequence $F_k(t)$ is defined by $F_{-1} = 1$, $F_0 = 1$, and $F_{k+1} = F_k - t^2 F_{k-1}$ for $k \geq 0$.

4 Partially directed bridges via heaps of cycles

In this section, we give alternative (and more combinatorial) proofs of the results of Section 3. In particular, these proofs explain why the numerators of series counting partially directed bridges of height k are so simple (t^k or $t^k(2-t^2)^k$, depending on the model).

Let $\Gamma = (V, E)$ be a directed graph. To each edge of this graph, we associate a weight taken in some commutative ring (typically, a ring of formal power series). A *cycle* of Γ is a path ending at its starting point, taken up to a cyclic permutation. A path is *self-avoiding* if it does not visit the same vertex twice. A self-avoiding cycle is called an *elementary cycle*. Two paths are *disjoint* if their vertex sets are disjoint. The *weight* $w(\pi)$ of a path (or cycle) π is the product of the weights of its edges. A *configuration of cycles* $\gamma = \{\gamma_1, \ldots, \gamma_r\}$ is a set of pairwise disjoint elementary cycles. The *signed weight* of γ is

$$\widetilde{w}(\gamma) := (-1)^r \prod_{i=1}^{r} w(\gamma_i).$$

For two vertices i and j, denote by $W_{i,j}$ the generating function of paths going from from i to j:

$$W_{i,j} = \sum_{\pi : i \rightsquigarrow j} w(\pi).$$

We assume that this sum is well-defined, which is always the case when considering length generating functions.

Proposition 6 *The generating function of paths going from i to j in the weighted digraph Γ is*

$$W_{i,j} = \frac{N_{i,j}}{G},$$

where $G = \sum_\gamma \widetilde{w}(\gamma)$ is the signed generating function of configuration of cycles, and

$$N_{i,j} = \sum_{\eta,\gamma} w(\eta)\widetilde{w}(\gamma),$$

where η is a self-avoiding path going from i to j and γ is a configuration of cycles disjoint from η.

This result can be proved as follows: one first identifies $N_{i,j}$ as the (i,j) coefficient of the matrix $(1-A)^{-1}$, where A is the adjacency matrix of Γ. Thanks to standard linear algebra, this coefficient can be expressed in terms of the determinant of $(1-A)$ and one of its cofactors. A simple expansion of these as sums over permutations shows that the determinant is G, and the cofactor $N_{i,j}$. Proposition 6 can also be proved without any reference to linear algebra, using the theory of *partially commutative monoids*, or *heaps of pieces* [11, 25]. In this context, configurations of cycles are called *trivial heaps of cycles*. This justifies the title of this section.

4.1 Bridges with large down steps

Let Γ_k be the graph with vertices $\{0,\dots,k\}$ and with the following weighted edges:
– *ascending* edges of height 1, $i \to i+1$, with weight A, for $i = 0,\dots,k-1$;
– *descending* edges of height h, $i \to i-h$, with weight D_h, for $i = h,\dots,k$ and $h \geq 0$.
For $k \geq 0$, denote by $C^{(k)}$ the generating function of paths from 0 to k in the graph Γ_k. These paths may be seen as pseudo-bridges of height k with general down steps.

Lemma 7 *The generating function of pseudo-bridges of height k is*

$$C^{(k)} = \frac{A^k}{H_k},$$

where the generating function of the denominators H_k is

$$\sum_{k\geq 0} H_k v^k = \frac{1-D(vA)}{1-v+vD(vA)}, \tag{3}$$

with $D(v)$ the generating function of descending steps:

$$D(v) = \sum_{h\geq 0} D_h v^h.$$

Proof: With the notation of Proposition 6, the series $C^{(k)}$ reads $N_{0,k}/G$. Since all ascending edges have height 1, the only self-avoiding path from 0 to k consists of k ascending edges, and has weight A^k. As it visits every vertex of the graph, the only configuration of cycles disjoint from it is the empty configuration. Therefore, the numerator $N_{0,k}$ is simply A^k. The elementary cycles consist of a descending step of height, say, h, followed by h ascending steps. The weight of this cycle is $D_h A^h$.

To underline the dependance of our graph in k, denote by H_k the denominator G of Proposition 6. Consider a configuration of cycles of Γ_k: either the vertex k is free, or it is occupied by a cycle; this gives the following recurrence relation, valid for $k \geq 0$:

$$H_k = H_{k-1} - \sum_{h=0}^{k} D_h A^h H_{k-h-1}.$$

with the initial condition $H_{-1} = 1$. This is equivalent to (3). ■

4.2 Partially directed self-avoiding walks as arbitrary paths

It is not straightforward to apply Proposition 6 (or Lemma 7) to the enumeration of partially directed bridges, because of the self-avoidance condition. To circumvent this difficulty, we will first prove that partially directed self-avoiding walks are arbitrary paths on a graph with generalized steps. We only deal with the horizontal model, but the diagonal model can be addressed in a similar way.

Let us say that a NES-walk is *proper* if it neither begins nor ends with a S step. All NES-pseudo-bridges are proper, whether in the horizontal or diagonal model. The following lemma explains how to see proper NES-walks as sequences of generalized steps.

Lemma 8 *Every proper* NES*-walk has a unique factorization into* N *steps and nonempty proper* ES*-walks with no consecutive* E *steps.*

This result, combined with Lemma 7, gives an alternative proof of Proposition 3.

5 Weakly directed walks: the horizontal model

We now return to the weakly directed walks defined in Section 2. We determine their generating function, study their asymptotic number and average end-to-end distance, and finally prove that the generating function we have obtained has infinitely many singularities, and hence, cannot be D-finite.

5.1 Generating function

By combining Propositions 2 and 3, it is now easy to count weakly directed *bridges*.

Proposition 9 *In the horizontal model, the generating function of weakly directed bridges is:*

$$W(t) = \frac{1}{1 + t - \frac{2tB}{1+tB}}$$

where $B := \sum_{k\geq 0} B^{(k)}(t)$ *is the generating function of* NES*-pseudo-bridges, given by Proposition* 3.

The generating function of *general* weakly directed walks is a bit more involved, but the numbers of weakly directed walks and bridges of length n only differ asymptotically by a multiplicative constant.

5.2 Asymptotic results

Proposition 10 *The generating function* W *of weakly directed bridges, given in Proposition* 9*, is meromorphic in the disk* $\mathcal{D} = \{z : |z| < \sqrt{2}-1\}$*. It has a unique dominant pole in this disk,* $\rho \simeq 0.3929$*. This pole is simple. Consequently, the number* w_n *of weakly directed bridges of length* n *satisfies* $w_n \sim \kappa\mu^n$*, with* $\mu = 1/\rho \simeq 2.5447$*.*

Let N_n *denote the number of irreducible bridges in a random weakly directed bridge of length* n*. The mean and variance of* N_n *satisfy:*

$$\mathbb{E}(N_n) \sim \mathfrak{m}\, n, \qquad \mathbb{V}(N_n) \sim \mathfrak{s}^2\, n,$$

where $\mathfrak{m} \simeq 0.318$ *and* $\mathfrak{s}^2 \simeq 0.7$*, and the random variable* $\frac{N_n - \mathfrak{m}\, n}{\mathfrak{s}\sqrt{n}}$ *converges in law to a standard normal distribution. In particular, the average end-to-end distance, being bounded from below by* $\mathbb{E}(N_n)$*, grows linearly with* n*.*

We have designed an algorithm for the random generation of weakly directed bridges, using a *Boltzmann sampler* [9]. A sample output of this algorithm is shown in Figure 5, confirming the linear form of weakly directed bridges.

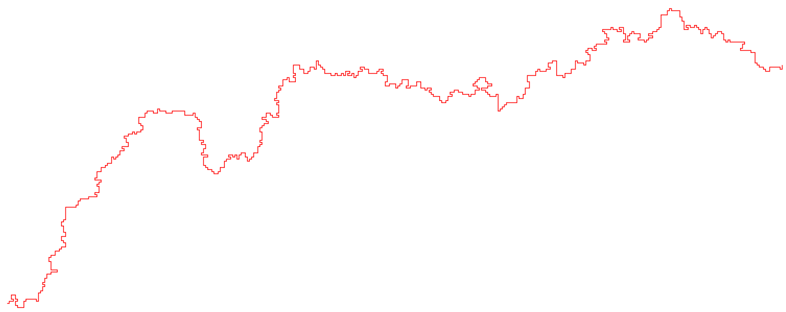


Fig. 5: A random weakly directed bridge of length 1009, rotated by $90°$.

5.3 Nature of the series

Proposition 11 *The generating function B of* NES*-pseudo-bridges, given in Proposition* 3*, converges around* 0 *and has a meromorphic continuation in $\mathbb{C} \setminus \mathcal{E}$, where $\mathcal{E}$ consists of the two real intervals $[-\sqrt{2}-1, -1]$ and $[\sqrt{2}-1, 1]$, and of the curve*

$$\mathcal{E}_0 = \left\{ x + iy : x \geq 0,\ y^2 = \frac{1 - x^2 - 2\,x^3}{1 + 2\,x} \right\}. \tag{4}$$

This curve, shown in Fig. 6*, is a natural boundary of B. That is, every point of $\mathcal{E}_0$ is a singularity of B.*

The above statements hold as well for the generating function W of weakly directed bridges. In particular, neither B nor W is D-finite.

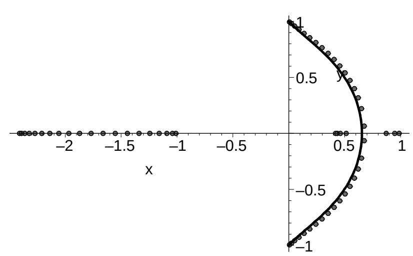


Fig. 6: The curve $\mathcal{E}_0$ and the zeroes of the polynomial G_{20}.

6 The diagonal model

We have defined weakly directed walks in Section 2 by requiring that the portion of the walk joining two visits to the same diagonal is partially directed. Here, we study a proper subclass of these walks, consisting of bridges formed of partially directed irreducible bridges.

Proposition 12 *The generating function of bridges formed of partially directed irreducible bridges is*

$$W_\Delta(t) = \frac{1}{1+2t-\dfrac{2tB_1}{1+tB_1}-\dfrac{4tB_2}{1+2tB_2}+\dfrac{2tB_0}{1+tB_0}},$$

where the series $B_i = \sum_{k\geq 0} B_i^{(k)}(t)$ *are given in Proposition* 5.

The growth constant is found to be a bit smaller than in the horizontal model (about 2.5378). The end-to-end distance is again linear.

References

[1] S. E. Alm and S. Janson. Random self-avoiding walks on one-dimensional lattices. *Comm. Statist. Stochastic Models*, 6(2):169–212, 1990.

[2] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy, and D. Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Math.*, 246(1-3):29–55, 2002.

[3] M. Bousquet-Mélou. Families of prudent self-avoiding walks. *J. Combin. Theory Ser. A*, 117(3):313–344, 2010. Arxiv:0804.4843.

[4] M. Bousquet-Mélou and M. Petkovšek. Linear recurrences with constant coefficients: the multivariate case. *Discrete Math.*, 225(1-3):51–75, 2000.

[5] M. Bousquet-Mélou and Y. Ponty. Culminating paths. *Discrete Math. Theoret. Comput. Sci.*, 10(2), 2008. ArXiv:0706.0694.

[6] P. G. de Gennes. Exponents for the excluded volume problem as derived by the Wilson method. *Phys. Lett. A*, 38(5):339–340, 1972.

[7] J. C. Dethridge and A. J. Guttmann. Prudent self-avoiding walks. *Entropy*, 8:283–294, 2008.

[8] E. Duchi. On some classes of prudent walks. In *FPSAC'05*, Taormina, Italy, 2005.

[9] Ph. Duchon, Ph. Flajolet, G. Louchard, and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combin. Probab. Comput.*, 13(4-5):577–625, 2004.

[10] B. Duplantier and I. K. Kostov. Geometrical critical phenomena on a random surface of arbitrary genus. *Nucl. Phys. B*, 340(2-3):491–541, 1990.

[11] D. Foata. A noncommutative version of the matrix inversion formula. *Adv. Math.*, 31:330–349, 1979.

[12] A. J. Guttmann and A. R. Conway. Square lattice self-avoiding walks and polygons. *Ann. Comb.*, 5(3-4):319–345, 2001.

[13] A. J. Guttmann and N. C. Wormald. On the number of spiral self-avoiding walks. *J. Phys. A: Math. Gen.*, 17:L271–L274, 1984.

[14] J. M. Hammersley and D. J. A. Welsh. Further results on the rate of convergence to the connective constant of the hypercubical lattice. *Q. J. Math., Oxf. II. Ser.*, 13:108–110, 1962.

[15] I. Jensen. Improved lower bounds on the connective constants for two-dimensional self-avoiding walks. *J. Phys. A: Math. Gen.*, 37(48):11521–11529, 2004.

[16] I. Jensen and A. J. Guttmann. Self-avoiding polygons on the square lattice. *J. Phys. A*, 32(26):4867–4876, 1999.

[17] T. Kennedy. A faster implementation of the pivot algorithm for self-avoiding walks. *J. Statist. Phys.*, 106(3-4):407–429, 2002.

[18] H. Kesten. On the number of self-avoiding walks. *J. Math. Phys.*, 4(7):960–969, 1963.

[19] G. F. Lawler, O. Schramm, and W. Werner. On the scaling limit of planar self-avoiding walk. In *Fractal geometry and applications: a jubilee of Benoît Mandelbrot, Part 2*, volume 72 of *Proc. Sympos. Pure Math.*, pages 339–364. Amer. Math. Soc., Providence, RI, 2004.

[20] N. Madras and G. Slade. *The self-avoiding walk*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, 1993.

[21] B. Nienhuis. Exact critical point and critical exponents of $o(n)$ models in two dimensions. *Phys. Rev. Lett.*, 49(15):1062–1065, 1982.

[22] V. Privman. Spiral self-avoiding walks. *J. Phys. A: Math. Gen.*, 16(15):L571–L573, 1983.

[23] H. Prodinger. The kernel method: a collection of examples. *Sém. Lothar. Combin.*, 50:Art. B50f, 19 pp. (electronic), 2003/04.

[24] A. Rechnitzer and E. J. Janse van Rensburg. Canonical Monte Carlo determination of the connective constant of self-avoiding walks. *J. Phys. A, Math. Gen.*, 35(42):L605–L612, 2002.

[25] X. G. Viennot. *Heaps of pieces, I : Basic definitions and combinatorial lemmas*, volume 1234/1986, pages 321–350. Springer Berlin / Heidelberg, 1986.

[26] D. Zeilberger. Symbol-crunching with the transfer-matrix method in order to count skinny physical creatures. *Integers*, pages A9, 34pp. (electronic), 2000.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 360–371

Involutions of the Symmetric Group and Congruence B-Orbits (Extended Abstract)

Eli Bagno[1] and Yonah Cherniavsky[2]

[1]*Jerusalem College of Technology, Israel*
[2]*Ariel University Center of Samaria, Israel*

Abstract. We study the poset of Borel congruence classes of symmetric matrices ordered by containment of closures. We give a combinatorial description of this poset and calculate its rank function. We discuss the relation between this poset and the Bruhat poset of involutions of the symmetric group. Also we present the poset of Borel congruence classes of anti-symmetric matrices ordered by containment of closures. We show that there exists a bijection between the set of these classes and the set of involutions of the symmetric group. We give two formulas for the rank function of this poset.

Résumé Nous étudions l'ensemble ordonné des classes de congruence de matrices symétriques ordonnées par containment de leurs fermetures. Nous donnons une description combinatoire de cet ensemble et calculons sa fonction rang. Nous étudions les relations entre cet ensemble et l'ensemble des involutions du groupe symérique ordonné selon l'ordre de Bruhat. Nous montrons qu'il existe une bijection parmi l'ensemble ordonné de classes de congruences de Borel des matrices anti-symétriques et l'ensemble des involutions du groupe symétrique. On termine en donnant deux formules pour la fonction rang pour ce dernier ensemble.

Keywords: Bruhat poset, congruence orbit, involutions of the symmetric group

1 Introduction

A remarkable property of the Bruhat decomposition of $GL_n(\mathbb{C})$ (i.e. the decomposition of $GL_n(\mathbb{C})$ into double cosets $\{B_1 \pi B_2\}$ where $\pi \in S_n$, $B_1, B_2 \in \mathbb{B}_n(\mathbb{C})$ – the subgroup of upper-triangular invertible matrices) is that the natural order on double cosets (defined by containment of closures) leads to the same poset as the combinatorially defined Bruhat order on permutations of S_n (for $\pi, \sigma \in S_n$, $\pi \leqslant \sigma$ if π is a subword of σ with respect to the reduced form in Coxeter generators). L. Renner introduced and developed the beautiful theory of Bruhat decomposition for not necessarily invertible matrices, see [10] and [9]. When the Borel group acts on all the matrices, the double cosets are in bijection with partial permutations which form a so called rook monoid R_n which is the finite monoid whose elements are the 0-1 matrices with at most one nonzero entry in each row and column. The group of invertible elements of R_n is isomorphic to the symmetric group S_n. Another efficient, combinatorial description of the Bruhat ordering on R_n and a useful, combinatorial formula for the length function on R_n are given by M. Can and L. Renner in [3].

The Bruhat poset of involutions of S_n was first studied by F. Incitti in [6] from a purely combinatorial point of view. He proved that this poset is graded, calculated the rank function and also showed several other important properties of this poset.

In this extended abstract we present a geometric interpretation of this poset and its natural generalization, considering the action of the Borel subgroup on symmetric matrices by congruence. Denote by $\mathbb{B}_n(\mathbb{C})$ the Borel subgroup of $GL_n(\mathbb{C})$, i.e. the group of invertible upper-triangular $n\times n$ matrices over the complex numbers. Denote by $\mathbb{S}(n,\mathbb{C})$ the set of all complex symmetric $n \times n$ matrices. The congruence action of $B \in \mathbb{B}_n(\mathbb{C})$ on $S \in \mathbb{S}(n,\mathbb{C})$ is defined in the following way:

$$S \longmapsto B^t SB \; .$$

The orbits of this action (to be precisely correct, we must say $S \mapsto \left(B^{-1}\right)^t SB^{-1}$ to get indeed a group action) are called the *congruence B-orbits*. It is known that the orbits of this action may be indexed by partial S_n-involutions (i.e. symmetric $n \times n$ matrices with at most one 1 in each row and in each column) (see [11]). Thus, if π is such a partial involution, we denote by $\mathcal{C}_\pi$ the corresponding congruence B-orbit of symmetric matrices. The poset of these orbits gives a natural extension of the Bruhat poset of regular (i.e. not partial) involutions of S_n. If we restrict this action to the set of invertible symmetric matrices we get a poset of orbits that is isomorphic to the Bruhat poset of involutions of S_n studied by F. Incitti.

Here, we give another view of the rank function of this poset, combining combinatorics with the geometric nature of it. The rank function equals to the dimension of the orbit variety. We define the combinatorial parameter $\mathfrak{D}$ which is an invariant of the orbit closure and give two combinatorial formulas for the rank function of the poset of partial involutions (Theorems 2 and 7). The result of Incitti that the Bruhat poset of involutions of S_n is graded and his formula for the rank function of this poset follow from our exposition (Corollary 1).

Also we present another graded poset of involutions of the symmetric group which also has the geometric nature, i.e. it can be described as a poset of matrix varieties ordered by containment of closures. Denote by $\mathbb{AS}(n,\mathbb{C})$ the set of all complex anti-symmetric $n \times n$ matrices. It is actually a vector space with respect to standard operations of addition and multiplication by complex scalars, also it is a Lie algebra usually denoted as $\mathfrak{so}$ with $[A,B] := AB - BA$. It is easy to see that $\mathbb{AS}(n,\mathbb{C})$ is closed under the congruence action. We consider the orbits of the congruence action of $\mathbb{B}_n(\mathbb{C})$ on $\mathbb{AS}(n,\mathbb{C})$.

The main points of this work are Proposition 2, Definition 8, Theorem **??** and Proposition 8. In Proposition 2 we show that the orbits of this action may be indexed by involutions of S_n. Then we consider the poset of these orbits ordered by containment of closures. In Definition 7 we introduce the parameter $\mathfrak{A}$ and then in Theorem 2 and Proposition 8 we give two different formulas for the rank function of the studied poset using the parameter $\mathfrak{A}$. This parameter is similar to the parameter $\mathfrak{D}$ introduced in [1] and it can be seen as a particular case of a certain unified approach to the calculation of the rank function for several "Bruhat-like" posets as we briefly discuss it at the last section of [1].

If we restrict this action on the set of invertible anti-symmetric matrices we get a poset of orbits that is isomorphic to the (reversed) Bruhat poset of involutions of S_n without fixed points which is a subposet of the poset studied by F. Incitti.

2 Preliminaries

2.1 Permutations and partial permutations. The Bruhat order

The Bruhat order on permutations of S_n is defined as follows: $\pi \leqslant \sigma$ if π is a subword of σ in Coxeter generators $s_1 = (1, 2)$, $s_2 = (2, 3)$,...,$s_{n-1} = (n-1, n)$. It it well studied from various points of view. The rank function is the length in Coxeter generators which is exactly the number of inversions in a permutation. A permutation matrix is a square matrix which has exactly one 1 in each row and each column while all other entries are zeros. A partial permutation is an injective map defined on a subset of $\{1, 2, .., n\}$. A partial permutation matrix is a square matrix which has at most one 1 at each row and each column and all other entries are zeros. So, if we delete the zero rows and columns from a partial permutation matrix we get a (regular) permutation matrix of smaller size, we will use this view later. See works of L. Renner [9] and [10] where the Bruhat order on partial permutations is introduced and studied.

2.2 Partial order on orbits

When an algebraic group acts on a set of matrices, the classical partial order on the set of all orbits is defined as follows:

$$\mathcal{O}_1 \leq \mathcal{O}_2 \iff \mathcal{O}_1 \subseteq \overline{\mathcal{O}_2}$$

where $\overline{S}$ is the (Zariski) closure of the set S.

Reminder 1 *Note that $\mathcal{O}_1 \subseteq \overline{\mathcal{O}_2} \Longrightarrow \overline{\mathcal{O}_1} \subseteq \overline{\mathcal{O}_2}$ for any two sets $\mathcal{O}_1, \mathcal{O}_2$.*

Definition 1 *As usual, a monomial matrix is a matrix which has at most one non-zero entry in each its row and in each its column.*

3 Rank-control matrices

In this section we define the *rank control matrix* which will turn out to be a key corner in the identification of our poset. We start with the following definition:

Definition 2 *Let $X = (x_{ij})$ be an $n \times m$ matrix. For each $1 \leq k \leq n$ and $1 \leq l \leq m$, denote by $X_{k\ell}$ the upper-left $k \times \ell$ submatrix of X. We denote by $R(X)$ the $n \times m$ matrix whose entries are: $r_{k\ell} = rank\,(X_{k\ell})$ and call it the rank control matrix of X.*

It follows from the definitions that for each matrix X, the entries of $R(X)$ are nonnegative integers which do not decrease in rows and columns and each entry is not greater than its row and column number. If X is symmetric, then $R(X)$ is symmetric as well.

Reminder 2 *This rank-control matrix is similar to the one introduced by A. Melnikov [7] when she studied the poset (with respect to the covering relation given in Definition 2.2) of adjoint B-orbits of certain nilpotent strictly upper-triangular matrices.*

The rank control matrix is connected also to the work of Incitti [6] where regular involutions of S_n are discussed.

Proposition 1 *Let $X, Y \in GL_n(\mathbb{F})$ be such that $Y = LXB$ for some invertible lower-triangular matrix L and some matrix $B \in \mathbb{B}_n(\mathbb{C})$. Denote by $X_{k\ell}$ and $Y_{k\ell}$ the upper-left $k \times \ell$ submatrices of X and Y respectively. Then for all $1 \leqslant k, \ell \leqslant n$*

$$rank\,(X_{k\ell}) = rank\,(Y_{k\ell})\ .$$

Proof:
$$\begin{pmatrix} L_{kk} & 0_{k\times(n-k)} \\ * & * \end{pmatrix} \begin{pmatrix} X_{k\ell} & * \\ * & * \end{pmatrix} \begin{pmatrix} B_{\ell\ell} & * \\ 0_{(n-\ell)\times\ell} & * \end{pmatrix} = \begin{pmatrix} L_{kk}X_{k\ell}B_{\ell\ell} & * \\ * & * \end{pmatrix},$$

and therefore, $Y_{k\ell} = L_{kk}X_{k\ell}B_{\ell\ell}$. The matrices L_{kk} and $B_{\ell\ell}$ are invertible, which implies that $Y_{k\ell}$ and $X_{k\ell}$ have equal ranks. □

The rank control matrices of two permutations can be used to compare between them in the sense of Bruhat order. This is the reasoning for the next definition:

Definition 3 *Define the following order on $n \times m$ matrices with positive integer entries: Let $P = (p_{ij})$ and $Q = (q_{ij})$ be two such matrices.*
Then
$$P \leqslant_{\mathcal{R}} Q \iff p_{ij} \leqslant q_{ij} \text{ for all } i, j\,.$$

The following lemma appears in another form as Theorem 2.1.5 of [2].

Lemma 1 *Denote by $\leqslant_B$ the Bruhat order of S_n and let $\pi, \sigma \in S_n$. Then*

$$\pi \leqslant_{\mathcal{B}} \sigma \quad\iff\quad R(\pi) \geqslant_{\mathcal{R}} R(\sigma)\,.$$

In other words, the Bruhat order on permutations corresponds to the inverse order of their rank-control matrices. □

4 Partial permutations, Partial Involutions and Congruence B-Orbits

Definition 4 *A* partial permutation *is an $n\times n$ $(0,1)$-matrix such that each row and each column contains at most one '1'.*

Definition 5 *If a partial permutation matrix is symmetric, then we call it a partial involution.*

The following easily verified lemma claims that partial permutations are completely characterized by their rank control matrices.

Lemma 2 *For two $n \times n$ partial permutation matrices π, σ we have*

$$R(\pi) = R(\sigma) \iff \pi = \sigma.$$

Proof: The statement of the lemma is implied by the following simple observation: let U be the $n \times n$ upper-triangular matrix with '1's on the main diagonal and in all upper triangle and let π be any partial permutation. Then
$$R(\pi) = U^t \pi U\ .$$

□ The following theorem can be found in [11] (Theorem 3.2). It is proved by performing a symmetric version of Gauss elimination process.

Theorem 1 *There exists a bijection between the set of congruence B-orbits of symmetric matrices over $\mathbb{C}$ and the set of partial involutions.*

5 A bijection between orbits and involutions

The following Proposition 2 is somewhat similar to Theorem 3.2 in [11].

Proposition 2 *There is a bijection between the set of congruence B-orbits of all anti-symmetric $n \times n$ matrices and the set of all involutions of S_n.*

Proof: The complete proof can be found in [4]. It is done by symmetric elimination process which starts with an anti-symmetric matrix and terminates with a certain monomial anti-symmetric matrix which has 1's in its upper triangle and -1's in its lower triangle. Such matrix is unique for the given orbit and there is a bihection between the set of such matrices and involutions of S_n. This bijection is illustrated in the Example 1. □

Example 1 *The monomial anti-symmetric matrix* $\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ *corresponds to the involution*
$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 3 & 1 & 2 & 6 \end{pmatrix} \in S_6$, *which can be written as the product of disjoint transpositions as* $(1,4)(2,5)$.
a

Observation 1 *The congruence B-orbits of invertible anti-symmetric $2n \times 2n$ matrices can be indexed by involutions of S_{2n} without fixed points.*

6 The Poset of Congruence B-Orbits of Symmetric Matrices

Here is a direct consequence of Lemma 2 and Proposition 1.

Proposition 3 *All the matrices of a fixed congruence B-Orbit share a comon rank-control matrix. In other words, if π is a partial S_n-involution, and C_π is the congruence B-orbit of symmetric matrices associated with π then*

$$\mathcal{C}_\pi = \{S \in \mathbb{S}(n, \mathbb{C}) \,|\, R(S) = R(\pi)\}\,.$$

The following lemma describes the orbits:

Lemma 3 *Let π be a partial involution and let $R(\pi)$ be its rank-control matrix. Then*

$$\overline{\mathcal{C}_\pi} = \{S \in \mathbb{S}(n, \mathbb{C}) \mid R(S) \leqslant_{\mathcal{R}} R(\pi)\}\ .$$

This lemma follows from Theorem 15.31 of [8]. Their exposition differs somewhat from ours as it deals with rectangular, not necessarily symmetric matrices but the differences can be easily overwhelmed by considering also equations of the form $a_{ij} = a_{ji}$ which are polynomial equations with regard to the entries of a matrix.

Reminder 3 *Over the fields $\mathbb{C}$ and $\mathbb{R}$ the closure in Lemma 5 may also be considered with respect to the metric topology.*

The next corollary follows from Lemma 5 and characterizes the order relation of the poset of B-orbits.

Let π and σ be partial S_n-involutions. Then

$$\mathcal{C}_\pi \leqslant_{\mathcal{O}} \mathcal{C}_\sigma \iff R(\pi) \leqslant_{\mathcal{R}} R(\sigma)$$

Explicit examples for $n = 3$ in the symmetric case can be found in [1] and for $n = 4$ in the antisymmetric case can be found in [4].

7 The Poset of Congruence B-Orbits of Anti-Symmetric Matrices

Here is a direct consequence of Proposition 1.

Proposition 4 *All the matrices of a fixed congruence B-Orbit have the same rank-control matrix. In other words, if $X \in \mathbb{AS}(n, \mathbb{C})$ and $\mathcal{A}_X$ is the congruence B-orbit of X, then*

$$\mathcal{A}_X = \{S \in \mathbb{AS}(n, \mathbb{C}) \,|\, R(S) = R(X)\} \ .$$

Similarly to the symmetric case we give the proposition which describes the orbit closures in the anti-symmetric case. This proposition also follows from Theorem 15.31 given by E. Miller and B. Sturmfels, see [8, Chapter 15, page 301]:

Proposition 5 *Let X be an anti-symmetric matrix and let $R(X)$ be its rank-control matrix. Then*

$$\overline{\mathcal{A}_X} = \{S \in \mathbb{AS}(n, \mathbb{C}) \,|\, R(S) \leqslant_{\mathcal{R}} R(X)\} \ .$$

The next corollary characterizes the order relation of the poset of B-orbits.

Corollary 1 *Let $X, Y \in \mathbb{AS}(n, \mathbb{C})$. Then*

$$\mathcal{A}_X \leqslant_{\mathcal{O}} \mathcal{A}_Y \iff R(X) \leqslant_{\mathcal{R}} R(Y)$$

8 The Rank Function

Definition 6 *A poset P is called graded (or ranked) if for every $x, y \in P$, any two maximal chains from x to y have the same length.*

Proposition 6 *The poset of congruence B-orbits of symmetric matrices and the poset of congruence B-orbits of anti-symmetric matrices(with respect to the order $\leqslant_{\mathcal{O}}$) are graded posets with the rank function given by the dimension of the closure.*

This proposition is a particular case of the following fact. Let G be a connected, solvable group acting on an irreducible, affine variety X. Suppose that there are a finite number of orbits. Let O be the set of G-orbits on X. For $x, y \in O$ define $x \leqslant y$ if $x \subseteq \overline{y}$. Then O is a graded poset.

This fact is given as an exercise in [10] (exercise 12, page 151) and can be proved using the proof of the theorem appearing of Section 8 of [9]. (Note that in our case the Borel group is solvable, the varieties of all symmetric and anti-symmetric matrices are irreducible because they are vector spaces and the number of orbits is finite since there are only finitely many partial permutation.)

A natural problem is to find an algorithm which calculates the dimension of the orbit closure from the monomial matrix or from its rank-control matrix. Here we present such an algorithm.

Definition 7 *Let π be a partial involution matrix and let $R(\pi) = (r_{ij})$ be its rank-control matrix. Add an extra 0 row to $R(\pi)$, pushed one place to the left, i.e. assume that $r_{0k} = 0$ for each $0 \leqslant k < n$.*

Denote

$$\mathfrak{D}(\pi) = \#\left\{(i,j) \,|\, 1 \leqslant i \leqslant j \leqslant n \quad \textit{and} \quad r_{ij} = r_{i-1,j-1}\right\}.$$

Definition 8 *Let $X \in \mathbb{AS}(n,\mathbb{C})$ and let $R(X) = (r_{ij})_{i,j=1}^{n}$ be the rank-control matrix of X. Add an extra 0 row to $R(X)$, pushed one place to the left, i.e. assume that $r_{0k} = 0$ for each $0 \leqslant k < n$. Denote*

$$\mathfrak{A}(X) = \#\left\{(i,j) \,|\, 1 \leqslant i < j \leqslant n \quad \textit{and} \quad r_{ij} = r_{i-1,j-1}\right\}.$$

The first parameter $\mathfrak{D}$ counts equalities in the diagonals of the upper triangle of the rank-control matrix including the main diagonal and the second parameter $\mathfrak{A}$ counts equalities in the diagonals of the upper triangle of the rank-control matrix without the main diagonal.

Theorem 2 *Let π be a partial S_n-involution. As above $\mathcal{C}_\pi$ denotes the orbit of symmetric matrices which corresponds to π. Then*

$$\dim \overline{\mathcal{C}_\pi} = \frac{n^2+n}{2} - \mathfrak{D}(\pi).$$

Proof: Consider the vector space

$$\mathbb{C}^{n^2} = \left\{ [a_{ij}]_{i,j=1}^{n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \,:\, a_{ij} \in \mathbb{C} \right\}.$$

Let X be some set of pairs of indexes, i.e. $X \subseteq \{(i,j) \,:\, 1 \leqslant i,j \leqslant n\}$. Define a subspace $W_X \subset \mathbb{C}^{n^2}$ of dimension $n^2 - |X|$ in the following way:

$$W_X = \left\{[a_{i_t j_t}] \,:\, (i_t, j_t) \notin X\right\},$$

i.e. W_X is spanned by the elements of the standard basis of $\mathbb{C}^{n^2}$ which we index by all pairs of indices not belonging to X.

Consider also the natural projection $p_X : \mathbb{C}^{n^2} \to W_X$. Since we consider elements of $\mathbb{C}^{n^2}$ as $n \times n$ matrices, we denote elements of W_X as matrices with empty boxes in the positions whose indexes are in X. For example, consider

$$\mathbb{C}^{3^2} = \left\{ [a_{ij}]_{i,j=1}^{3} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \,:\, a_{ij} \in \mathbb{C} \right\}$$

and let $X = \{(2,3),(3,2),(3,3)\}$. Then

$$W_X = \left\{ [a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{31}] = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & \square \\ a_{31} & \square & \square \end{bmatrix} \,:\, a_{ij} \in \mathbb{C} \right\} \subset \mathbb{C}^{3^2}.$$

In this example the natural projection $p_X : \mathbb{C}^{3^2} \to W_X$ is

$$p_X\left([a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}]\right) = [a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{31}]$$

or in the matrix notation

$$p_X\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & \square \\ a_{31} & \square & \square \end{bmatrix}.$$

By a *fragment* of an $n \times n$ matrix we mean the image of this matrix under the projection p_X with certain X.

Denote

$$V^{kn} = \overline{p_X(\overline{\mathcal{C}_\pi})}$$

where $X = \{(k+1,n),(n,k+1),(k+2,n),(n,k+2),...(n,n)\}$.

The variety V^{kn} corresponds to the fragments of V with empty entries in the n-th row and column: the last non-empty entry in the n-th column is in the row number k, all further positions in the n-th row and column are empty.

Observation 2 *Let V be a variety in $\mathbb{C}^n$ which is described by the polynomial equations*

$$f_1(x_1,...,x_n) = 0, f_2(x_1,...,x_n) = 0, ..., f_k(x_1,...,x_n) = 0$$

and let $p : \mathbb{C}^n \to \mathbb{C}^{n-k}$ be the natural projection

$$p(x_1,x_2,...,x_{n-k},...,x_{n-1},x_n) = (x_1,x_2,...,x_{n-k}).$$

Then

$$\overline{p(V)} = \{(x_1,x_2,...,x_{n-k}) \in \mathbb{C}^{n-k} \ : \ f_{i_1} = 0, f_{i_2} = 0, ..., f_{i_t} = 0\}$$

where the equations $f_{i_j} = 0$ appearing here are only those which do not include the variables x_{n-k+1}, x_{n-k+2},...,x_n, i.e. only those f_i whose partial derivatives by with respect to the variables x_{n-k+1}, x_{n-k+2},...,x_n are zeros.

Observation 3 *Note that since V^{kn} and $V^{k-1,n}$ are projections of the same variety $\overline{\mathcal{C}_\pi}$ and V^{kn} has one more coordinate than $V^{k-1,n}$, there are only two possibilities for their dimensions:* $\dim V^{kn} = \dim V^{k-1,n}$ *or* $\dim V^{kn} = \dim V^{k-1,n} + 1$.

(This is true since the rank of the Jacobian matrix can change only by 1 when we delete the rows corresponding to the coordinates.)

Now, let us start the course of the proof, by induction on n. For $n = 1$ the statement is obviously true.

Let π_n be any partial S_n involution. Denote by π_{n-1} its upper-left $n-1 \times n-1$ submatrix (which is an S_{n-1} partial involution by itself). Denote by $R(\pi_n)$, $(R(\pi_{n-1}))$ the corresponding rank-control matrices.

By the induction hypothesis, $\dim \overline{\mathcal{C}_{\pi_{n-1}}} = \frac{n^2-n}{2} - \mathfrak{D}(\pi_{n-1})$. Now we add to π_{n-1} the n-th column and consider the n-th column of $R(\pi_n)$. (We also add the n-th row but since our matrices are symmetric it suffices to check the dimension when we add the n-th column.) We added n new coordinates to the variety $\overline{\mathcal{C}_{\pi_{n-1}}}$ and we have to show that

$$\dim \overline{\mathcal{C}_\pi} = \dim \overline{\mathcal{C}_{\pi_{n-1}}} + n - \#\left\{(i,n) \,|\, 1 \leqslant i \leqslant n \quad \text{and} \quad r_{in} = r_{i-1,n-1}\right\}, \qquad (*)$$

The equality $(*)$ implies the statement of our theorem since $\frac{n^2-n}{2} + n = \frac{n^2+n}{2}$ and

$$\mathfrak{D}(\pi) = \mathfrak{D}(\pi_{n-1}) + \#\left\{(i,n) \,|\, 1 \leqslant i \leqslant n \quad \text{and} \quad r_{in} = r_{i-1,n-1}\right\}.$$

Obviously, if $r_{1,n}=0$, then $a_{1,n}=0$ for any $A=(a_{ij})_{i,j=1}^{n}\in\overline{\mathcal{C}_\pi}$. This itself is a polynomial equation which decreases the dimension by 1.

If, on the other hand, $r_{1,n}=1$, it means that the rank of the first row is maximal and therefore, no equation is involved. In other words, the dimension of the variety V^{1n} is one more than the dimension of the variety V^{0n}, corresponding to and they have equal dimensions when $r_{1,n}=0$.

Now move down along the n-th column of $R(\pi_n)$. Again, by induction, this time on the number of rows, assume that for each $1\leqslant i\leqslant k-1$ $dimV^{in}=dimV^{i-1,n}$ if and only if $r_{i-1,n-1}=r_{i,n}$ while $\dim V^{in}=\dim V^{i-1,n}+1$ if and only if $r_{i-1,n-1}<r_{i,n}$.

First, let $r_{k-1,n-1}=r_{k,n}=c$. Consider a matrix $A=(a_{ij})_{i,j=1}^{n}\in\overline{\mathcal{C}_\pi}$ and its upper-left $(k-1)\times(n-1)$ submatrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1}\\ a_{21} & a_{22} & \cdots & a_{2,n-1}\\ \cdots & \cdots & \cdots & \cdots\\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,n-1}\end{bmatrix}.$$

Using the notation introduced in Proposition 1, we denote this submatrix as $A_{k-1,n-1}$.

If $c=0$, then $rankA_{kn}=0$, so A_{kn} is a zero matrix and thus $\dim V^{in}=\dim V^{i-1,n}=0$.

Let $c\neq 0$. Since $rank\,(A_{k-1,n-1})=c$, we can take c linearly independent columns $\begin{bmatrix} a_{1,j_1}\\ a_{2,j_1}\\ \cdots\\ a_{k-1,j_1}\end{bmatrix}$, ... , $\begin{bmatrix} a_{1,j_c}\\ a_{2,j_c}\\ \cdots\\ a_{k-1,j_c}\end{bmatrix}$ which span its column space. Now take only the linearly independent rows of the $(k-1)\times c$ matrix $\begin{bmatrix} a_{1,j_1} & \cdots & a_{1,j_c}\\ a_{2,j_1} & \cdots & a_{2,j_c}\\ \cdots & \cdots & \cdots\\ a_{k-1,j_1} & \cdots & a_{k-1,j_c}\end{bmatrix}$ to get a nonsingular $c\times c$ matrix $T_c=\begin{bmatrix} a_{i_1,j_1} & \cdots & a_{i_1,j_c}\\ a_{i_2,j_1} & \cdots & a_{i_2,j_c}\\ \cdots & \cdots & \cdots\\ a_{i_c,j_1} & \cdots & a_{i_c,j_c}\end{bmatrix}$.

The equality $r_{k-1,n-1}=r_{k,n}=c\leqslant k-1$ implies that any $(c+1)\times(c+1)$ minor of the matrix A_{kn} is zero, in particular

$$\det\begin{bmatrix} a_{i_1,j_1} & \cdots & a_{i_1,j_c} & a_{i_1,n}\\ a_{i_2,j_1} & \cdots & a_{i_2,j_c} & a_{i_2,n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{i_c,j_1} & \cdots & a_{i_c,j_c} & a_{i_c,n}\\ a_{k,j_1} & \cdots & a_{k,j_c} & a_{k,n}\end{bmatrix}=0$$

which is a polynomial equation. This equation is algebraically independent of the similar equations obtained for $1\leqslant i\leqslant k-1$ since it contains a "new" variable – the entry $a_{k,n}$. It indeed involves the entry $a_{k,n}$ since $\det T_c\neq 0$. This equation means that the variable $a_{k,n}$ is not independent of the coordinates of the variety $V^{k-1,n}$, and therefore $\dim V^{k-1,n}=\dim V^{kn}$.

Now let $r_{k-1,n-1}<r_{k,n}=c$. We have to show that in this case the variable a_{nk} is independent of the coordinates of $V^{k-1,n}$, in other words, we have to show that there is no new equation. Consider the

fragment $\begin{bmatrix} r_{k-1,n-1} & r_{k-1,n} \\ r_{k-1,n} & r_{k,n} \end{bmatrix}$. There are four possible cases:

$$\begin{bmatrix} r_{k-1,n-1} & r_{k-1,n} \\ r_{k-1,n} & r_{k,n} \end{bmatrix} = \begin{bmatrix} c-1 & c-1 \\ c-1 & c \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} c-2 & c-1 \\ c-1 & c \end{bmatrix} \quad \text{or}$$
$$\begin{bmatrix} c-1 & c \\ c-1 & c \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} c-1 & c-1 \\ c & c \end{bmatrix}.$$

The equality $r_{k,n} = c$ implies that each $(c+1)\times(c+1)$ minor of A_{kn} is equal to zero, but we shall see that each such equation is not new, i.e. it is implied by the equality $r_{k,n-1} = c-1$ or by the equality $r_{k-1,n} = c-1$. In the first three cases we decompose the $(c+1)\times(c+1)$ determinant $\det \begin{bmatrix} \cdots & \cdots \\ \cdots & a_{k,n} \end{bmatrix}$ using the last column. Since in all these cases $r_{k,n-1} = c-1$, each $c\times c$ minor of this decomposition (i.e. each $c\times c$ minor of $A_{k,n-1}$) is zero and therefore, this determinant is zero. In the fourth case we get the same if we decompose the determinant using the last row instead of the last column: since $r_{k-1,n} = c-1$, all the $c\times c$ minors of this decomposition (i.e. all $c\times c$ minor of $A_{k-1,n}$) are zeros and thus, our $(c+1)\times(c+1)$ determinant equals to zero. So there is no algebraic dependence between a_{kn} and the coordinates of $V^{k-1,n}$. Therefore, $\dim V^{kn} = \dim V^{k-1,n} + 1$. The case $k = n$ is the same as other cases when $k \leqslant n-1$. The proof is completed. □

Theorem 3 *Let $\pi \in S_n$ be an involution. Denote by $\mathcal{A}_\pi$ the orbit of anti-symmetric matrices which corresponds to π.Then*

$$\dim \overline{\mathcal{A}_\pi} = \frac{n^2-n}{2} - \mathfrak{A}(\pi).$$

The proof is similar to the proof of Theorem 2 and can be found in [4].

9 Another formula for the rank function.

9.1 The symmetric case.

Obviously, an $n\times n$ partial involution matrix π can be described uniquely by the pair $(\tilde{\pi}, \{i_1, ..., i_k\})$, where $n-k$ is the rank of the matrix π, $\tilde{\pi} \in S_{n-k}$ such that $\tilde{\pi}^2 = Id$ is the regular (not partial) involution of the symmetric group S_{n-k} and the integers $i_1, ..., i_k$ are the numbers of zero rows (columns) in the matrix π.

The following theorem is a generalization of the formula for the rank function of the Bruhat poset of the involutions of S_n given by Incitti in [6]. It is indeed the rank function because we already know that the rank function is the dimension (Proposition 6) and the dimension is determined by the parameter $\mathfrak{D}$ (Theorem 2). Recall that for $\sigma \in S_n$, $inv(\sigma) = \#\{(i,j) | i < j \,\&\, \sigma(i) > \sigma(j)\}$ and $exc(\sigma) = \#\{i | \sigma(i) > i\}$.

Proposition 7 *Following Incitti, denote by $Invol(G)$ the set of all involutions in the group G. Then for a partial permutation $\pi = (\tilde{\pi}, \{i_1, ..., i_k\})$, where $\tilde{\pi} \in Invol(S_{n-k})$ and the integers $i_1, ..., i_k$ are the numbers of zero rows (columns) in the matrix π is:*

$$\mathfrak{D}(\pi) = \frac{exc(\tilde{\pi}) + inv(\tilde{\pi})}{2} + \sum_{t=1}^{k} (n+1-i_t)$$

In other words, $\mathfrak{D}(\pi)$ equals to the length of $\tilde{\pi}$ in the poset of the involutions of the group S_{n-k} plus the sum of the numbers of zero rows of the matrix π, where the numbers are taken in the opposite order, i.e. the n-th row is labeled by 1, the $(n-1)$-th row is labeled by 2,..., the first row is labeled by n.

Comment 1 *The Bruhat poset of regular (not partial) involutions of S_n is a graded poset with the rank function given by the formula*

$$\mathfrak{D}(\sigma) = \frac{exc(\sigma) + inv(\sigma)}{2} ,$$

where $\sigma \in Invol(S_n)$.

The proofs of Proposition 7 and Corollary 1 can be found in [1].

9.2 The anti-symmetric case.

Here we don't distinguish between an involution $\pi \in S_n$ and the monomial anti-symmetric matrix (with minuses in the lower triangle) associated to π by the bijection presented in Proposition 2.

Definition 9 *Let $\pi \in S_n$ be an involution. It is always possible to write it as product of disjoint transpositions*

$$\pi = (i_1, j_1)\,(i_2, j_2)\cdots(i_k, j_k)$$

in such a way that for all $1 \leqslant t \leqslant k$, $i_t < j_t$ and $i_1 < i_2 < \cdots < i_k$. Let us call it "the canonic form".

Denote by $\mathfrak{I}(\pi)$ the **number of inversions** *in the word $i_1 j_1 i_2 j_2 \cdots i_k j_k$.*

Proposition 8 *Let $\pi \in S_n$ be an involution. Then*

$$\mathfrak{A}(\pi) = \mathfrak{I}(\pi) + \sum_{a\,:\,\pi(a)=a} (n-a)\,.$$

The proof can be found in [4].

The proofs of Propositions 7 and 8 are done by induction and use Theorems 2 and 3 respectively.

References

[1] E. Bagno and Y. Cherniavsky, *Congruence B-Orbits and the Bruhat Poset of Involutions of the Symmetric Group*, preprint, available from http://arxiv.org/abs/0912.1819.

[2] A. Björner and F. Brenti, Combinatorics of Coxeter groups,Springer GTM 231, 2004.

[3] M. B. Can and L. E. Renner, *Bruhat-Chevalley order on the rook monoid*, preprint (2008), available from http://arxiv.org/abs/0803.0491.

[4] Y. Cherniavsky, *Involutions of the Symmetric Group and Congruence B-orbits of Anti-Symmetric Matrices*, preprint, available from http://arxiv.org/abs/0910.4743.

[5] A. Hultmnan, *On the combinatorics of twisted involutions in Coxeter groups*, Trans. AMS **359** (2007), 2787-2798; arXiv/0411429.

[6] F. Incitti, *The Bruhat order on the involutions of the symmetric group*, Journal of Algebraic Combinatorics **20** (2004) 243-261.

[7] A. Melnikov, *Description of B-orbit closures of order 2 in upper-triangular matrices*, Transformation Groups, Vol. 11 No. 2, 2006, pp. 217-247.

[8] E. Miller and B. Sturmfels, Combinatorial Commutative Algebra, Springer GTM 227, 2005.

[9] L. E. Renner, *Analogue of the Bruhat decomposition for algebraic monoids*, J. Algebra **101** (1986), 303–338.

[10] L. E. Renner, Linear Algebraic Monoids, Springer, 2005.

[11] F. Szechtman, *Equivalence and Congruence of Matrices under the Action of Standard Parabolic Subgroups*, Electron. J. Linear Algebra **16** (2007), 325-333.

Harmonics for deformed Steenrod operators (Extended Abstract)

François Bergeron[1†], Adriano Garsia[2‡] and Nolan Wallach[2§]

[1] *Dépt. de Math., UQAM, Montréal, Québec, H3C 3P8, CANADA.*
[2] *UCSD, 9500 Gilman Drive # 0112 La Jolla, CA 92093-0112, USA.*

received 4 Nov 2009,

Abstract. We explore in this paper the spaces of common zeros of several deformations of Steenrod operators. Proofs are omitted in view of pages limitation for the extended abstract.

Résumé. Nous explorons dans cet article l'espace des zéros communs de plusieurs déformations d'opérateurs de Steenrod. Faute de place, les preuves sont omises.

Keywords: Harmonic polynomials, Steenrod Operators

1 Introduction

In recent years many authors have studied variations on a striking classical result of invariant theory holding for any finite group W of real $n \times n$ matrices generated by reflections. Roughly stated, this result asserts that there is a natural decomposition

$$\mathbb{R}[\mathbf{x}] \simeq \mathbb{R}[\mathbf{x}]^W \otimes \mathbb{R}[\mathbf{x}]_W \tag{1}$$

of the ring of polynomials $\mathbb{R}[\mathbf{x}]$, in n variables $\mathbf{x} = x_1, x_2, \ldots, x_n$, as a tensor product of the ring $\mathbb{R}[\mathbf{x}]^W$ of W-invariant polynomials, and the "W-coinvariant-space" $\mathbb{R}[\mathbf{x}]_W$. This last is simply the space obtained as the quotient of the ring $\mathbb{R}[\mathbf{x}]$ by the ideal generated by constant-term-free W-invariant polynomials. It is well known that $\mathbb{R}[\mathbf{x}]_W$ is isomorphic as a W-module to the space $\mathcal{H}_W$ of *W-harmonic polynomials*, i.e.: the set of polynomials $f(\mathbf{x})$ that satisfy all partial differential equations of the form $p(\partial_\mathbf{x})f(\mathbf{x}) = 0$, where $p(\partial_\mathbf{x})$ any constant-term-free W-invariant polynomial in the partial derivatives ∂_i.

The purpose of this work is to study twisted versions of this setup. Typically, we replace symmetric operators $\partial_1^k + \ldots + \partial_n^k$, by operators of the form

$$D_k := \sum_{i=1}^{n} a_{i,k} x_i \partial_i^{k+1} + b_{i,k} \partial_i^k, \tag{2}$$

†supported by NSERC-Canada and FQRNT-Québec
‡supported by NSF
§supported by NSF

with some parameters $a_{i,k}$ and $b_{i,k}$. We then consider the solution set $\mathcal{H}_{\mathbf{x}}$ of the system of partial differential equations $D_k f(\mathbf{x}) = 0$, for $k \geq 1$. Observe that the operators D_k are homogeneous. We say that they are of degree $-k$ since they lower degree of polynomials by k. It follows that $\mathcal{H}_{\mathbf{x}}$ is graded by degree. In particular, it makes sense to consider the *Hilbert series*

$$H_n(t) := \sum_{d \geq 0} t^d \dim(\pi_d(\mathcal{H}_{\mathbf{x}})), \tag{3}$$

with π_d denoting the projection onto the homogeneous component of degree d. Clearly the right-hand side of (3) depends on the choice of the parameters $a_{i,k}$ and $b_{i,k}$. Recall that the Hilbert series of the space $\mathcal{H}_{\mathfrak{S}_n}$, of $\mathfrak{S}_n$-harmonic polynomials (which corresponds to setting $a_{i,k} = 0$ and $b_{i,k} = 1$) is the classical t-analog of $n!$. As we will see later, this is a "generic" value for $H_n(t)$.

Before going on with our discussion, let us consider an interesting dual point of view. Following a terminology of Wood [5], we shall say that a polynomial is a *hit-polynomial* if it can be expressed in the form

$$f(\mathbf{x}) = \sum_k D_k^* \, g_k(\mathbf{x}), \tag{4}$$

for some polynomials $g_k(\mathbf{x})$, with D_k^* standing for the dual operator of D_k with respect to the following scalar product on the ring of polynomials.

For two polynomials f and g in $\mathbb{R}[\mathbf{x}]$, one sets

$$\langle f, g \rangle := f(\partial_{\mathbf{x}}) g(\mathbf{x})\big|_{\mathbf{x}=0}. \tag{5}$$

In other words, this corresponds to the constant term of the polynomial resulting from the application of the differential operator $f(\partial_{\mathbf{x}})$ to $g(\mathbf{x})$. A straightforward computation reveals that, for two monomials $\mathbf{x}^{\mathbf{a}}$ and $\mathbf{x}^{\mathbf{b}}$, we have $\langle \mathbf{x}^{\mathbf{a}}, \mathbf{x}^{\mathbf{b}} \rangle = \mathbf{a}!$, if $\mathbf{a} = \mathbf{b}$, and 0 otherwise. Here, as is now almost usual, $\mathbf{a}!$ stands for $a_1! a_2! \cdots a_n!$. This observation makes it clear that (5) indeed defines a scalar product on $\mathbb{R}[\mathbf{x}]$. Moreover, the dual of the operator ∂_i^k is easily checked to be multiplication by x_i^k. It follows that

$$D_k^* = \sum_{i=1}^n a_{i,k} \, x_i^{k+1} \partial_i + b_{i,k} \, x_i^k.$$

From general basic linear algebra principles, it follows that the space of $\mathcal{H}_{\mathbf{x}}$, of general harmonic polynomials, is orthogonal to the space of hit-polynomials. Moreover, since the subspace of hit-polynomials is homogeneous, the corresponding quotient $\mathbf{C}$ of $\mathbb{R}[\mathbf{x}]$, by this subspace, is isomorphic to $\mathcal{H}_{\mathbf{x}}$ as a graded space.

For, the special case corresponding to setting $a_{i,k} = q$, and $b_{i,k} = 1$, for all i and k, we denote $\mathcal{H}_{\mathbf{x};q}$ resulting space which has been considered by Hivert and Thiéry (in [3]). Using the notation

$$D_{k;q} := \sum_{i=1}^n q \, x_i \partial_i^{k+1} + \partial_i^k,$$

Using a simply Lie-bracket calculation, Hivert and Thiéry have observed that $\mathcal{H}_{\mathbf{x};q}$ is simply characterized as the common solutions of the two equations $D_{1;q} f(\mathbf{x}) = 0$, and $D_{2;q} f(\mathbf{x}) = 0$. Recall that the ring of polynomials $\mathbb{R}[\mathbf{x}]$ can be considered as a $\mathfrak{S}_n$-module for the action that corresponds to permutation of the variables. This action restricts to a natural $\mathfrak{S}_n$-action on the space $\mathcal{H}_{\mathbf{x};q}$, since the operators $D_{k;q}$ are symmetric. It is classical that $\mathcal{H}_{\mathfrak{S}_n} = \mathcal{H}_{\mathbf{x};0}$ is isomorphic, as a $\mathfrak{S}_n$-module, to the regular representation of $\mathfrak{S}_n$. Hivert-Thiéry go on to state that

Conjecture 1 (Hivert-Thiéry) *As $\mathfrak{S}_n$-modules, the spaces $\mathcal{H}_{\mathbf{x};q}$ is isomorphic to $\mathcal{H}_{\mathfrak{S}_n}$, when $q > 0$. In particular, this implies that the Hilbert series of $\mathcal{H}_{\mathbf{x};q}$ is $[n]!_t$.*

It follows from (1) that the graded Frobenius characteristic $F_n(t)$ of $\mathcal{H}_{\mathbf{x};q}$ (and $\mathcal{H}_{\mathfrak{S}_n}$) is

$$F_n(t) = [n]!_t\,(1-t)^n \sum_{\lambda\vdash n} \prod_{k=1}^{n} \frac{1}{d_k!} \left(\frac{p_k}{k\,(1-t^k)}\right)^{d_k}, \tag{6}$$

where $d_k = d_k(\lambda)$ is the number of size k parts of λ.

In this work we generalize and extend the scope of the above conjecture to include the more general operators of (2). Along the way we prove several related results.

2 Tilde-Harmonics and Hat-Harmonics

We first consider another interesting special case of (2). Namely, we suppose that all $b_{i,k}$'s vanish, and all $a_{i,k}$'s are equal to 1. Thus, we consider the space of common zeros of the operators $\widetilde{D}_k := \sum_{i=1}^n x_i\,\partial_i^{k+1}$, which is called the space of *tilde-harmonics*, and denoted $\widetilde{\mathcal{H}}_{\mathbf{x}}$. We easily check that

$$[\widetilde{D}_k, \widetilde{D}_j] = (k-j)\widetilde{D}_{k+j}, \tag{7}$$

hence $\widetilde{\mathcal{H}}_{\mathbf{x}}$ is simply the set of common zeros of the two equations $\widetilde{D}_1 f(\mathbf{x}) = 0$, and $\widetilde{D}_2 f(\mathbf{x}) = 0$. The space $\widetilde{\mathcal{H}}_{\mathbf{x}}$ affords a natural action of the symmetric group, and the associated graded Frobenius characteristic is denoted $\widetilde{F}_n(t)$. Computer experimentations suggest that the Hilbert series of $\widetilde{\mathcal{H}}_{\mathbf{x}}$ seems to be

$$\widetilde{\mathcal{H}}_n(t) = \sum_{k=0}^{n} \binom{n}{k} t^k [k]_t!. \tag{8}$$

Modulo a natural conjecture, this follows from a very explicit description of $\widetilde{\mathcal{H}}_{\mathbf{x}}$ outlined below. To state it we need one more family of operators and yet another version of harmonic polynomials. For each $k \geq 1$, consider the operator $\widehat{D}_k = \sum_{i=1}^n x_i \partial_i^{k+1} + (k+1)\,\partial_i^k$, and introduce the space

$$\widehat{\mathcal{H}}_{\mathbf{x}} := \big\{ f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}] \mid \widehat{D}_k f(\mathbf{x}) = 0, \quad \forall k \geq 1 \big\},$$

whose elements are said to be "*hat-harmonics*". We will soon relate the two notions of tilde and hat harmonics. Experimentation suggest that $\widehat{\mathcal{H}}_{\mathbf{x}}$ has dimension $n!$, and that even more precisely we have the following.

Conjecture 2 *As a graded $\mathfrak{S}_n$-module, $\widehat{\mathcal{H}}_{\mathbf{x}}$ is isomorphic to the space of $\mathfrak{S}_n$-harmonics.*

Now, for any given k-subset $\mathbf{y}$ of the n variables $\mathbf{x}$, let us consider the space $\widehat{\mathbf{H}}_{\mathbf{y}}$, and write

$$e_{\mathbf{y}} := \prod_{x\in\mathbf{y}} x,$$

for the elementary symmetric polynomial of degree k in the variables $\mathbf{y}$. As usual, we define the *support* of a monomial to be the set of variable that appear in it, with non-zero exponent. Clearly, $\mathbf{y}^{\mathbf{a}}$ has support $\mathbf{y}$ if and only if $\mathbf{y}^{\mathbf{a}} = e_{\mathbf{y}}\,\mathbf{y}^{\mathbf{b}}$, for some $\mathbf{b}$. Observe that we have the operator identity we can easily check the operator identity

$$\widetilde{D}_k\, e_{\mathbf{x}} = e_{\mathbf{x}}\, \widehat{D}_k, \tag{9}$$

where $e_{\mathbf{x}}$ stands for the operator of multiplication by $e_{\mathbf{x}}$. We can now state the following remarkable fact.

Theorem 1 *The space of tilde-harmonics has the direct sum decomposition*

$$\widetilde{\mathcal{H}}_{\mathbf{x}} = \bigoplus_{\mathbf{y}\subseteq\mathbf{x}} e_{\mathbf{y}} \widehat{\mathcal{H}}_{\mathbf{y}}, \tag{10}$$

if we consider that hat-harmonics for $\mathbf{y} = \emptyset$ *are simply the scalars.*

The same holds for the more general case of operators $a_k\, \widetilde{D}_k$ and $a_k\, \widehat{D}_k$, with the a_k's equal to 0 or 1. The intent here is to restrict the set of equations considered to those k for which a_k takes the value 1. The corresponding spaces are denoted $\widetilde{\mathcal{H}}^{\mathbf{a}}_{\mathbf{x}}$ and $\widehat{\mathbf{H}}^{\mathbf{a}}_{\mathbf{x}}$, with similar convention for the corresponding Hilbert series and graded Frobenius characteristics. It follows that, even in this more general context, we have

Corollary 2 *For all choices of* a_k*,*

$$\widetilde{\mathcal{H}}^{\mathbf{a}}_n(t) = \sum_{k=0}^{n} \binom{n}{k} t^k\, \widehat{\mathcal{H}}^{\mathbf{a}}_k(t). \tag{11}$$

In particular, if conjecture 2 holds then (8) holds. There is an even finer corollary of Theorem 1.

Corollary 3 *The graded Frobenius characteristic of* $\widetilde{\mathcal{H}}^{\mathbf{a}}_{\mathbf{x}}$ *is given by the symmetric function*

$$\widetilde{F}^{\mathbf{a}}_n(t) = \sum_{k=0}^{n} t^k\, \widehat{F}^{\mathbf{a}}_k(t)\, h_{n-k}(\mathbf{z}) \tag{12}$$

A conjecture of Wood [5, conjecture 7.3] is thus partially addressed in a very explicit manner. Indeed, in view of Theorem 1, Wood's conjecture is a consequence of Conjecture 2 and the fact that $\widetilde{\mathbf{C}}$ is isomorphic to $\widetilde{\mathcal{H}}_{\mathbf{x}}$ as a graded $\mathfrak{S}_n$-module.

3 More on q-harmonics

We now link the study of harmonics of the $\widetilde{D}_k$ to further our understanding of the common zeros of the operators $D_{k:q}$, in the case when q is considered as a formal parameter. Our point of departure is the following important fact. Denote by $\nabla_k := \sum_{i=1}^{n} \partial_i^k$ the generalized Laplacian, and observe that $D_{k:q} = q\widetilde{D}_k + \nabla_k$, then we get the following.

Theorem 4 *Up to a power of* q*, every* q*-harmonic polynomial* f *may be written in the form*

$$f = f_0 + qf_1 + q^2 f_2 + \cdots + q^m f_m \tag{13}$$

with $f_i \in \mathbb{R}[\mathbf{x}]$, and such that for all $k \geq 1$ we have

$$\begin{aligned}
&\text{(a) } \nabla_k f_0 = 0 \\
&\text{(b) } \nabla_k f_i = -\widetilde{D}_k f_{i-1}, \qquad \text{for all} \quad i = 2, \ldots, m-1, \\
&\text{(c) } \widetilde{D}_k f_m = 0.
\end{aligned} \tag{14}$$

In particular, it follows that for any $r \geq 0$, and any choice of $k_1, k_2, \ldots, k_r \geq 1$, the element

$$\nabla_{k_1} \nabla_{k_2} \cdots \nabla_{k_r} f_r \tag{15}$$

is a $\mathfrak{S}_n$-harmonic polynomial in the usual sense.

Let us now reformulate the expansion of (13) in the form

$$f = q^r(f_0 + q f_1 + \cdots + q^m f_m) \qquad (\text{with each } f_i \in \mathbb{R}[\mathbf{x}],\ f_i \neq 0)$$

We call f_0 the *first term* of f and denote it "$\mathrm{FT}(f)$". Analogously we say that f_m is the *last term* of f and denote it "$\mathrm{LT}(f)$". The integer m will be called the *length* of f. We also set

$$\mathcal{H}^F_{\mathbf{x}} := \mathcal{L}[\mathrm{FT}(f) \mid f \in \mathcal{H}_{\mathbf{x};q}] \quad \text{and} \quad \mathcal{H}^L_{\mathbf{x}} := \mathcal{L}[\mathrm{LT}(f) \mid f \in \mathcal{H}_{\mathbf{x};q}] \tag{16}$$

to respectively stand for the span of first terms of q-harmonics and last terms. Using a theorem of [3] we then get the following remarkable corollary.

Corollary 5 *The three spaces $\mathcal{H}^F_{\mathbf{x}}$, $\mathcal{H}^L_{\mathbf{x}}$ and $\mathcal{H}_{\mathbf{x};q}$ are isomorphic as graded $\mathfrak{S}_n$-modules and therefore they are all isomorphic to a submodule of the Harmonics of $\mathfrak{S}_n$.*

Since the dimension of $\mathcal{H}_{\mathbf{x};q}$ is thus bounded above, the single equality $\dim \mathcal{H}_{\mathbf{x};q} = n!$ would imply that $\mathcal{H}_{\mathbf{x};q}$ affords the regular representation of $\mathfrak{S}_n$. In particular this would yield that $\mathcal{H}^F_{\mathbf{x}}$ is none other than the space of harmonics of $\mathfrak{S}_n$. Since $\mathcal{H}_{\mathbf{x};q}$ is isomorphic to $\mathcal{H}^F_{\mathbf{x}}$, as a graded $\mathfrak{S}_n$-module, it would follow that $\mathcal{H}_{\mathbf{x};q}$ itself is isomorphic to the space of harmonics of $\mathfrak{S}_n$ (as a graded $\mathfrak{S}_n$-module). Thus the Hivert-Thiéry conjecture would result.

4 The Kernel of D_k

To compute the general space $\mathcal{H}_{\mathbf{x}}$ of harmonic polynomials, we need to find common solutions of the differential equations $D_k f(\mathbf{x}) = 0$, for $k > 0$. For each k, the kernel of the operator D_k may be given a precise explicit description whenever $a_{i,k}\, d + b_{i,k} \neq 0$, for all $d \in \mathbb{N}$.We lighten the notation by writing simply a_i instead of $a_{i,k}$.

The case $k = 1$ illustrates all aspects of the method. We construct a set

$$\{\mathbf{y}^{\mathbf{r}} + \Psi_1(\mathbf{y}^{\mathbf{r}})\}_{\mathbf{r} \in \mathbb{N}^{n-1}} \tag{17}$$

which is a basis of the solution set of $D_1 f(\mathbf{x}) = 0$. Here, Ψ_1 is a linear operator described below. Simply writing x for x_n, and $\mathbf{y}$ for the set of variables $x_1, \ldots, x_{n-1}$, we expand $f \in \mathbb{R}[\mathbf{x}]$ as polynomials in x:

$$f = \sum_d f_d \frac{x^d}{d!}, \qquad \text{with } f_d \in \mathbb{R}[\mathbf{y}]. \tag{18}$$

The effect of D_1 can then be described in the format

$$D_1\left(\sum_d f_d \frac{x^d}{d!}\right) = \sum_d \left[D_1(f_d) + (d\,a_n + b_n) f_{d+1}\right] \frac{x^d}{d!}. \tag{19}$$

Setting $a := a_n$ and $b := b_n$, we now assume that $a\,d + b \neq 0$, for all $d \in \mathbb{N}$. Then, the right-hand side of (19) vanishes if and only we choose f to be such that

$$f_{d+1} = \frac{-1}{a\,d+b}\, D_1(f_d), \tag{20}$$

for all $d \geq 0$. Unfolding this recurrence for the f_d's, we find that every element of the kernel of D_1 can be written as $f_0 + \Psi_1(f_0)$, if we define the linear operator Ψ_1 as

$$\Psi_1(g) := \sum_{m \geq 1} (-1)^m \frac{D_1^m(g)}{[a;b]_m} \frac{x^m}{m!}, \qquad \text{for } g \in \mathbb{R}[\mathbf{y}]. \tag{21}$$

Here we use the notation $[a;b]_m := b\,(a+b)\,(2\,a+b)\cdots((m-1)\,a+b)$. This leads to the following theorem.

Theorem 6 *The collection of polynomials $\mathbf{y}^\mathbf{r} + \Psi_1(\mathbf{y}^\mathbf{r})$ is a basis for the kernel of D_1. In fact, given any polynomial f in the kernel of D_1, its expansion in terms of this basis is simply obtained as*

$$f = \sum_{\mathbf{r}} a_\mathbf{r}(\mathbf{y}^\mathbf{r} + \Psi_1(\mathbf{y}^\mathbf{r})) \tag{22}$$

with $(f \bmod x) = \sum_\mathbf{r} a_\mathbf{r} \mathbf{y}^\mathbf{r}$.

It follows readily that, whenever $a\,d + b \neq 0$ for all $d \in \mathbb{N}$, the Hilbert series of the dimension of the kernel of D_1 is $(1-t)^{1-n}$. In view of Theorem 1, this implies that the Hilbert series of the kernel of $\widetilde{D}_1$ is

$$1 + \sum_{k=1}^{n} \binom{n}{k} t^k \frac{1}{(1-t)^{k-1}}. \tag{23}$$

In fact, we can get an explicit description of this kernel using (9).

We can generalize formula (21) to get a description of the kernel of D_k as follows. Observe as before that

$$D_k\left(\sum_d f_d \frac{x^d}{d!}\right) = \sum_d \left[D_k(f_d) + (a\,d + b) f_{d+k}\right] \frac{x^d}{d!}. \tag{24}$$

For this expression to be zero, we must have

$$f_{d+k} = \frac{-1}{a\,d+b}\, D_k(f_d),$$

with the same conditions as before on a and b. This recurrence has a unique solution given initial values for f_d, $0 \leq d \leq k-1$. Clearly these can be fixed at leisure. Substituting the solution of the recurrence in f, we get an element of the kernel of D_k if and only if f is of the form

$$f = (f \bmod x^k) + \Psi_k(f \bmod x^k),$$

with Ψ_k the linear operator defined as

$$\Psi_k\left(\sum_{r=0}^{k-1} f_r \frac{x^r}{r!}\right) := \sum_{m\geq 1}\sum_{r=0}^{k-1}(-1)^m \frac{D_k^m(f_r)}{[a\,k;a\,r+b]_m}\,\frac{x^{k\,m+r}}{(k\,m+r)!}. \tag{25}$$

In particular, it follows that the Hilbert series of the kernel of D_k is $(1+t+\ldots t^{k-1})\,(1-t)^{1-n}$.

5 Some explicit harmonic polynomials

Common zeros of all D_k's are exactly what we are looking for. Some of these are easy to find when the D_k's are symmetric. Let λ be any partition of n, and consider a tableau τ of shape λ, this is to say a bijection

$$\tau : \lambda \longrightarrow \{1, 2, \ldots, n\},$$

with λ identified with the set of cells of its Ferrers diagram. Recall that, for $\lambda = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > 0$, the *cells* of λ are the n pairs (i,j) in $\mathbb{N}^2$, such that

$$1 \leq i \leq \lambda_j, \qquad 1 \leq j \leq k.$$

The value $\tau(i,j)$ is called an *entry* of τ, and it is said to lie in *column* i of τ. The *Garnir polynomial* of a λ-shape tableau τ, is defined to be

$$\Delta_\tau(\mathbf{x}) := \prod_{i,\; j<k} (x_{\tau(i,j)} - x_{\tau(i,k)}).$$

In other terms, the factors that appear in $\Delta_\tau(\mathbf{x})$ are differences of entries of τ that lie in the same column.

Now, define $\mathcal{V}_\lambda$ to be the linear span of the polynomials Δ_τ, for τ varying in the set of tableaux of shape λ. In formula,

$$\mathcal{V}_\lambda := \mathbb{R}\{\Delta_\tau \mid \tau \text{ tableau of shape } \lambda\}.$$

In other words, $\mathcal{V}_\lambda$ is the linear span of the Δ_τ. It is well known that this homogeneous (invariant) subspace is an irreducible representation of of $\mathfrak{S}_n$ of dimension equal to the number of standard Young tableaux. Moreover, in the ring $\mathbb{R}[\mathbf{x}]$, there exists no isomorphic copy of this irreducible representation lying in some homogeneous component of degree lower then that in which lies $\mathcal{V}_\lambda$. It is easy to check that the degrees of all of the Δ_τ's, for a tableau of shape λ, are all equal to

$$\sum_{i=1}^{\ell(\lambda)} (i-1)\,\lambda_i,$$

which is usually denoted $n(\lambda)$ in the literature (see [4]). This is the smallest possible value for the cocharge of a standard tableau of shape λ. This fact has the following easy implication.

Proposition 7 *For any tableau τ of shape λ, the Garnir polynomial $\Delta_\tau(\mathbf{x})$ is a zero of D_k, for $k \geq 1$, whenever D_k is symmetric.*

A direct consequence of this is that there is at least one copy of each irreducible representation of $\mathfrak{S}_n$ in $\mathcal{H}_{\mathbf{x}}$, when the D_k' s are all symmetric. Moreover, under the same conditions, we have

$$\sum_{\lambda \vdash n} f_\lambda \, t^{n(\lambda)} \ll H_n(t),$$

with "$\ll$" denoting coefficient wise inequality, and $H_n(t)$ as in (3).

6 A new regular sequence and a universal dimension bound

The goal of this section is to establish a bound for the dimension of $\mathcal{H}_{\mathbf{x};q}$ which is valid for all values of q. To carry this out we need some auxiliary results from commutative algebra. Let $\mathbb{F}$ be an algebraically closed field and let $\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \ldots, \theta_n(\mathbf{x})$ be homogeneous polynomials of $\mathbb{F}[\mathbf{x}]$ of respective degrees $d_1, d_2, \ldots, d_n$. The following result is basic.

Proposition 8 *The polynomials $\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \ldots, \theta_n(\mathbf{x})$ form a regular sequence in $\mathbb{F}[\mathbf{x}]$ if and only if the system of equations*

$$\theta_1(\mathbf{x}) = 0\,,\ \theta_2(\mathbf{x}) = 0\,,\ \ldots\,,\ \theta_n(\mathbf{x}) = 0$$

has, for $\mathbf{x} \in \mathbb{F}^n$, the unique solution

$$x_1 = 0\,,\ x_2 = 0\,,\ \ldots\,,\ x_n = 0.$$

We next make use of this proposition to study the sequence of polynomials

$$\varphi_m(\mathbf{x}) := \sum_{i=1}^{n} a_i x_i^m,$$

for $m \geq 0$. More precisely we seek to obtain conditions on the coefficient sequence

$$\mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{F}^n \tag{26}$$

which assure that, for a given $k \geq 1$, that the polynomials

$$\varphi_k(\mathbf{x}),\ \varphi_{k+1}(\mathbf{x}),\ \ldots,\ \varphi_{k+n-1}(\mathbf{x})$$

form a regular sequence in $\mathbb{F}[\mathbf{x}]$.

We first observe that the polynomials $\varphi_m(\mathbf{x})$, for $m > n$, may be expressed in term of the $\varphi_k(\mathbf{x})$'s, for $1 \leq k \leq n$. Indeed, recall that the ordinary elementary symmetric functions $e_r(\mathbf{x})$ may be presented in the form of the identity

$$(t - x_1)(t - x_2) \cdots (t - x_n) = \sum_{r=0}^{n} (-1)^r e_r(\mathbf{x})\, t^{n-r}.$$

Setting $t = x_i$, we obtain

$$\sum_{r=0}^{n} (-1)^r e_r(\mathbf{x})\, x_i^{n-r} = 0.$$

Multiplying both sides by $a_i\, x_i^{m-n}$ and isolating $a_i\, x_i^m$, we get

$$a_i\, x_i^m = -\sum_{r=1}^{n} (-1)^r e_r(\mathbf{x})\, a_i\, x_i^{m-r}.$$

Thus, summing up on i, the following recurrence results

$$\varphi_m(\mathbf{x}) = \sum_{r=1}^{n} (-1)^{r+1} e_r(\mathbf{x})\, \varphi_{m-r}(\mathbf{x}). \tag{27}$$

Unfolding this recurrence, we conclude that φ_m lies in the ideal $(\varphi_1, \varphi_2, \dots, \varphi_n)_{\mathbb{F}[\mathbf{x}]}$, for all $m \geq 1$.

Remark 1 It is interesting to observe that identity (27) yields that

$$\varphi_1(\mathbf{x}),\ \varphi_2(\mathbf{x}),\ \dots,\ \varphi_n(\mathbf{x}) \tag{28}$$

is never a regular sequence when $a_1 + a_2 + \cdots + a_n = 0$. Indeed, setting $m = n$ in (27), we get

$$\varphi_m(\mathbf{x}) = \sum_{r=1}^{n-1} \varphi_{m-r}(\mathbf{x})(-1)^{r+1} e_r(\mathbf{x}) + (-1)^{n+1} e_n(\mathbf{x})\big(a_1 + a_2 + \cdots + a_n\big)$$

and thus the vanishing of $a_1 + a_2 + \cdots + a_n$ forces $\varphi_n(\mathbf{x})$ to vanish modulo the ideal

$$(\varphi_1, \varphi_2, \dots, \varphi_{n-1})_{\mathbb{F}[\mathbf{x}]}.$$

Let us now denote

$$\Phi_n^k := (\varphi_k, \varphi_{k+1}, \dots, \varphi_{k+n-1})_{\mathbb{F}[\mathbf{x}]},$$

the ideal in $\mathbb{F}[\mathbf{x}]$ generated by the n polynomials $\varphi_\ell(\mathbf{x})$, with $k \leq \ell \leq k+n-1$. We also write Φ_n for Φ_n^1. Proposition 8 and (27) combine to yield the following remarkable result.

Theorem 9 *For any $k \geq 1$ the sequence*

$$\varphi_k(\mathbf{x}),\ \varphi_{k+1}(\mathbf{x}),\ \dots,\ \varphi_{k+n-1}(\mathbf{x}), \tag{29}$$

is regular if and only if the sequence

$$\varphi_1(\mathbf{x}),\ \varphi_2(\mathbf{x}),\ \dots,\ \varphi_n(\mathbf{x}), \tag{30}$$

is regular.

This given, here and after we need only be concerned with finding conditions on $a_1, a_2, \dots, a_n$ that assure the regularity of sequence $\varphi_1, \varphi_2, \dots, \varphi_n$. The following result offers a useful criterion.

Theorem 10 *In the ring $\mathbb{F}[\mathbf{x}]$, the polynomials*

$$\varphi_1, \varphi_2, \dots, \varphi_n$$

form a regular sequence if and only if we have

$$x_i^{\binom{n}{2}+1} \in \Phi_n. \tag{31}$$

When this happens we have the Hilbert series equalities

$$F_{\mathbb{F}[x]/\Phi_n^k}(t) = [k]_t[k+1]_t \cdots [k+n-1]_t \tag{32}$$

and, in particular,

$$\dim \mathbb{F}[x]/\Phi_n^k = (k)(k+1)\cdots(k+n-1).$$

Going along the lines of Remark 1, we are now ready to assert the following characterization of the a_i's for which we have regularity.

Theorem 11 *For $k > 1$, the sequences*

$$\varphi_k, \varphi_{k+1}, \ldots, \varphi_{k+n-1} \tag{33}$$

is regular if and only if we have

$$a_{i_1} + a_{i_2} + \cdots + a_{i_k} \neq 0, \tag{34}$$

for all $1 \leq i_1 < i_2 < \cdots < i_k \leq n$.

We intend to derive the consequences of this assumption in the theory of q-harmonics. First, we simply reformulated every statement modulo the substitution of variables

$$\begin{aligned} (a_1, a_2, \ldots, a_n) &\mapsto \mathbf{x} = (x_1, x_2, \ldots, x_n), \\ \mathbf{x} &\mapsto \xi = (\xi_1, \xi_2, \ldots, \xi_n), \end{aligned}$$

and we now have

$$\Phi_n^k = (\varphi_k(\xi), \varphi_{k+1}(\xi), \ldots, \varphi_{k+n-1}(\xi))_{\mathbb{F}_{\mathbf{x}}[\xi]}.$$

This given, from Theorem (10) we can derive the following facts about the ring

$$\mathbb{F}_{\mathbf{x}}[\xi_1, \xi_2, \ldots, \xi_n],$$

where now, $\mathbb{F}_{\mathbf{x}}$ denotes the field of rational functions in $\mathbf{x}$ with coefficients in $\mathbb{F}$.

Theorem 12 *Let*

$$u_1(\xi),\ u_2(\xi),\ \ldots,\ u_{(n+1)!}(\xi)$$

be a monomial basis for the quotient

$$\mathbb{F}_{\mathbf{x}}[\xi]/\Phi_n^k,$$

and let $\deg(u_i) = d_i$. Then every polynomial $f(\xi) \in \mathbb{F}_{\mathbf{x}}[\xi]$, which is homogeneous of degree d, has a unique expansion of the form

$$f(\xi) = \sum_{i=1}^{(n+1)!} u_i(\xi) \sum_{\sum_k r_k(k+1)=d-d_i} a_{i;\mathbf{r}}(\mathbf{x}) \varphi_1^{r_1}(\xi), \varphi_2^{r_2}(\xi) \cdots \varphi_n^{r_n}(\xi), \tag{35}$$

where the coefficients $a_{i;\mathbf{r}}(\mathbf{x})$ are rational functions of $\mathbf{x}$, for $\mathbf{r} \in \mathbb{N}^n$. In particular if $d > \binom{n+1}{2}$ then

$$f(\xi) \equiv 0 \bmod \Phi_n^k. \tag{36}$$

Let us now denote by $\mathcal{D}(\mathbf{x})$ the algebra of differential operators with coefficients in $\mathbb{F}_{\mathbf{x}}$. Moreover, let $\mathcal{D}_d(\mathbf{x})$ denote the subspace of $\mathcal{D}(\mathbf{x})$ consisting of operators of order d. More precisely we have $D \in \mathcal{D}_d(\mathbf{x})$ if and only if D may be expanded in the form

$$D = \sum_{|\mathbf{r}|\leq d} a_{\mathbf{r}}(\mathbf{x})\, \partial_{\mathbf{x}}^{\mathbf{r}} \tag{37}$$

with coefficients $a_{\mathbf{r}}(\mathbf{x}) \in \mathbb{F}_{\mathbf{x}}$ such that $a_{\mathbf{r}}(\mathbf{x}) \neq 0$ at least once when $|\mathbf{r}| = d$. We are here extending our vectorial notation to operators, so that

$$\partial_{\mathbf{x}}^{\mathbf{r}} = \partial_1^{r_1} \partial_2^{r_2} \cdots \partial_n^{r_n}$$

is an operator of order $|\mathbf{r}| = r_1 + r_2 + \ldots + r_n$. The degree condition in (37) imply that the polynomial

$$\sigma(D) := \sum_{|\mathbf{r}|=d} a_{\mathbf{r}}(\mathbf{x})\, \xi^{\mathbf{r}}.$$

does not identically vanish. We will refer to $\sigma(D)$ as the "*symbol*" of D.

This given, as a corollary of Theorem (10), we obtain the following basic result for Steenrod operators

Theorem 13 *Every operator $D \in \mathcal{D}_d(\mathbf{x})$ has an expansion of the form*

$$D = \sum_{i=1}^{(n+1)!} \sum_{\sum_\ell r_k(k+1)\leq d-d_i} a_{i;\mathbf{r}}(\mathbf{x}) u_i(\partial_{\mathbf{x}}) D_{1;q}^{r_1} D_{2;q}^{r_2} \cdots D_{n;q}^{r_n}$$

where $d_i = \deg(u_i)$ and $a_{i;\mathbf{r}}(\mathbf{x}) \in \mathbb{F}_{\mathbf{x}}$. Note that this holds true for any rational value of q.

We may now establish the main goal of this section.

Theorem 14 *For any value of q the dimension of the space of q-Harmonic polynomials in $\mathbf{x}$ does not exceed $(n+1)!$*

7 Last Considerations

Further computer experiments suggest that we have

Conjecture 3 *The set $\mathcal{D}_n^{\mathbf{a}}$ of common polynomial zeros of the operators*

$$\sum_{i=1}^{n} a_i\, \partial_{x_i}^k \partial_{y_i}^j,$$

for all $k, j \in \mathbb{N}$ such that $k + j > 0$, is of a bigraded space of dimension $(n+1)^{n-1}$, whenever we have $\mathbf{a} = (a_1, \ldots, a_n)$ such that

$$\sum_{k\in K} a_k \neq 0, \tag{38}$$

for all nonempty subsets K of $\{1, \ldots, n\}$.

Another interesting experimental observation concerning the space of common zeros of D_1 and D_2 with general operators

$$
\begin{aligned}
D_1 &:= \sum_{i=1}^{n} a_i\, x_i \partial_i^2 + b_i\, \partial_i, \\
D_2 &:= \sum_{i=1}^{n} c_i\, x_i \partial_i^3 + d_i\, \partial_i^2,
\end{aligned}
$$

is that there seem to be conditions, similar to (38), for which this space is always $n!$-dimensional.

References

[1] F. Bergeron, N. Bergeron, A. Garsia, M. Haiman et G. Tesler, *Lattice Diagram Polynomials and Extended Pieri Rules*, Advances in Mathematics, 1999.

[2] M. Haiman, *Combinatorics, symmetric functions and Hilbert schemes*, In CDM 2002: Current Developments in Mathematics in Honor of Wilfried Schmid & George Lusztig, International Press Books (2003) 39–112.

[3] F. Hivert and N. Thiéry, *Non-commutative deformation of symmetric functions and the integral Steenrod algebra*, in: Invariant theory in all characteristics, volume 35 of CRM Proc. Lecture Notes, pages 91-125, Amer. Math. Soc., Providence, RI, 2004.

[4] I. G. Macdonald, *Symmetric functions and Hall polynomials*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1995, With contributions by A. Zelevinsky, Oxford Science Publications.

[5] R.M.W. Wood, *Problems in the Steenrod algebra*, Bull. London Math. Soc. 30 (1998) 449–517.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 384–395

Words and Noncommutative Invariants of the Hyperoctahedral Group

Anouk Bergeron-Brlek[1]

[1] *York University, Mathematics and Statistics, 4700 Keele Street, Toronto, Ontario, M3J 1P3, Canada*

Abstract. Let $\mathcal{B}_n$ be the hyperoctahedral group acting on a complex vector space $\mathcal{V}$. We present a combinatorial method to decompose the tensor algebra $T(\mathcal{V})$ on $\mathcal{V}$ into simple modules via certain words in a particular Cayley graph of $\mathcal{B}_n$. We then give combinatorial interpretations for the graded dimension and the number of free generators of the subalgebra $T(\mathcal{V})^{\mathcal{B}_n}$ of invariants of $\mathcal{B}_n$, in terms of these words, and make explicit the case of the signed permutation module. To this end, we require a morphism from the Mantaci-Reutenauer algebra onto the algebra of characters due to Bonnafé and Hohlweg.

Résumé. Soit $\mathcal{B}_n$ le groupe hyperoctaédral agissant sur un espace vectoriel complexe $\mathcal{V}$. Nous présentons une méthode combinatoire donnant la décomposition de l'algèbre $T(\mathcal{V})$ des tenseurs sur $\mathcal{V}$ en modules simples via certains mots dans un graphe de Cayley donné. Nous donnons ensuite des interprétations combinatoires pour la dimension graduée et le nombre de générateurs libres de la sous-algèbre $T(\mathcal{V})^{\mathcal{B}_n}$ des invariants de $\mathcal{B}_n$, en termes de ces mots, et explicitons le cas du module de permutation signé. À cette fin, nous utilisons un morphisme entre l'algèbre de Mantaci-Reutenauer et l'algèbre des charactères introduit par Bonnafé et Hohlweg.

Keywords: Tensor algebras, invariants of finite groups, hyperoctahedral group, signed permutation module, Cayley graph, words.

1 Introduction

Let $\mathcal{V}$ be a vector space over the field $\mathbb{C}$ of complex numbers with basis $\{x_1, x_2, \ldots, x_n\}$. Then the tensor algebra

$$T(\mathcal{V}) = \mathbb{C} \oplus \mathcal{V} \oplus \mathcal{V}^{\otimes 2} \oplus \mathcal{V}^{\otimes 3} \oplus \cdots$$

can be identified with the ring $\mathbb{C}\langle \mathbf{x} \rangle$ of polynomials in noncommutative variables $\mathbf{x} = x_1, x_2, \ldots, x_n$, where we use the notation $\mathcal{V}^{\otimes d}$ to represent the d-fold tensor space. Any action of a finite group G on $\mathcal{V}$ can be extended to the tensor algebra and the graded character can be found in terms by what we might identify as an analogue of MacMahon's Master Theorem [10] for the tensor space,

$$\chi_{\mathcal{V}^{\otimes d}}(g) = \mathrm{tr}(M(g))^d = \left[q^d\right] \frac{1}{1 - \mathrm{tr}\, M(g) q},$$

where $\left[q^d\right]$ represents taking the coefficient of q^d in the expression to the right and $M(g)$ is a matrix which represents the action of the group element g on a basis of $\mathcal{V}$. In particular, we consider the *algebra* *of invariants of* G, denoted $T(\mathcal{V})^G$, as the subspace of elements of $T(\mathcal{V})$ which are fixed under the action of G. The analogue of Molien's Theorem [6] for the tensor algebra allows us to calculate the graded dimension of this space

$$\mathcal{P}(T(\mathcal{V})^G) = \sum_{d\geq 0} \dim(\mathcal{V}^{\otimes d})^G q^d = \frac{1}{|G|} \sum_{g\in G} \frac{1}{1-\operatorname{tr} M(g)\, q}\,. \tag{1.1}$$

It is well-known that the algebra of invariants of G is freely generated [7, 8] by an infinite set of generators (except when G is scalar) [6].

These algebraic tools do not clearly show the underlying combinatorial structure of these algebras. Our main goal is to find a combinatorial method to decompose $T(\mathcal{V})$ into simple G-modules. The idea is to associate to a module $\mathcal{V}$ of G, a special subalgebra of the group algebra together with a surjective morphism of algebras into the algebra of characters of G. Then we get as a consequence a combinatorial way to decompose $T(\mathcal{V})$ by counting words generated by a particular Cayley graph of the group G. To compute the graded dimension of $T(\mathcal{V})^G$, it then suffices to look at the multiplicity of the trivial module in $T(\mathcal{V})$. This leads to combinatorial descriptions for the graded dimension and the number of free generators of the algebras of invariants of G, which unifie their interpretations.

At this point, we treated the cases of the cyclic, dihedral and symmetric groups [3]. For the symmetric group, the main bridge to link the words in a particular Cayley graph and the decomposition of the tensor algebra into simple modules is a morphism from the theory of the descent algebra [12, 15]. In order to handle cases beyond those already considered, we must find a relation between the group algebra and the algebra of characters.

We present in this paper the case of the hyperoctahedral group $\mathcal{B}_n$, where the main bridge comes from a surjective morphism from the Mantaci-Reutenauer algebra [11] onto the characters of $\mathcal{B}_n$ due to Bonnafé and Hohlweg [1]. More precisely, we present a combinatorial way to decompose the $\mathcal{B}_n$-module $T(\mathcal{V})$ into simple modules using words in a Cayley graph of $\mathcal{B}_n$ and study the subalgebra $T(\mathcal{V})^{\mathcal{B}_n}$ of invariants. This technique applies to modules that can be realized in the Mantaci-Reutenauer algebra, for example for modules indexed by bipartitions of hook shapes, and we make explicit the case of the signed permutation module $\mathcal{V}_{[n-1],[1]}$. We also give combinatorial descriptions for the graded dimensions and the number of free generators of the algebra $T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n}$ of invariants, using words in a particular Cayley graph of $\mathcal{B}_n$. Finally, we present an application to set partitions, since the dimension of $T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n}$ is also given by the set partitions of at most n even parts [2].

The paper is organized as follows. We recall in Section 2.1 the definition of a Cayley graph, and introduce its weighted version. Section 2.2 fixes some notation about bipartitions and bitableaux. Section 2.3 is dedicated to the hyperoctahedral group and recalls its representation theory. In Section 2.4, we describe the generalized Robinson-Schensted correspondence from [16, 5] needed in the statement of the Main Theorem. The bridge between the words in a particular Cayley graph of $\mathcal{B}_n$ and the decomposition of the tensor algebra is a morphism from the Mantaci-Reutenauer algebra into the character algebra of $\mathcal{B}_n$, which we present in Section 2.5. In Section 3, we prove the Main Theorem which gives a combinatorial way to decompose the tensor algebra on any $\mathcal{B}_n$-module into simple modules. As a consequence, we give in Section 4 a combinatorial way to compute the graded dimension of the space of invariants of $\mathcal{B}_n$, and give a description for its number of free generators as an algebra. We then investigate in Sections 3.1 and 4.1 the case of the signed permutation module and give an application to set partitions in Section 4.2.

2 Preliminairies

2.1 Cayley graph

For our purpose, let us recall the definition of a Cayley graph. Let G be a finite group and let $S \subseteq G$ be a set of group elements. The *Cayley graph* associated with (G, S) is defined as the oriented graph $\Gamma = \Gamma(G, S)$ having one vertex for each element of G and the edges associated with elements in S. Two vertices g_1 and g_2 are joined by a directed edge associated to $s \in S$ if $g_2 = g_1 s$. If the resulting Cayley graph of G is connected, then the set S generates G.

A path along the edges of Γ corresponds to a word in the alphabet S. We denote by S^* the *free monoid* on S, *i.e.* the set of all words in the alphabet S. Naturally, the *length* of a word is the number of its letters. We say that a word *reduces to an element* $g \in G$ in the Cayley graph Γ if it corresponds to a path along the edges from the vertex labelled by the identity to the one labelled by g. Such a word, when simplified with respect to the group relations, corresponds to the reduced word g. We denote by $\mathcal{W}_d(g)$ the set of words of length d which reduce to g. A word w is called a *prefix* of a word u if there exits a word v such that $u = wv$. The prefix is *proper* if v is not the empty word. We say that a word *does not cross the identity* if it has no proper prefix which reduces to the identity.

We also consider *weighted Cayley graphs*, where we associate a weight $\nu(s)$ to each letter $s \in S$. We define the *weight of a word* $w = s_1 s_2 \cdots s_r$ in S^* to be the product of the weights of its letter,

$$\nu(w) = \nu(s_1)\nu(s_2)\cdots\nu(s_r).$$

For sake of simplicity, we use undirected edges to represent bidirectional edges and nonlabelled edges to represent edges of weight one.

Example 2.1 *The Cayley graph of the hyperoctahedral group* $\mathcal{B}_2 = \{12, 21, 1\bar{2}, 2\bar{1}, \bar{1}2, \bar{2}1, \bar{1}\bar{2}, \bar{2}\bar{1}\}$ *of signed permutations of* $\{1, 2\}$ *with generators* $\bar{1}2$ *and* $\bar{2}1$ *of weigth one is represented in Figure 1.*

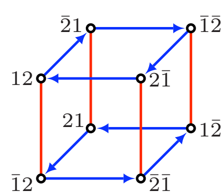


Fig. 1: $\Gamma(\mathcal{B}_2, \{\bar{1}2, \bar{2}1\})$.

2.2 Bipartitions and bitableaux

To fix the notation, we recall some definitions. A *partition* λ of a positive integer n is a decreasing sequence $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_\ell > 0$ of positive integers such that $n = |\lambda| = \lambda_1 + \lambda_2 + \ldots + \lambda_\ell$. We write $\lambda = [\lambda_1, \lambda_2, \ldots \lambda_\ell] \vdash n$. It is natural to represent a partition by its *Young diagram* which is the finite subset $\mathrm{diag}(\lambda) = \{(a, b) \,|\, 0 \leq a \leq \ell - 1 \text{ and } 0 \leq b \leq \lambda_{a+1} - 1\}$ of $\mathbb{N}^2$. Visually, each element of $\mathrm{diag}(\lambda)$ corresponds to the bottom left corner of a box of dimension 1×1 in $\mathbb{N}^2$.

A *bipartition of* n, denoted $\boldsymbol{\lambda} \vdash n$, is a couple $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ of partitions such that $|\boldsymbol{\lambda}| = |\lambda_1| + |\lambda_2| = n$.

Example 2.2 *The bipartitions of 2 are*

$$\square\square, \emptyset \qquad \begin{array}{c}\square\\ \square\end{array}, \emptyset \qquad \square, \square \qquad \emptyset, \begin{array}{c}\square\\ \square\end{array} \qquad \emptyset, \square\square\ .$$

A *tableau* t of shape $\lambda \vdash n$ with values in $T = \{1, 2, \ldots, n\}$ is a function $t : \mathrm{diag}(\lambda) \to T$. We denote by $\mathrm{sh}(t)$ the shape of t. We can visualize it by filling each box c of $\mathrm{diag}(\lambda)$ with the value $t(c)$. A *standard Young tableau* of shape $\lambda \vdash n$ is a tableau with filling $\{1, 2, \ldots, n\}$ and strictly increasing values along each row and each column.

A *bitableau* is a pair $\mathbf{T} = (t_1, t_2)$ of tableaux. The *shape* of a bitableau is the couple $\mathrm{sh}(\mathbf{T}) = (\mathrm{sh}(t_1), \mathrm{sh}(t_2))$. A *standard Young bitableau* is a bitableau $\mathbf{T} = (t_1, t_2)$ where t_1 and t_2 have strictly increasing values along each row and each column, $|\mathrm{sh}(\mathbf{T})| = n$ and the filling of t_1 and t_2 is the set $\{1, 2, \ldots, n\}$. We denote by $\mathrm{SYB}(\boldsymbol{\lambda})$ the set of standard Young bitableaux of shape $\boldsymbol{\lambda}$ and by SYB_n the set of standard Young bitableaux with n boxes.

Example 2.3 *The standard Young bitableaux of shape* $\boldsymbol{\lambda} \vdash 2$ *are*

$$\boxed{1}\boxed{2}, \emptyset \qquad \begin{array}{c}\boxed{2}\\ \boxed{1}\end{array}, \emptyset \qquad \boxed{1}, \boxed{2} \qquad \boxed{2}, \boxed{1} \qquad \emptyset, \begin{array}{c}\boxed{2}\\ \boxed{1}\end{array} \qquad \emptyset, \boxed{1}\boxed{2}\ .$$

2.3 The hyperoctahedral group $\mathcal{B}_n$

Denote by $[n]$ the set $\{1, 2, \ldots, n\}$ and by $\overline{m}$ the integer $-m$. The *hyperoctahedral group* is the group of signed permutations of $[n]$ of order $2^n n!$ which can be seen as the wreath product of the cyclic group of order two $\mathbb{Z}/2\mathbb{Z}$ with the symmetric group $\mathcal{S}_n$ of permutations of $[n]$. We will often represent an element π of $\mathcal{B}_n$ as a word

$$\pi = \pi(1)\pi(2)\cdots\pi(n),$$

where each $\pi(i)$ is an integer whose absolute value is in $[n]$. Note that if we forget the signs in π, we get a permutation of $[n]$. We denote by e the identity element in the hyperoctahedral group.

Example 2.4 $\bar{1}7\bar{6}\bar{5}243$ *is an element of* $\mathcal{B}_7$.

Since the conjugacy classes of $\mathcal{B}_n$ are characterized by bipartitions of n (see [9], Appendix B), it is natural to index the simple modules of $\mathcal{B}_n$ with bipartitions $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$ such that $|\lambda_1| + |\lambda_2| = n$. We denote them by $\mathcal{V}_{\boldsymbol{\lambda}}$ with associated irreducible characters $\chi_{\boldsymbol{\lambda}}$. In particular, $\mathcal{V}_{[n],\emptyset}$ is the trivial module and $\mathcal{V}_{[n-1],[1]}$ the signed permutation module (see Example 4.4). Let us denote by $\mathbb{Z}\mathrm{Irr}(\mathcal{B}_n)$ the *algebra of characters* of $\mathcal{B}_n$.

2.4 Generalized Robinson-Schensted correspondence

The *Robinson-Schensted correspondence* [13, 14] is a bijection between the elements σ of the symmetric group $\mathcal{S}_n$ and pairs $(P(\sigma), Q(\sigma))$ of standard Young tableaux of the same shape. In this section, we present a generalization of this correspondence to the hyperoctahedral group defined as in [16, 5].

Consider the element π of $\mathcal{B}_n$ as a word. Define $\mathbf{P}(\pi)$ to be the standard Young bitableau $(P^+(\pi), P^-(\pi))$ where $P^+(\pi)$ and $P^-(\pi)$ are the insertion tableaux (from the Robinson-Schensted correspondence) of π with respectivley positive and negative letters of π. Similarly, $\mathbf{Q}(\pi) = (Q^+(\pi), Q^-(\pi))$ is the standard Young bitableau where $Q^+(\pi)$ and $Q^-(\pi)$ are the recording tableaux of π for the insertion of respectivley positive and negative letters of π. The map

$$\pi \longleftrightarrow (\mathbf{P}(\pi), \mathbf{Q}(\pi))$$

is a bijection from $\mathcal{B}_n$ onto the set of all pairs of standard Young bitableaux of the same shape. We say that $\mathbf{P}(\pi)$ and $\mathbf{Q}(\pi)$ are respectively the *insertion* and *recording bitableaux* of π.

Example 2.5 *Consider the element* $\bar{1}7\bar{6}\bar{5}243$ *of* $\mathcal{B}_7$. *Then we find*

$$\mathbf{P}(\pi) = \begin{array}{|c|c|}\hline 7 \\ \cline{1-1} 4 \\ \hline 2 & 3 \\ \hline\end{array}\,,\ \begin{array}{|c|c|}\hline 6 \\ \hline 1 & 5 \\ \hline\end{array} \qquad \textit{and} \qquad \mathbf{Q}(\pi) = \begin{array}{|c|c|}\hline 7 \\ \cline{1-1} 5 \\ \hline 2 & 6 \\ \hline\end{array}\,,\ \begin{array}{|c|c|}\hline 4 \\ \hline 1 & 3 \\ \hline\end{array}.$$

2.5 Mantaci-Reutenauer algebra and special morphism

Surprisingly, the key to prove our main result comes from a morphism from the Mantaci-Reutenauer algebra onto the characters of $\mathcal{B}_n$ due to Bonnafé and Hohlweg [4].

A *signed composition of* n, denoted $\mathbf{c} \models n$, is a sequence of nonzero integers $\mathbf{c} = (c_1, c_2, \ldots, c_k)$ such that $|\mathbf{c}| = |c_1| + |c_2| + \cdots + |c_k| = n$. Following Mantaci and Reutenauer [11], we associate to each element $\pi \in \mathcal{B}_n$ a *descent composition* $\mathbf{Des}(\pi)$ constructed by recording the length of the increasing runs (in absolute value) with constant sign, and then recording that sign.

Example 2.6 *The descent composition of* $\bar{1}7\bar{6}\bar{5}243 \in \mathcal{B}_7$ *is* $\mathbf{Des}(\bar{1}.7.\bar{6}.\bar{5}.24.3) = (\bar{1}, 1, \bar{1}, \bar{1}, 2, 1)$.

The *descent composition* $\mathbf{Des}(\mathbf{T})$ of a standard Young bitableau $\mathbf{T} = (t^+, t^-)$ with n boxes is defined in [1] in the following way. First, look for maximal subwords $j\ j+1\ j+2 \cdots k$ of consecutive letters of the word $12 \cdots n$ such that either the numbers $j, j+1, j+2, \ldots, k$ can be read in this order in t^+ when one goes from left to right and top to bottom, or they can be read in t^- in the same manner. The concatenation of these subwords is the word $12 \cdots n$ and the descent composition $\mathbf{Des}(\mathbf{T})$ is the signed composition of n obtained by recording the lengths of these subwords, and the sign of their tableau.

Example 2.7 *Consider the bitableau* $\mathbf{T}= \begin{array}{|c|c|}\hline 7 \\ \cline{1-1} 5 \\ \hline 2 & 6 \\ \hline\end{array}\,,\ \begin{array}{|c|c|}\hline 4 \\ \hline 1 & 3 \\ \hline\end{array}$. *The partition of* 1234567 *in maximal subwords is* $1|2|3|4|56|7$ *hence we can deduce that* $\mathbf{Des}(\mathbf{T}) = (\bar{1}, 1, \bar{1}, \bar{1}, 2, 1)$.

Given a signed composition $\mathbf{c} \models n$, define the element of the group algebra of $\mathcal{B}_n$

$$D_{\mathbf{c}} = \sum_{\substack{\pi \in \mathcal{B}_n \\ \mathbf{Des}(\pi) = \mathbf{c}}} \pi.$$

These elements form a basis of the Mantaci-Reutenauer algebra $\mathcal{MR}_n$, which is a subalgebra of the group algebra of $\mathcal{B}_n$ containing the Solomon's descent algebra of $\mathcal{B}_n$ [11]. Given a standard Young bitableau $\mathbf{T}$ with n boxes, define the element of the group algebra of $\mathcal{B}_n$

$$Z_{\mathbf{T}} = \sum_{\substack{\pi \in \mathcal{B}_n \\ \mathbf{Q}(\pi) = \mathbf{T}}} \pi,$$

where $\mathbf{Q}(\pi)$ corresponds to the recording bitableau resulting from the generalized Robinson-Schensted correspondence. These elements are linearly independent and the space $\mathcal{Q}_n$ that they span is called the *coplactic space*, introduced by Bonnafé and Hohlweg [4]. Note that this space in not an algebra in general. By Lemma 5.7 of [1], the descent composition of an element $\pi \in \mathcal{B}_n$ coincides with the descent composition of its recording bitableau $\mathbf{Q}(\pi)$. Therefore we can rewrite $D_{\mathbf{c}}$ as

$$D_{\mathbf{c}} = \sum_{\substack{\pi \in \mathcal{B}_n \\ \mathbf{Des}(\pi)=\mathbf{c}}} \pi = \sum_{\substack{\mathbf{T} \in \mathrm{SYB}_n \\ \mathbf{Des}(\mathbf{T})=\mathbf{c}}} Z_{\mathbf{T}}, \tag{2.1}$$

hence $\mathcal{MR}_n \subseteq \mathcal{Q}_n$ (see [1] Corollary 5.8). There is a surjective algebra morphism from the Mantaci-Reutenauer algebra onto the character algebra $\Theta : \mathcal{MR}_n \to \mathbb{Z}\mathrm{Irr}(\mathcal{B}_n)$ due to Bonnafé and Hohlweg [4], and a linear map

$$\tilde{\Theta} : \mathcal{Q}_n \to \mathbb{Z}\mathrm{Irr}(\mathcal{B}_n) \tag{2.2}$$

defined by $\tilde{\Theta}(Z_{\mathbf{T}}) = \chi_{\mathrm{sh}(\mathbf{T})}$ such that $\tilde{\Theta}$ restricted to $\mathcal{MR}_n$ corresponds to Θ.

3 Decomposition of $T(\mathcal{V})$ into simple modules

In this section, we develop a combinatorial method to decompose the d-fold tensor of any $\mathcal{B}_n$-module into simple modules. To achieve this, we use the algebra morphism $\Theta : \mathcal{MR}_n \to \mathbb{Z}\mathrm{Irr}(\mathcal{B}_n)$ introduced in Section 2.5 from the Mantaci-Reutenauer algebra onto the algebra of characters of $\mathcal{B}_n$. The next proposition says that the multiplicity of a simple module in the d-fold tensor of any module $\mathcal{V}$ is given as some coefficients in f^d, where f is an element of $\mathcal{MR}_n$ whose image under Θ is the character of $\mathcal{V}$.

Proposition 3.1 *Let $\mathcal{V}$ be a $\mathcal{B}_n$-module such that $\Theta(f) = \chi_{\mathcal{V}}$, for some element f in $\mathcal{MR}_n$. For $\boldsymbol{\lambda} \vdash n$, the multiplicity of $\mathcal{V}_{\boldsymbol{\lambda}}$ in $\mathcal{V}^{\otimes d}$ is equal to*

$$\sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} [Z_{\mathbf{T}}] f^d,$$

where $[Z_{\mathbf{T}}]f^d$ means taking the coefficient of $Z_{\mathbf{T}}$ in f^d.

Proof: By Equation (2.1), we can write f^d as

$$f^d = \sum_{\boldsymbol{\lambda} \vdash n} \sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} c_{\mathbf{T}} Z_{\mathbf{T}}.$$

Applying the linear map (2.2), we get

$$\tilde{\Theta}(f^d) = \sum_{\boldsymbol{\lambda} \vdash n} \sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} c_{\mathbf{T}} \tilde{\Theta}(Z_{\mathbf{T}}) = \sum_{\boldsymbol{\lambda} \vdash n} \sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} c_{\mathbf{T}} \chi_{\boldsymbol{\lambda}}.$$

Since the restriction of $\tilde{\Theta}$ to $\mathcal{MR}_n$ is Θ, we get $\tilde{\Theta}(f^d) = \Theta(f^d) = \Theta(f)^d = {\chi_{\mathcal{V}}}^d$ and thus

$$[\chi_{\boldsymbol{\lambda}}]{\chi_{\mathcal{V}}}^d = \sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} c_{\mathbf{T}} = \sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} [Z_{\mathbf{T}}] f^d.$$

□

The subsequent theorem provide us with an interesting interpretation for the multiplicity of $\mathcal{V}_{\boldsymbol{\lambda}}$ in the d-fold tensor of a $\mathcal{B}_n$-module. This multiplicity is the weighted sum of words in a particular Cayley graph of $\mathcal{B}_n$ which reduce to $\pi_{\mathbf{T}}$, an element of $\mathcal{B}_n$ having recording bitableau $\mathbf{T}$ of shape $\boldsymbol{\lambda}$ (after performing the generalized Robinson-Schensted correspondence). But first, the following key lemma will allow us to link some coefficients of an element of the group algebra to some weighted words in a Cayley graph of G.

Lemma 3.2 ([3]) *Let $\Gamma(G, \{s_1, s_2, \ldots, s_r\})$ be a Cayley graph of G with weights $\nu(s_i) = \nu_i$. Then the coefficient of $\pi \in G$ in the element $(\nu_1 s_1 + \nu_2 s_2 + \cdots + \nu_r s_r)^d$ of the group algebra $\mathbb{C}G$ equals*

$$\sum_{w \in \mathcal{W}_d(\pi)} \nu(w).$$

Before stating the Main Theorem, we need to recall the following. The *support* of an element f of the group algebra of $\mathcal{B}_n$ is defined by $\mathrm{supp}(f) = \{\pi \in \mathcal{B}_n | [\pi]f \neq 0\}$, where $[\pi]f$ is the coefficient of π in f.

Theorem 3.3 *Let $\mathcal{V}$ be a $\mathcal{B}_n$-module such that $\Theta(f) = \chi_{\mathcal{V}}$, for some element f of $\mathcal{MR}_n$, and consider the Cayley graph $\Gamma(\mathcal{B}_n, \mathrm{supp}(f))$ with weights $\nu(\pi) = [\pi](f)$ for each $\pi \in \mathrm{supp}(f)$. For $\boldsymbol{\lambda} \vdash n$, the multiplicity of $\mathcal{V}_{\boldsymbol{\lambda}}$ in $\mathcal{V}^{\otimes d}$ is equal to*

$$\sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} \sum_{w \in \mathcal{W}_d(\pi_{\mathbf{T}})} \nu(w),$$

where $\pi_{\mathbf{T}} \in \mathcal{B}_n$ is such that $\mathbf{Q}(\pi_{\mathbf{T}}) = \mathbf{T}$ and $\mathcal{W}_d(\pi_{\mathbf{T}})$ is the set of words of length d which reduce to $\pi_{\mathbf{T}}$.

Proof: From Proposition 3.1, the multiplicity of $\mathcal{V}_{\boldsymbol{\lambda}}$ in $\mathcal{V}^{\otimes d}$ is

$$\sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} [Z_{\mathbf{T}}] f^d.$$

Since by definition $\pi \in \mathrm{supp}(Z_{\mathbf{T}})$ if and only if π has recording bitableau $\mathbf{T}$, the coefficient of $Z_{\mathbf{T}}$ in f^d is also the coefficient of $\pi_{\mathbf{T}}$ in f^d with $\mathbf{Q}(\pi_{\mathbf{T}}) = \mathbf{T}$ and the result follows from Lemma 3.2. $\square$

3.1 Decomposition of $T(\mathcal{V}_{[n-1],[1]})$ into simple modules

When the hyperoctahedral group $\mathcal{B}_n$ acts as a reflection group on the ring of polynomials in n noncommutative variables, this action corresponds to the signed permutation module $\mathcal{V}_{[n-1],[1]}$. We use the following two corollaries of Proposition 3.1 and Theorem 3.3 respectively, for establishing a connection between the multiplicity of a simple module in ${\mathcal{V}_{[n-1],[1]}}^{\otimes d}$ and words of length d in a particular Cayley graph of $\mathcal{B}_n$. To this end, consider the basis element $D_{(\bar{1},n-1)}$ of the Mantaci-Reutenauer algebra $\mathcal{MR}_n$, which is the sum of all elements of $\mathcal{B}_n$ having descent composition $(\bar{1}, n-1)$. Since

$$\Theta\big(D_{(\bar{1},n-1)}\big) = \tilde{\Theta}\Big(Z_{\boxed{2}\boxed{3}\boxed{4}\boxed{\cdots}\boxed{n}\,,\,\boxed{1}}\Big) = \chi_{[n-1],[1]},$$

we have the following formulas for the multiplicity.

Corollary 3.4 *For $\boldsymbol{\lambda} \vdash n$, the multiplicity of $\mathcal{V}_{\boldsymbol{\lambda}}$ in ${\mathcal{V}_{[n-1],[1]}}^{\otimes d}$ is equal to*

$$\sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} [Z_{\mathbf{T}}] {D_{(\bar{1},n-1)}}^d.$$

Corollary 3.5 *Consider the Cayley graph $\Gamma(\mathcal{B}_n, \operatorname{supp}(D_{(\bar{1},n-1)}))$. For $\boldsymbol{\lambda} \vdash n$, the multiplicity of $\mathcal{V}_{\boldsymbol{\lambda}}$ in ${\mathcal{V}_{[n-1],[1]}}^{\otimes d}$ is equal to*

$$\sum_{\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})} |\mathcal{W}_d(\pi_{\mathbf{T}})|,$$

where $\pi_{\mathbf{T}} \in \mathcal{B}_n$ is such that $\mathbf{Q}(\pi_{\mathbf{T}}) = \mathbf{T}$.

Example 3.6 *Using Corollary 3.4, the $\mathcal{B}_3$-module ${\mathcal{V}_{[2],[1]}}^{\otimes 4}$ decomposes into simple modules as*

$${\mathcal{V}_{[2],[1]}}^{\otimes 4} \cong 4\,\mathcal{V}_{[3],\emptyset} \oplus 7\,\mathcal{V}_{[2,1],\emptyset} \oplus 3\,\mathcal{V}_{[1,1,1],\emptyset} \oplus 10\,\mathcal{V}_{[1],[2]} \oplus 10\,\mathcal{V}_{[1],[1,1]}.$$

Indeed, the element ${D_{(\bar{1},2)}}^4$ of the Mantaci-Reutenauer algebra equals

$$4\,Z_{\boxed{1}\boxed{2}\boxed{3},\emptyset} + 3\,Z_{\begin{array}{l}\boxed{3}\\\boxed{1}\boxed{2}\end{array},\emptyset} + 4\,Z_{\begin{array}{l}\boxed{2}\\\boxed{1}\boxed{3}\end{array},\emptyset} + 3\,Z_{\begin{array}{l}\boxed{3}\\\boxed{2}\\\boxed{1}\end{array},\emptyset} + 5\,Z_{\boxed{3},\boxed{2}\boxed{1}} + 2\,Z_{\boxed{2},\boxed{1}\boxed{3}} + 3\,Z_{\boxed{1},\boxed{2}\boxed{3}} + 5\,Z_{\boxed{3},\begin{array}{l}\boxed{2}\\\boxed{1}\end{array}} + 3\,Z_{\boxed{1},\begin{array}{l}\boxed{3}\\\boxed{2}\end{array}} + 2\,Z_{\boxed{2},\begin{array}{l}\boxed{3}\\\boxed{1}\end{array}}$$

and is sent to

$$4\,\chi_{[3],\emptyset} + 7\,\chi_{[2,1],\emptyset} + 3\,\chi_{[1,1,1],\emptyset} + 10\,\chi_{[1],[2]} + 10\,\chi_{[1],[1,1]}$$

via the map $\tilde{\Theta}$. Table 1 shows how these multiplicities can also be computed using Corollary 3.5 by considering words of length four in the Cayley graph of $\mathcal{B}_3$ with generators $\{\bar{1}23, \bar{2}13, \bar{3}12\}$.

$V_{\boldsymbol{\lambda}}$	$\mathbf{T} \in \mathrm{SYB}(\boldsymbol{\lambda})$	$\pi_{\mathbf{T}} \in \mathcal{B}_3$ $\mathbf{Q}(\pi_{\mathbf{T}}) = \mathbf{T}$	$\mathcal{W}_4(\pi_{\mathbf{T}})$	mult. of $V_{\boldsymbol{\lambda}}$ in ${V_{[2],[1]}}^{\otimes 4}$
$V_{[3],\emptyset}$	$\boxed{1}\boxed{2}\boxed{3}\,,\emptyset$	123	$aaaa$ $abab$ $baba$ $bbbb$	4
$V_{[2,1],\emptyset}$	$\begin{array}{l}\boxed{3}\\\boxed{1}\boxed{2}\end{array},\emptyset$	132	$acab$ $caba$ $cbbb$	
	$\begin{array}{l}\boxed{2}\\\boxed{1}\boxed{3}\end{array},\emptyset$	213	$aaba$ $abbb$ $baaa$ $bbab$	7
$V_{[1,1,1],\emptyset}$	$\begin{array}{l}\boxed{3}\\\boxed{2}\\\boxed{1}\end{array},\emptyset$	321	$baca$ $bcbb$ $ccab$	3
$V_{[1],[2]}$	$\boxed{1}\,,\,\boxed{2}\boxed{3}$	$3\bar{1}\bar{2}$	$bbca$ $bccb$ $cccc$	
	$\boxed{2}\,,\,\boxed{1}\boxed{3}$	$\bar{1}3\bar{2}$	$bcac$ $ccbc$	
	$\boxed{3}\,,\,\boxed{1}\boxed{2}$	$\bar{1}\bar{2}3$	$aabb$ $abba$ $baab$ $bbaa$ $cacc$	10
$V_{[1],[1,1]}$	$\boxed{1}\,,\begin{array}{l}\boxed{3}\\\boxed{2}\end{array}$	$3\bar{2}\bar{1}$	$bcab$ $caca$ $ccbb$	
	$\boxed{2}\,,\begin{array}{l}\boxed{3}\\\boxed{1}\end{array}$	$\bar{2}3\bar{1}$	$acac$ $cbbc$	
	$\boxed{3}\,,\begin{array}{l}\boxed{2}\\\boxed{1}\end{array}$	$\bar{2}\bar{1}3$	$aaab$ $abaa$ $babb$ $bbba$ $cabc$	10

Tab. 1: Decomposition of ${\mathcal{V}_{[2],[1]}}^{\otimes 4}$ using words in $\Gamma(\mathcal{B}_3, \{a, b, c\})$ where $a = \bar{1}23$, $b = \bar{2}13$ and $c = \bar{3}12$.

4 Algebra $T(\mathcal{V})^{\mathcal{B}_n}$ of invariants of $\mathcal{B}_n$

As a consequence of Theorem 3.3, we have a combinatorial interpretation for the graded dimension of the algebra $T(\mathcal{V})^{\mathcal{B}_n}$ of invariants of $\mathcal{B}_n$ in terms of words in a particular Cayley graph of $\mathcal{B}_n$.

Corollary 4.1 *Let $\mathcal{V}$ be a $\mathcal{B}_n$-module such that $\theta(f) = \chi_{\mathcal{V}}$, for some $f \in \mathcal{MR}_n$, and consider the Cayley graph $\Gamma(\mathcal{B}_n, \mathrm{supp}(f))$ with weight $\nu(\pi) = [\pi](f)$ for each $\pi \in \mathrm{supp}(f)$. Then*

$$\dim \left(\mathcal{V}^{\otimes d}\right)^{\mathcal{B}_n} = \sum_{w \in \mathcal{W}_d(e)} \nu(w).$$

Proof: The dimension of the space of invariants of $\mathcal{B}_n$ in $\mathcal{V}^{\otimes d}$ is equal to the multiplicity of the trivial module in $\mathcal{V}^{\otimes d}$. Then the result follows from Theorem 3.3. □

Another interesting result is that the number of free generators of the algebra of invariants of $\mathcal{B}_n$ can be counted by some special words in a particular Cayley graph of $\mathcal{B}_n$. These are the weighted words corresponding to paths which begin and end at the identity vertex, but without crossing the identity vertex.

Proposition 4.2 *Let $\mathcal{V}$ be a $\mathcal{B}_n$-module such that $\theta(f) = \chi_{\mathcal{V}}$, for some $f \in \mathcal{MR}_n$. Then the number of free generators of $T(\mathcal{V})^{\mathcal{B}_n}$ as an algebra are counted by the words which reduce to the identity without crossing the identity in the Cayley graph $\Gamma(\mathcal{B}_n, \mathrm{supp}(f))$ with weight $\nu(\pi) = [\pi](f)$ for each $\pi \in \mathrm{supp}(f)$.*

4.1 Algebra $T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n}$ of invariants of $\mathcal{B}_n$

We have an interpretation for the graded dimension of the space $T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n}$ of invariants of $\mathcal{B}_n$ in terms of paths starting from and ending at the identity vertex in the Cayley graph of $\mathcal{B}_n$ generated by the elements of $\mathcal{B}_n$ having descent composition $(\bar{1}, n-1)$. As a consequence of Corollary 4.1, we can easily compute these dimensions since

$$\Theta\big(D_{(\bar{1},n-1)}\big) = \chi_{[n-1],[1]}.$$

Corollary 4.3 *The dimension of $(\mathcal{V}_{[n-1],[1]}{}^{\otimes d})^{\mathcal{B}_n}$ is equal to the number of words of length d which reduce to the identity in the Cayley graph $\Gamma(\mathcal{B}_n, \mathrm{supp}(D_{(\bar{1},n-1)}))$.*

Example 4.4 *When the group $\mathcal{B}_3$ acts on the polynomial ring $\mathbb{C}\langle x_1, x_2, x_3\rangle$ by $\pi(x_i) = sgn(\pi(i))\, x_{|\pi(i)|}$, the space $\mathbb{C}\langle x_1, x_2, x_3\rangle_4^{\mathcal{B}_3} \cong (\mathcal{V}_{[2],[1]}{}^{\otimes 4})^{\mathcal{B}_3}$ of invariants of $\mathcal{B}_3$ has a monomial basis indexed by the set partitions of $[4]$ with at most 3 parts of even cardinality (see Section 4.2):*

$$\begin{aligned}
\mathbf{m}_{\{1234\}}(x_1,x_2,x_3) &= x_1x_1x_1x_1 + x_2x_2x_2x_2 + x_3x_3x_3x_3,\\
\mathbf{m}_{\{12,34\}}(x_1,x_2,x_3) &= x_1x_1x_2x_2 + x_1x_1x_3x_3 + x_2x_2x_1x_1 + x_2x_2x_3x_3 + x_3x_3x_1x_1 + x_3x_3x_2x_2,\\
\mathbf{m}_{\{13,24\}}(x_1,x_2,x_3) &= x_1x_2x_1x_2 + x_1x_3x_1x_3 + x_2x_1x_2x_1 + x_2x_3x_2x_3 + x_3x_1x_3x_1 + x_3x_2x_3x_2,\\
\mathbf{m}_{\{14,23\}}(x_1,x_2,x_3) &= x_1x_2x_2x_1 + x_1x_3x_3x_1 + x_2x_1x_1x_2 + x_2x_3x_3x_2 + x_3x_1x_1x_3 + x_3x_2x_2x_3.
\end{aligned}$$

As recorded in Table 2, its cardinality equals the one of the set

$$\{aaaa, abab, baba, bbbb\}$$

of words of length 4 in the letters $a = \bar{1}23$, $b = \bar{2}13$ and $c = \bar{3}12$ which reduce to the identity in the Cayley graph $\Gamma(\mathcal{B}_3, \{a, b, c\})$.

In general, for any module $\mathcal{V}$, the algebra $T(\mathcal{V})^{\mathcal{B}_n}$ of invariants of $\mathcal{B}_n$ is freely generated [6], therefore we have the following relation between its Poincaré series and the generating series $\mathcal{F}(T(\mathcal{V})^{\mathcal{B}_n})$ counting the number of its free generators:

$$\mathcal{P}(T(\mathcal{V})^{\mathcal{B}_n}) = \frac{1}{1 - \mathcal{F}(T(\mathcal{V})^{\mathcal{B}_n})}. \tag{4.1}$$

The next corollary of Proposition 4.2 presents a nice interpretation for the number of these free generators.

Corollary 4.5 *The number of free generators of $T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n}$ as an algebra are counted by the words which reduce to the identity without crossing the identity in the Cayley graph $\Gamma(\mathcal{B}_n, \operatorname{supp}(D_{(\bar{1},n-1)}))$.*

Example 4.6 *The free generators of $T(\mathcal{V}_{[2],[1]})^{\mathcal{B}_3}$ are counted by the number of words which reduce to the identity without crossing the identity in the Cayley graph $\Gamma(\mathcal{B}_3, \{a, b, c\})$ where $a = \bar{1}23$, $b = \bar{2}13$ and $c = \bar{3}12$. They are*

		$abaaab$	$abbabb$	$abbbba$	$baaaba$	
		$baabbb$	$babbab$	$bbaabb$	$bbabba$	
	$abab$	$bbbaab$	$abcabc$	$acacac$	$accbbc$	
aa	$baba$	$bbcacc$	$bcabca$	$bcacca$	$bccbcc$	$\cdots$
	$bbbb$	$baccab$	$cacaca$	$caccbb$	$cbbcac$	
		$cbccbc$	$ccbbca$	$ccbccb$	$cccccc$	

Using relation (4.1) *and the analogue of Molien's Theorem* (1.1)*, the generating series for the number of free generators is given by*

$$\begin{aligned}
\mathcal{F}(T(\mathcal{V}_{[2],[1]})^{\mathcal{B}_3}) &= 1 - \mathcal{P}(T(\mathcal{V}_{[2],[1]})^{\mathcal{B}_3})^{-1} \\
&= 1 - \left(\frac{1}{48}\left\{\frac{1}{(1-3q)} + \frac{15}{(1-q)} + 16 + \frac{15}{(1+q)} + \frac{1}{(1+3q)}\right\}\right)^{-1} \\
&= \frac{q^2 - 6q^4}{1 - 9q^2 + 3q^4},
\end{aligned}$$

with series expansion $q^2 + 3q^4 + 24q^6 + 207q^8 + 1791q^{10} + 15498q^{12} + 134109q^{14} + 1160487q^{16} + \cdots$

4.2 Applications to set partitions

A *set partition of* $[n]$, denoted by $A \vdash [n]$, is a family of disjoint nonempty subsets $A_1, A_2, \ldots, A_k \subseteq [n]$ such that $A_1 \cup A_2 \cup \ldots \cup A_k = [n]$. The subsets A_i are called the *parts* of A. The algebra

$$T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n} \cong \mathbb{C}\langle \mathbf{x} \rangle^{\mathcal{B}_n}$$

corresponds to the space of polynomials in noncommutative variables $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ which are invariant under the action of $\mathcal{B}_n$ defined in Example 4.4. Using the fact that a monomial basis for the space $\mathbb{C}\langle \mathbf{x} \rangle^{\mathcal{B}_n}$ of invariants of $\mathcal{B}_n$ is indexed by the set partitions with at most n parts of even cardinality, a closed formula for the Poincaré series of $T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n}$ has been proved in [2] and is given by

$$\mathcal{P}(T(\mathcal{V}_{[n-1],[1]})^{\mathcal{B}_n}) = 1 + \sum_{k=1}^{n} \frac{1 \cdot 3 \cdot \ldots \cdot (2k-1)q^{2k}}{(1-q^2)\,(1-4\,q^2) \cdots (1-k^2\,q^2)}.$$

The words considered in the Cayley graph of $\mathcal{B}_n$, with generators the elements having descent composition $(\bar{1}, n-1)$, have a different nature to that of set partitions. But from Corollary 4.3, we can show for instance the following result.

Corollary 4.7 *The number of set partitions of* $[2d]$ *into at most* n *parts is the number of words of length* $2d$ *which reduce to the identity in the Cayley graph* $\Gamma(\mathcal{B}_n, \mathrm{supp}(D_{(\bar{1},n-1)}))$.

5 Appendix

The table in this section represents the words of length 2, 3 and 4 which reduce to a specific element in the Cayley graph $\Gamma(\mathcal{B}_3, \{\bar{1}23, \bar{2}13, \bar{3}12\})$.

123	132	231	213	312	321	$23\bar{1}$	$13\bar{2}$	$12\bar{3}$	$2\bar{1}3$	$1\bar{2}3$	$1\bar{3}2$
aa			ba	ca							
									aba bbb	abb bba	acb cba
$aaaa$ $abab$ $baba$ $bbbb$	$acab$ $caba$ $cbbb$	$bcab$ $caca$ $ccbb$	$aaba$ $abbb$ $baaa$ $bbab$	$aaca$ $acbb$ $caaa$ $cbab$	$baca$ $bcbb$ $ccab$						
$\bar{1}23$	$\bar{2}13$	$\bar{3}12$	$32\bar{1}$	$31\bar{2}$	$21\bar{3}$	$3\bar{1}2$	$3\bar{2}1$	$2\bar{3}1$	$\bar{1}32$	$\bar{2}31$	$\bar{3}21$
aaa bab	aab baa	aac caa				aca cbb	bca ccb	bcb cca	cab	cac	bac
$3\bar{1}\bar{2}$	$2\bar{1}\bar{3}$	$1\bar{2}\bar{3}$	$\bar{1}3\bar{2}$	$\bar{1}2\bar{3}$	$\bar{2}1\bar{3}$	$\bar{1}\bar{2}3$	$\bar{1}\bar{3}2$	$\bar{2}\bar{3}1$	$3\bar{2}\bar{1}$	$2\bar{3}\bar{1}$	$1\bar{3}\bar{2}$
						bb	cb	cc			
$bbca$ $bccb$ $cccc$	$bccc$ $cbca$ $cccb$	$accc$ $cbcb$ $ccca$	$bcac$ $ccbc$	$bcbc$ $ccac$	$acbc$ $cbac$	$aabb$ $abba$ $baab$ $bbaa$ $cacc$	$aacb$ $acbc$ $bacc$ $caab$ $cbaa$	$aacc$ $baca$ $bcba$ $caac$ $ccaa$	$bcab$ $caca$ $ccbb$	$abca$ $accb$ $cbcc$	$abcb$ $acca$ $bbcc$
$\bar{2}3\bar{1}$	$\bar{3}2\bar{1}$	$\bar{3}1\bar{2}$	$\bar{2}\bar{1}3$	$\bar{3}\bar{1}2$	$\bar{3}\bar{2}1$	$\bar{3}\bar{2}\bar{1}$	$\bar{2}\bar{3}\bar{1}$	$\bar{1}\bar{3}\bar{2}$	$\bar{3}\bar{1}\bar{2}$	$\bar{2}\bar{1}\bar{3}$	$\bar{1}\bar{2}\bar{3}$
			ab	ac	bc						
						abc	acc	bcc	bbc	cbc	ccc
$acac$ $cbbc$	$abac$ $bbbc$	$abbc$ $bbac$	$aaab$ $abaa$ $babb$ $bbba$ $cabc$	$aaac$ $acaa$ $babc$ $cabb$ $cbba$	$aabc$ $baac$ $bcaa$ $cacb$ $ccba$						

Tab. 2: Words in $\Gamma(\mathcal{B}_3, \{a, b, c\})$, where $a = \bar{1}23$, $b = \bar{2}13$ and $c = \bar{3}12$.

References

[1] P. Baumann and C. Hohlweg. A Solomon descent theory for the wreath products $G \wr S_n$. *Trans. Amer. Math. Soc.*, 360(3):1475–1538 (electronic), 2008.

[2] A. Bergeron-Brlek. Invariants in noncommutative variables of the symmetric and hyperoctahedral Groups. *DMTCS Proceedings* **AJ**, *20th International conference on Formal Power Series and Algebraic Combinatorics, Valparaiso, Chile*, pages 653–654, 2008.

[3] A. Bergeron-Brlek, C. Hohlweg, and M. Zabrocki. Words and polynomial invariants of finite groups in noncommutative variables. 36 pages, 2009. To appear in Annals of Combinatorics.

[4] C. Bonnafé and C. Hohlweg. Generalized descent algebra and construction of irreducible characters of hyperoctahedral groups. *Ann. Inst. Fourier (Grenoble)*, 56(1):131–181, 2006. With an appendix by Pierre Baumann and Hohlweg.

[5] C. Bonnafé and L. Iancu. Left cells in type B_n with unequal parameters. *Represent. Theory*, 7:587–609 (electronic), 2003.

[6] W. Dicks and E. Formanek. Poincaré series and a problem of S. Montgomery. *Linear and Multilinear Algebra*, 12(1):21–30, 1982/83.

[7] V. K. Kharchenko. Algebras of invariants of free algebras. *Algebra i Logika*, 17(4):478–487, 491, 1978.

[8] D. R. Lane. *Free algebras of rank two and their automorphisms*. Ph.D thesis, London, 1976.

[9] I. G. Macdonald. *Symmetric functions and Hall polynomials*. The Clarendon Press Oxford University Press, New York, 1979. Oxford Mathematical Monographs.

[10] P. A. MacMahon. *Combinatory analysis*. Two volumes (bound as one). Chelsea Publishing Co., New York, 1960.

[11] R. Mantaci and C. Reutenauer. A generalization of Solomon's algebra for hyperoctahedral groups and other wreath products. *Comm. Algebra*, 23(1):27–56, 1995.

[12] S. Poirier and C. Reutenauer. Algèbres de Hopf de tableaux. *Ann. Sci. Math. Québec*, 19(1):79–90, 1995.

[13] G. de B. Robinson. On the Representations of the Symmetric Group. *Amer. J. Math.*, 60(3):745–760, 1938.

[14] C. Schensted. Longest increasing and decreasing subsequences. *Canad. J. Math.*, 13:179–191, 1961.

[15] L. Solomon. A Mackey formula in the group ring of a Coxeter group. *J. Algebra*, 41(2):255–264, 1976.

[16] R. P. Stanley. Some aspects of groups acting on finite posets. *J. Combin. Theory Ser. A*, 32(2):132–161, 1982.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 396–407

A unified bijective method for maps: application to two classes with boundaries

Olivier Bernardi[1†] and Éric Fusy[‡2]

[1] *Department of Mathematics, MIT, 77 Massachusetts Avenue, Cambridge MA 02139, USA.*
[2] *LIX, École Polytechnique, 91128 Palaiseau Cedex, France.*

Abstract. Based on a construction of the first author, we present a general bijection between certain decorated plane trees and certain orientations of planar maps with no counterclockwise circuit. Many natural classes of maps (e.g. Eulerian maps, simple triangulations,...) are in bijection with a subset of these orientations, and our construction restricts in a simple way on the subset. This gives a general bijective strategy for classes of maps. As a non-trivial application of our method we give the first bijective proofs for counting (rooted) simple triangulations and quadrangulations with a boundary of arbitrary size, recovering enumeration results found by Brown using Tutte's recursive method.

Résumé. En nous appuyant sur une construction du premier auteur, nous donnons une bijection générale entre certains arbres décorés et certaines orientations de cartes planaires sans cycle direct. De nombreuses classes de cartes (par exemple les eulériennes, les triangulations) sont en bijection avec un sous-ensemble de ces orientations, et notre construction se spécialise de manière simple sur le sous-ensemble. Cela donne un cadre bijectif général pour traiter les familles de cartes. Comme application non-triviale de notre méthode nous donnons les premières preuves bijectives pour l'énumération des triangulations et quadrangulations simples (enracinées) ayant un bord de taille arbitraire, et retrouvons ainsi des formules de comptage trouvées par Brown en utilisant la méthode récursive de Tutte.

Keywords: Triangulation, quadrangulation, maps with boundaries, mobiles, bijection, counting

1 Introduction

The enumeration of planar maps (connected graphs embedded on the sphere) has received a lot of attention since the seminal work of Tutte in the 60's [Tut63]. Tutte's recursive method consists in translating the decomposition of a class of maps (typically obtained by deleting an edge) into a functional equation satisfied by the corresponding generating function. The translation usually requires an additional "catalytic" variable, and the obtained functional equation is solved using the so-called "quadratic method" [GJ83, sec.2.9] or its extensions [BMJ06]. The final result is, for many classes of maps, a strikingly simple counting formula. For instance, the number of (rooted) maps with n edges is $\frac{2\cdot 3^n}{(n+1)(n+2)}\binom{2n}{n}$. Tutte's

† Supported by the French ANR project A3 and the European project ExploreMaps – ERC StG 208471
‡ Supported by the European project ExploreMaps – ERC StG 208471

method has the advantage of being systematic, but is quite technical in the way of solving the equation and does not give a combinatorial understanding of the simple-looking enumerative formulas.

As an alternative method, bijective constructions have been developed to obtain more direct combinatorial proofs of the counting formulas, with nice algorithmic applications (random generation and asymptotically optimal encoding in linear time). The first bijections appeared in [CV81] and later in [Sch98] where direct bijections for several classes of maps are described. Typically bijections are from a class of "decorated" plane trees to a class of maps and operate on trees by progressively closing facial cycles. Even if it has been successfully applied to many classes, e.g. in [PS06, PS03, FPS08, BFG04], the bijective method for maps is up to now not as systematic as Tutte's recursive method, since for each class of maps one has to "guess" the tree family to match with, and one has to specify the construction from trees to maps.

This article contributes to fill this gap. Based on a construction of the first author [Ber07, BC10], we provide in Section 3 a general bijection Φ between a set $\mathcal{D}$ of certain decorated plane trees which we call *mobile*[(i)] and a set $\mathcal{O}$ of certain orientations on planar maps with no counterclockwise circuit. As it turns out, a map class is often in bijection with a subfamily $\mathcal{S}$ of $\mathcal{O}$ on which our construction restricts nicely; typically the orientations in $\mathcal{S}$ are characterized by degree constraints which can be traced through our construction and yields a degree characterization of the associated mobiles. The mobiles family is then specifiable by a decomposition grammar and amenable to the Lagrange inversion formula for counting. To summarize, our method makes the bijective method more systematic, since it consists in specializing a "master bijection" Φ to the class of maps under consideration. The problem of enumerating a class of map $\mathcal{M}$ therefore reduces to guessing a family of "canonical" orientations (in $\mathcal{O}$) for $\mathcal{M}$ instead of guessing a family of trees to match with $\mathcal{M}$ (the first task being often simpler than the second).

We focus here, in Section 4 and Section 5 respectively, on two classes that were not completely covered before, namely simple triangulations and simple quadrangulations with a polygonal boundary and a root-corner incident to the boundary. We show bijectively that the number $t_{n,k}$ of rooted simple triangulations with $n+k$ vertices and boundary of length k and the number $q_{n,k}$ of rooted simple quadrangulations with $n+2k$ vertices and boundary of length $2k$ satisfy

$$t_n^{(k)} = \frac{2(2k-3)!}{(k-1)!(k-3)!}\frac{(4n+2k-5)!}{n!(3n+2k-3)!}, \qquad q_n^{(k)} = \frac{3(3k-2)!}{(k-2)!(2k-1)!}\frac{(3n+3k-4)!}{n!(2n+3k-2)!},$$

recovering results found by Brown respectively in [Bro64] and [Bro65] using Tutte's recursive method. The case without boundaries ($k=3$ for triangulations, $k=2$ for quadrangulations) have already received bijective proofs in [PS06, FPS08] (for triangulations) and [Fus07, Sch98] (for quadrangulations); our construction actually coincides with [FPS08, Theo.4.10] for triangulations and with [Sch98, Sec.2.3.3] for quadrangulations. The case of triangulations with boundaries has also received a partial bijective interpretation, different from ours, in [PS06] (only one direction is given, from trees to maps, which by injection shows that $t_{n,k}$ is *at least* the number above, but does not suffice to prove equality).

2 Maps and orientations

Maps. A (planar) *map* is a connected planar graph embedded in the oriented sphere and considered up to continuous deformation. A map is *simple* if it has no loop nor multiple edge. The *faces* are the connected components of the complementary of the graph. A *plane tree* is a map with a unique face. Cutting an

(i) The term *mobile* is borrowed from a bijection by Bouttier et al. [BFG04] which can be seen as a specialization of Φ.

edge e at its middle point gives two half-edges, each incident to an endpoint of e (they are both incident to the same vertex if e is a loop). We shall also consider some maps decorated with dangling half-edges called *stems* (see e.g. Figure 2(a)). A *corner* is the angular section between two consecutive half-edges around a vertex. The degree of a vertex or face is the number of incident corners. A map is a *triangulation* (resp. *quadrangulation*) if every face has degree 3 (resp. 4).

A map is said to be *vertex-rooted* if a vertex is marked, *face-rooted* if a face is marked, and *corner-rooted* if a corner is marked[ii]. The marked vertex, face or corner are called the *root-vertex*, *root-face* or *root-corner*. For a corner-rooted map, the marked corner is indicated by a dangling half-edge pointing to that corner; see Figure 1. A corner-rooted map is said to *induce* the vertex-rooted map (resp. face-rooted map) obtained by keeping the root-vertex (resp. root-face) as marked, but otherwise forgetting the root-corner. Given a face-rooted (or corner-rooted) map, vertices and edges are said to be *outer* or *inner* depending on whether they are incident to the root-face or not.

Orientations. An *orientation* O of a map M is the choice of a direction for each edge of M. A *circuit* is a directed cycle. A *source* is a vertex incident to no ingoing edge. If M is face-rooted (resp. vertex-rooted, corner-rooted), then the pair (M, O) is called a *face-rooted orientation* (resp. *vertex-rooted orientation*, *corner-rooted orientation*). A corner-rooted orientation naturally *induces* a face-rooted orientation and a vertex-rooted orientation. For a vertex v of M, the indegree $\mathrm{in}(v)$ is the number of edges going into v; the outdegree $\mathrm{out}(v)$ is the number of edges going out of v. For a face $f \in M$, the clockwise-degree $\mathrm{cw}(f)$ is the number of edges incident to f that have f on their right; the counterclockwise-degree $\mathrm{ccw}(f)$ is the number of edges that have f on their left. For corner-rooted maps, the half-edge indicating the root-corner increases by 1 the indegree of the root-vertex and the clockwise-degree of the root-face.

A vertex-rooted orientation is said to be *accessible* if every vertex is accessible from the root-vertex by a directed path; it is *source-accessible* if in addition the root-vertex is a source. A circuit of a face-rooted (or corner-rooted) orientation is said *clockwise* if the root-face is on its left. The orientation is *minimal* if every circuit is clockwise; it is *clockwise-minimal* if in addition the root-face is a (clockwise) circuit. We extend the definition of accessibility to (face-rooted) clockwise-minimal orientations O by calling O *accessible* if it is accessible from one of the vertices incident to the root-face. Observe that O is in fact accessible from any vertex on the root-face in this case Similarly, we call a source-accessible orientation O *minimal* if O is minimal for one of the faces incident to the root-vertex. Observe that O is in fact minimal for every face incident to the root-vertex in this case.

Let d be a positive integer. We denote by $\mathcal{S}_d$ the set of source-accessible minimal orientations such that the root-vertex has degree d. We denote by $\mathcal{O}_d$ the set of clockwise-minimal accessible orientations such that the root-face has degree d. We denote by $\widetilde{\mathcal{S}}_d$ the subset of $\mathcal{S}_d$ such that every face incident to the root-vertex has clockwise degree 1. We denote by $\widetilde{\mathcal{O}}_d$ the subset of $\mathcal{O}_d$ such that every vertex incident to the root-face has indegree 1.

Given a map M with vertex-set V and given a function $\alpha : V \mapsto \mathbb{N}$, an *$\alpha$-orientation* is an orientation of M such that $\mathrm{in}(v) = \alpha(v)$ for each $v \in V$. The following result is well-known [Fel04]:

Lemma 1 *If a face-rooted map M has an α-orientation, then M has a unique minimal α-orientation.*

Duality. The *dual* M^* of a map M is the map obtained by the following two step process; see Figure 1.

[ii] Corner-rooted map are usually simply called *rooted maps* in the literature. A face-rooted map can be thought as a *plane map* (a connected graph embedded in the plane) by thinking of the root-face as the infinite face.

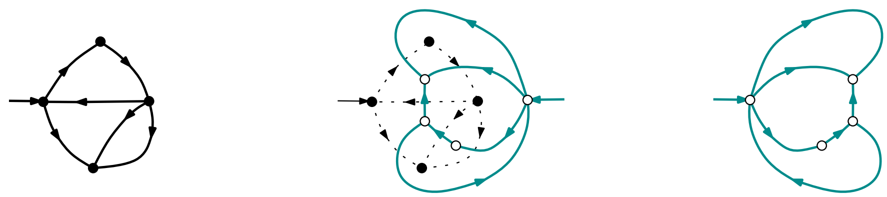


Fig. 1: The dual of an oriented map.

1. In each face f of M, draw a vertex v_f of M^*. For each edge e of M separating faces f and f' (which can be equal), draw the *dual edge* e^* of M^* going from v_f to $v_{f'}$ across e.
2. Flip the drawing of M^*, that is, inverse the orientation of the sphere.

The dual of a face-rooted map is a vertex-rooted map. Corners of a map and of its dual are in natural correspondence (they face each other); this gives the way of defining the root-corner of the dual of a corner-rooted map; see Figure 1. Duality is involutive on maps and rooted maps.

The *dual* of an orientation of M is the orientation of the dual map M^* obtained by applying the following rule at step 1: the dual-edge e^* of an edge $e \in M$ is oriented from the left of e to the right of e. Observe that duality is an involution for oriented map (this is the motivation for step 2 in the definition of duality). The clockwise degree (resp. counterclockwise degree) of a face f of M is equal to the indegree (resp. outdegree) of the vertex v_f of M^* (this is true also with the special convention applying to corner-rooted maps). Also one easily checks that minimality is equivalent to accessibility in the dual:

Lemma 2 *A face-rooted orientation is minimal (resp. clockwise-minimal) if and only if the dual vertex-rooted orientation is accessible (resp. source-accessible).*

Observe that duality maps the set of orientations $\mathcal{S}_d$ (resp. $\widetilde{\mathcal{S}}_d$) to the set $\mathcal{O}_d$ (resp. $\widetilde{\mathcal{O}}_d$). Also, minimal accessible orientations (of corner-rooted maps) are self dual. We mention that these orientations, which play an important role below, are in bijection with spanning trees [Ber07].

3 Bijections between mobiles and orientations

In this section, we first recall a bijection Φ originally due to the first author [Ber07].We then present some extensions of Φ which will be convenient for our subsequent goals. Indeed, in the next two sections we will show how to use these extensions in order to count several families of maps.

The bijection Φ maps minimal accessible (corner-rooted) orientations with n edges and pairs of corner-rooted plane trees (B, T) with $n+1$ and n edges respectively. The tree B is called the (rooted) *mobile* and its vertices are *bicolored* in black and white (in such a way that edges always connect a black and a white vertex). Informally, the bijection Φ consists in folding the tree T (oriented from the root to leaves) around the mobile. More precisely, one glues the vertices of T on the black corners of the mobile and then erases the edges and white vertices of B (leaving the edges of T as edges of a minimal accessible orientation). In what follows we adopt a slightly different presentation, in which the tree T only appears implicitly in certain *decorations* added to the mobile B.

A *decorated mobile* is a bicolored (unrooted) plane tree with *outgoing stems* (dangling outgoing half-edges) possibly attached to each black corner; see Figure 2(a). The *excess* of a decorated mobile is the number of edges minus the number of (outgoing) stems. A mobile with excess δ is called a *δ-mobile*. A *fully decorated* mobile is obtained from a decorated mobile by inserting an *ingoing stem* (dangling ingoing

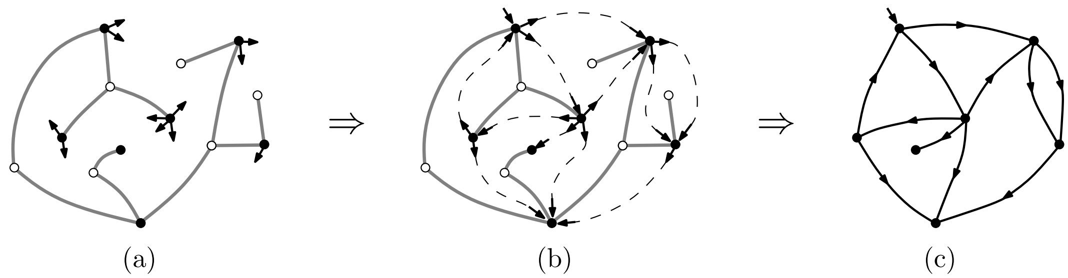


Fig. 2: The rooted closure of a mobile of excess $\delta = 1$.

half-edge) in each black corner following an edge of the mobile (and not a stem) in clockwise order around the vertex; the fully decorated mobile is represented in solid lines in Figure 2(b). The *degree* $deg(v)$ of a vertex v of a decorated mobile is the total number of incident half-edges (including the outgoing stems). For a black vertex b the indegree $in(b)$ and out-degree $out(b)$ are respectively the number of incident ingoing and outgoing stems incident to b in the fully-decorated mobile (so $deg(b) = in(b) + out(b)$).

3.1 Bijection between 1-mobiles and minimal accessible orientations

We now recall the bijection given in [Ber07] between 1-mobiles and minimal accessible orientations.

Closure. Let D be a decorated mobile with p edges and q outgoing stems (hence excess $\delta = p - q$). The corresponding fully decorated mobile D' has p ingoing and q outgoing stems. A clockwise walk around D' (with the face area on the left of the walker) sees a succession of outgoing stems and ingoing stems. Associating an opening parenthesis to outgoing stems and a closing parenthesis to ingoing stems, one obtains thus a cyclic binary word with q opening and p closing parentheses. This yields in turn a matching of outgoing stems with ingoing stems, leaving $|\delta|$ stems unmatched, which are ingoing if $\delta > 0$ and outgoing if $\delta < 0$; see Figure 2. The *partial closure* C of the decorated mobile D is obtained by forming a directed edge out of each matched pair, see Figure 2(a)-(b). We consider C as a planar map with two types of edges (those of the mobile, which are undirected, and the new formed edges, which are directed) and $|\delta|$ stems. Note that, if $\delta \geq 0$, there are δ white corners incident to the root-face of C, because initially the number of such corners is equal to the number of edges, and then each matched pair of stems decreases this number by 1. These corners, which stay incident to the root-face throughout the partial closure, are called *exposed white corners*.

The *rooted-closure* of the decorated mobile D is obtained from the partial closure C by erasing every white vertex and edge of the mobile (this might result in a disconnected embedded graph in general).

Opening. Let M be an oriented map (rooted or not) with vertex set V and face set F. The *partial opening* of M is the map C with two types of vertices (black vertices in V and white vertices in $W = \{w_f,\ f \in F\}$) and two types of edges (directed and undirected) obtained as follows.

- Insert a white vertex w_f inside each face f of M.
- Draw an undirected edge between w_f and each corner incident to f which precedes an ingoing half-edge in clockwise order around its incident vertex. If M is corner-rooted, then the stem indicating the root-corner is interpreted as an ingoing half-edge and gives rise to an edge of C.

If M is a corner-rooted orientation, the *rooted-opening* of M is obtained from the partial opening C by

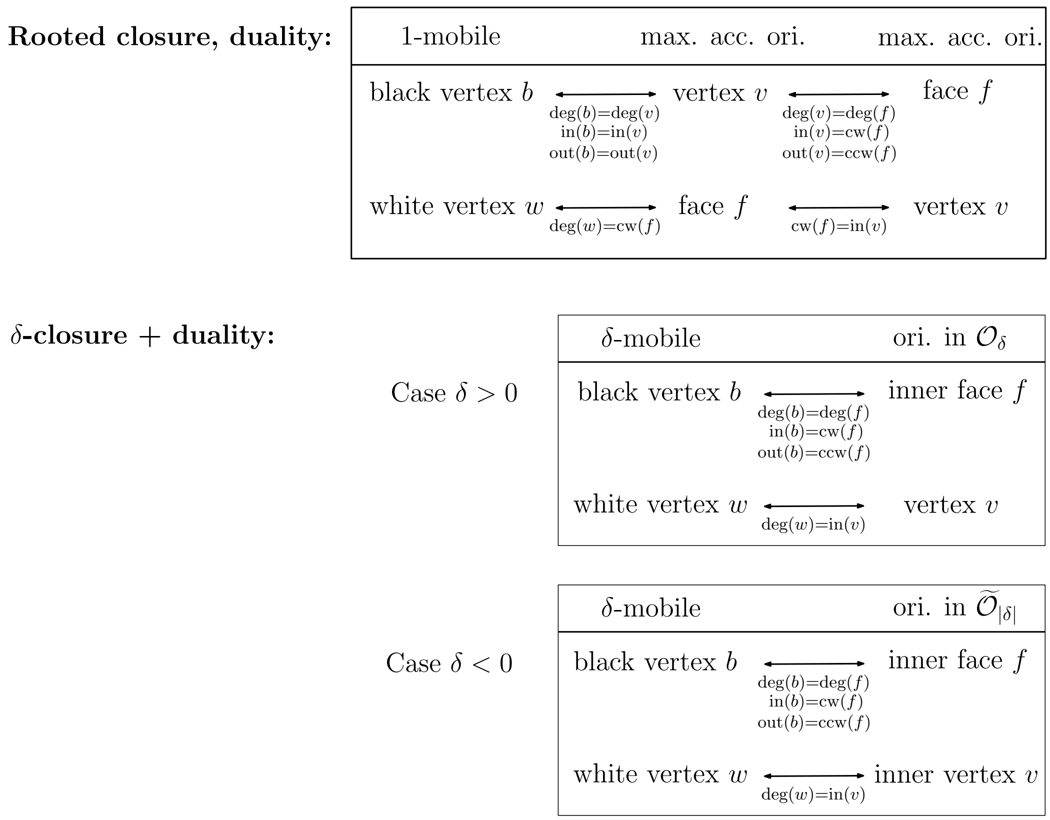


Fig. 3: The closure-bijections, with the parameter correspondences.

erasing all the ingoing half-edges of M, thereby creating an undirected embedded bicolored graph with some outgoing stems incident to black corners.

We recall the result from [Ber07] (see also [BC10]) that we shall generalize.

Theorem 3 *The rooted closure is a bijection between decorated mobiles of excess $\delta = 1$ and (corner-rooted) minimal accessible orientations. The rooted opening is the inverse mapping. Lastly, the parameter-correspondence is shown in Figure 3, top-part.*

3.2 Bijection for δ-mobiles

δ-closure. We now define the δ-closure of a δ-mobile (the definition depends on the sign of δ). Let D be a δ-mobile and let C be the partial closure of D. The *δ-closure* M of D is defined as follows.

- If $\delta > 0$, then C has δ ingoing stems (incident to the root-face). The vertex-rooted orientation M is obtained from C by first creating a root-vertex v of M in the root-face of C and connecting it to each ingoing stem (stems thus become part of an edge of M directed away from v); second erasing the edges and white vertices of the mobile.

- If $\delta < 0$, then C has δ outgoing stems (incident to the root-face). The vertex-rooted orientation M is obtained from C by first creating a root-vertex v of M in the root-face of C and connecting it to each outgoing stem and then reorienting these edges (stems thus become part of an edge of M directed away from v); second erasing the edges and white vertices of the mobile.
- If $\delta = 0$, then M is the face-rooted orientation obtained from C by erasing the edges and white vertices of the mobile.

Actually, it is not obvious from our definitions that the δ-closures give *connected* orientations but we prove this and more below.

Theorem 4 *Let δ be in $\mathbb{Z}$.*

- *For $\delta > 0$, the δ-closure is a bijection between δ-mobiles and the set $\mathcal{S}_\delta$, which (by duality) is itself in bijection with the set $\mathcal{O}_\delta$. The parameter-correspondence is shown in Figure 3 bottom-part.*
- *For $\delta < 0$, the δ-closure is a bijection between δ-mobiles and the subset $\widetilde{\mathcal{S}}_{|\delta|} \subset \mathcal{S}_{|\delta|}$, which (by duality) is itself in bijection with the subset $\widetilde{\mathcal{O}}_{|\delta|} \subset \mathcal{O}_{|\delta|}$ The parameter-correspondence is shown in Figure 3 bottom-part.*
- *For $\delta = 0$, the δ-closure is a bijection between δ-mobiles and minimal orientations.*

The remaining of this section is devoted to the proof of Theorem 4 (the proof for $\delta = 0$, which is similar, is omitted since we will not use it in this article).

Case $\delta > 0$. We first prove that the δ-closure of a δ-mobile is in $\mathcal{S}_\delta$. Let D be a δ-mobile, let C be its partial closure and let M be its δ-closure. As observed above, the mobile D has $\delta > 0$ exposed white corners. Let D' be the decorated mobile obtained from D by creating a new black vertex b, joining b to an exposed white corner, and adding δ outgoing stems to b. The excess of D' is 1, hence by Theorem 3 the rooted closure of D' gives a minimal accessible orientation M'. Moreover, it is easily seen (Figure 4) that the root-corner of M' is incident to the new vertex b (because the ingoing stem incident to b is not matched during the partial closure). Moreover (provided the ingoing root half-edge is not counted) b is a source of the orientation M', and the vertex-rooted orientation M is induced by the corner-rooted orientation M'. Thus, the orientation M is in $\mathcal{S}_\delta$.

The following comment will be useful later (for the case $\delta < 0$): the closure M of D is in $\widetilde{\mathcal{S}}_\delta$ if and only if each of the exposed white corners of D is incident to a (white) leaf of D. Indeed, a white vertex w_f of D has an exposed white corner if and only if it corresponds to a face f of M incident to the root-vertex b. Moreover, the clockwise degree of f is (as always) the degree of w_f.

We now prove that the δ-closure is a bijection by defining the inverse mapping. Let M be a vertex-rooted orientation in $\mathcal{S}_\delta$. By applying the partial opening of M and then erasing every ingoing half-edge of M, one obtains an embedded graph with stems $\widehat{D}$. The embedded graph $\widehat{D}$ is in fact disconnected since the root-vertex b of M is incident to no edge of $\widehat{D}$ (since b is a source of M). The δ-*opening* D of M is obtained from $\widehat{D}$ by erasing the vertex b. In order to prove that D is a decorated mobile (i.e. a tree with stems), we consider a minimal accessible orientation M' obtained from M by choosing a root-corner for M among the corners incident to the root-vertex b. By Theorem 3, the rooted opening of M' gives a decorated mobile D'. Clearly, D is obtained from D' by erasing the black vertex b. Moreover, b is a leaf of D' (since b is incident to no ingoing half-edge except the stem indicating the root-corner of O), hence D is a mobile, and it has excess δ. Lastly, since the rooted closure and rooted opening are inverse mappings, it is clear that δ-closure and δ-opening are inverse mappings, hence bijections.

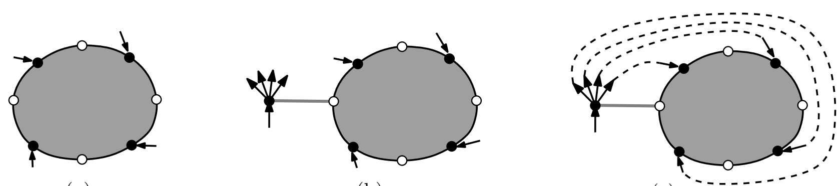


Fig. 4: Formulation of the δ-closure, for $\delta > 0$, as a reduction to the rooted closure. Figure (a) shows generically the partial closure of a δ-mobile with $\delta = 4$, in (b) one creates a black vertex b with δ outgoing stems, and connects it to an exposed white corner, in (c) one performs the remaining matchings of stems to complete the δ-closure.

Case $\delta < 0$**.** We denote $d = -\delta$. Let D be a δ-mobile. We associate to D a d-mobile $\phi(D)$ obtained from D by transforming each of its d unmatched outgoing stems into an edge of $\phi(D)$ connected to a newly created white leaf. Observe that the δ-closure of D and the d-closure of $\phi(D)$ coincide. Hence the δ-closure is the composition of the mapping ϕ and of the d-closure. Moreover, the mapping ϕ is a bijection between the set of δ-mobiles and the set $\mathcal{D}_d$ of d-mobiles such that every exposed white corner belongs to a leaf. Indeed, $\phi(D)$ belongs to $\mathcal{D}_d$ since the unique incident corner for each of the d newly created white leaves remains exposed during the partial closure; and the inverse mapping ϕ^{-1} is obtained by replacing each edge incident to an exposed leaf by an outgoing stem. Lastly, by the observations above (case $\delta > 0$), the d-closure induces a bijection between the set $\mathcal{D}_d$ and the set $\widetilde{\mathcal{S}}_d$. The inverse mapping to the δ-closure, called the δ-opening, is obtained as the composition of ϕ^{-1} with the d-opening. This completes the proof of Theorem 4 (in the cases $\delta \neq 0$).

4 Bijective counting of triangulations with boundaries

In this section we obtain bijections for simple triangulations (a.k.a. *3-connected triangulations*, *maximal planar graphs*) and for triangulations with boundaries. The bijections are obtained by specializing the closures defined in the previous section to certain classes of orientations characterizing simple triangulations.

Let T be a face-rooted triangulation. A *3-orientation* of T is an orientation such that inner vertices have indegree 3 and outer vertices have indegree 1. Schnyder proved in [Sch89] that any simple face-rooted triangulation admits a 3-orientation, that any 3-orientation is accessible from the outer vertices and that the root-face is always directed. Moreover, one easily checks (using Euler's relation) that loops and double edges are obstructions to the existence of a 3-orientation. Thus, a planar triangulation admits a 3-orientation if and only if it is simple. In the following we simply call *3-orientation* a 3-orientation of a face-rooted triangulation. From Lemma 1 one obtains:

Lemma 5 *Face-rooted simple triangulations are in bijection with minimal 3-orientations. Such orientations are clockwise-minimal and accessible.*

Minimal 3-orientations are the orientations in $\widetilde{\mathcal{O}}_3$ such that all inner vertices have indegree 3 and all faces have degree 3. Thus, we can use the case $\delta = -3$ of Theorem 4 to conclude that face-rooted simple triangulations are in bijection with (-3)-mobiles having every vertex of degree 3 (recall that outgoing stems count in the degree of a black vertex). In fact, the constraint that the excess is -3 can be omitted, since it is a consequence of all vertices having degree 3 (as easily seen by induction on the number of vertices). Call *trivalent* the decorated mobiles with all vertices of degree 3. We obtain:

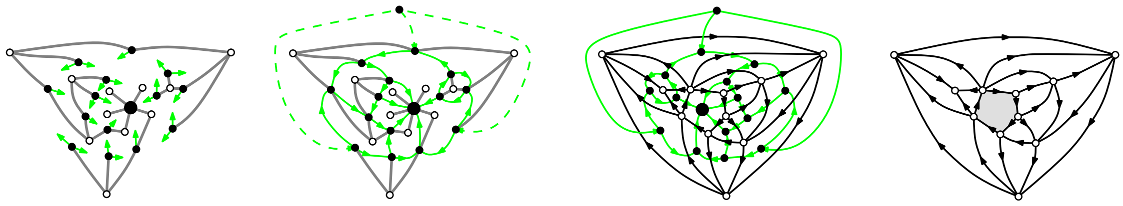


Fig. 5: (a) A mobile in $\mathcal{B}_k$. (b) The 3-closure. (c) Duality. (d) The non-separated k-annular triangulation (with its minimal pseudo 3-orientation).

Proposition 6 (Recovering [FPS08]) *The δ-closure, case $\delta = -3$ (together with duality) induces a bijection between face-rooted triangulations with $n+3$ vertices and trivalent mobiles with n white vertices.*

Proposition 7 (Counting rooted simple triangulations) *For $n \geq 0$, let t_n be the number of corner-rooted simple triangulations with $n+3$ vertices. The generating function $T(x) = \sum_{n\geq 0}(2n+1)t_n x^n$ satisfies*

$$T(x) = u^3, \ \text{ where } u = 1 + x\, u^4.$$

Consequently, the Lagrange inversion formula gives: $t_n = 2\dfrac{(4n+1)!}{(n+1)!(3n+2)!}$.

Proof: The Euler relation easily implies that a triangulation with $n+3$ vertices has $2n+1$ non-root faces. Hence $(2n+1)t_n$ is the cardinality of the set $\mathcal{H}_n$ of face-rooted triangulations with n inner vertices having an additional marked corner c not incident to the root-face (think of obtaining this map by first marking a corner and then a face). Marking the corner c is equivalent to marking a black corner of the associated mobile (since the black vertices are in correspondence to the triangular faces). In other words, $\mathcal{H}_n$ is in bijection, via the (-3)-closure, with trivalent mobiles that have n white vertices and a marked black corner, and $T(x)$ is the generating function of this class of mobiles. Finally, the expression of $T(x)$ above is just the translation of a recursive decomposition for the mobiles (details are omitted here). $\square$

We now proceed to count bijectively the triangulations with boundaries. In the following, k is an integer greater than 3. A *k-gonal triangulation* is a map having one face of degree k whose contour is *simple* (incident to k distinct vertices) and all other faces of degree 3. The k-gonal face is called *boundary face*, and the vertices are called *boundary* or *non-boundary* depending on whether they are incident to the boundary face. A *pseudo 3-orientation* of a k-gonal triangulation is an orientation such that all non-boundary vertices have indegree 3, and the boundary face is directed. A pseudo 3-orientation of a k-gonal triangulation is shown in Figure 5(d). By the Euler relation, a k-gonal triangulation with n non-boundary vertices has $3n+2k-3$ edges. Hence, the sum of indegrees of the boundary vertices is $2k-3$. A *k-annular triangulation* is a face-rooted k-gonal triangulation whose root-face is not the boundary face; see Figure 5(d). Let T be a simple k-annular triangulation, with root-face f and boundary face f'. A 3-cycle C of T is called *separating* if C is different from (the contour of) the root-face and has f on one side and f' on the other side; T is said to be *non-separated* if it has no separating 3-cycle. We denote by $\mathcal{T}_k$ the set of k-annular triangulations, and by $\mathcal{N}_k$ the subset of non-separated ones.

Lemma 8 *A k-annular triangulation A admits a pseudo 3-orientation if and only if it is simple. If A is simple, it admits a unique minimal pseudo 3-orientation. This orientation is accessible from all outer* *vertices if and only if A is non-separated. Moreover, the root-face is directed (in clockwise direction) in this case.*

To summarize, the set $\mathcal{N}_k$ of non-separated k-annular triangulations is in bijection with the set of clockwise-accessible minimal pseudo 3-orientations.

The proof of the lemma, which is ommited, essentially relies on Lemma 1 and on the Euler relation.

By definition, the clockwise-accessible minimal pseudo 3-orientations are the orientations in $\mathcal{O}_3$ such that every vertex has indegree 3 and every face has degree 3 except for one *boundary* face b which has degree k and is directed clockwise. Moreover, a counting argument (using the Euler relation) shows that the indegrees of the boundary vertices must add up to $2k-3$. Let D be the 3-mobile giving an orientation $O \in \mathcal{N}_k$ (by the 3-closure followed by duality) and let v_b be the black vertex of D corresponding to the boundary face b. Since b is counterclockwise, v_b has no outgoing stem and has k white neighbors, which clearly (by definition of closure) corresponds to the boundary vertices; see Figure 5. Hence, the degree of these white vertices of the mobile must add up to $2k-3$. Lastly, as in the case of triangulations without boundary, the condition of the excess being 3 is implied by the degree conditions on black and white vertices, so can be omitted. To conclude, by specialization of the 3-closure, Lemma 8 translates into:

Theorem 9 *The family $\mathcal{N}_k$ of non-separated k-annular triangulations is in bijection with the family $\mathcal{B}_k$ of decorated mobiles having every vertex of degree 3 except for one black vertex b of degree k carrying no outgoing stem and such that the degrees of its k (white) neighbors add up to $2k-3$.*

Theorem 10 (Counting rooted k-gonal triangulations) *Let $k > 3$, $n \geq 0$, and let $t_{k,n}$ be the number of simple corner-rooted k-gonal triangulations with $n+k$ vertices having the root-corner in the k-gonal face. The generating function $T_k(x) = \sum_{n\geq 0}(2n+k-2)\, t_{k,n}x^n$ satisfies*

$$T_k(x) = \binom{2k-4}{k-3} u^{2k-3}, \quad \text{where } u = 1 + x\, u^4.$$

Consequently, the Lagrange inversion formula gives: $t_{k,n} = \dfrac{2(2k-3)!}{(k-1)!(k-3)!}\dfrac{(4n+2k-5)!}{n!(3n+2k-3)!}$.

Proof: Let $\vec{\mathcal{T}}_k$ be the set of k-annular triangulations with a marked corner in the boundary face (equivalently, a marked boundary vertex). Let $\vec{\mathcal{N}}_k$ be the subset of these k-annular triangulations that are non-separated. Let also $\vec{\mathcal{T}}$ be the set of corner-rooted simple triangulations with a marked inner face. The separating 3-cycles of a k-annular triangulation are linearly ordered by inclusion of their boundary region (the region which contains the boundary face). Thus, there is a unique decomposition of k-annular triangulations $A \in \vec{\mathcal{T}}_k$ into a pair $(N,T) \in \vec{\mathcal{N}}_k \times \vec{\mathcal{T}}$. This decomposition is a bijection and translates into the generating function equation $T_k(x) = N_k(x)T(x)$, where $T_k(x), N_k(x), T(x)$ are respectively the generating function of the maps in $\vec{\mathcal{T}}_k$, $\vec{\mathcal{N}}_k$, $\vec{\mathcal{T}}$ counted by number of non-boundary vertices. The generating function $\vec{\mathcal{T}}_k(x)$ is $\sum_{n\geq 0}(2n+k-2)\, t_{k,n}x^n$ because maps in $\vec{\mathcal{T}}_k$ with n non-boundary vertices have $2n+k-2$ non-boundary faces. Moreover, by Proposition 7, $T(x) = u^3$, where $u = 1 + xu^4$. It remains to express $N_k(x)$ in terms of u. By Theorem 9 (and the fact that marking a corner in the boundary face accounts to marking a corner incident to the special black vertex in the associated mobile), $N_k(x)$ is the generating function of corner-rooted mobile (counted by number of white vertices) such that the root-vertex is a black corner of degree k whose (white) neighbors have total degree $2k-3$. The white vertices have a total of $k-3$ hanging subtrees, which are trivalent trees whose generating funcion is $v=u^2$ (easy proof omitted). In addition, there are $\binom{2k-4}{k-3}$ ways to distribute the $k-3$ hanging trees on the k white vertices. Hence, $N_k(x)=\binom{2k-4}{k-3}u^{2k-6}$. □

5 Bijective counting of quadrangulations with boundaries

In this section we reiterate the strategy used in previous section to the case of quadrangulations.

We call *2-orientation* a face-rooted orientation, in which faces have degree 4, inner vertices have indegree 2, and outer vertices have indegree 1. De Fraysseix *et al.* [dFOdM01] have shown that any simple face-rooted quadangulation admits a 2-orientation, that any 2-orientation is accessible from the outer vertices and that the root-face is always directed. Moreover, one easily checks that a double edge is an obstruction for 2-orientation. Hence minimal 2-orientations are in bijection with simple face-rooted quadrangulations. By definition, minimal 2-orientations are the orientations in $\widetilde{\mathcal{O}}_4$ having faces of degree 4 and inner vertices of indegree 2. The (-4)-closure (followed by duality) gives a bijection between such orientations and (-4)-mobiles whise white vertices have degree 2 and whose black vertices have degree 4. Call these mobiles *tetravalent*. We obtain:

Proposition 11 (Counting rooted simple quadrangulations) *For $n\geq 0$, let q_n be the number of rooted simple quadrangulations with $n+4$ vertices. Then the generating function $Q(x)=\sum_{n\geq 0}(n+1)q_nx^n$ satisfies*

$$Q(x)=u^4,\ \text{ where } u=1+x\,u^3.$$

Consequently, the Lagrange inversion formula gives: $q_n=2\dfrac{(3n+3)!}{(n+2)!(2n+3)!}$.

Call $2k$*-gonal quadrangulation* a map with faces of degree 4 except for one face of degree $2k>4$. The strategy for counting simple $2k$-gonal quadrangulations parallels the case of triangulations. One first defines a $2k$*-annular quadrangulation* to be a simple quadrangle-rooted $2k$-gonal quadrangulation. One then proves that any such map admits a unique minimal *pseudo 2-orientation* (orientation such that the boundary face is clockwise and non-boundary vertices have indegree 2), and that this orientation is clockwise-accessible if and only if the map is *non-separated* (no 4-cycle separates the root-face from the boundary face). Any $2k$-annular quadrangulation decomposes uniquely into a pair made of a face-rooted quadrangulation with an additional marked face and a non-separated $2k$-annular quadrangulation. Moreover, the later maps are in bijection (via the 4-closure) with a family of pseudo-tetravalent decorated mobiles which is easy to enumerate. We obtain:

Theorem 12 (Counting rooted simple $2k$-gonal quadrangulations) *Let $k>2$, $n\geq 0$, and let $q_{k,n}$ be the number of rooted simple $2k$-gonal quadrangulations with $n+2k$ vertices and with the root-corner incident to the $2k$-gonal face. The generating function $Q_k(x)=\sum_{n\geq 0}(n+k-1)q_{k,n}x^n$ satisfies*

$$Q_k(x)=\binom{3k-3}{k-2}u^{3k-2},\quad \text{where } \ u=1+x\,u^3.$$

Consequently, the Lagrange inversion formula gives: $q_{k,n}=\dfrac{3(3k-2)!}{(k-2)!(2k-1)!}\,\dfrac{(3n+3k-4)!}{n!(2n+3k-2)!}$.

References

[BC10] O. Bernardi and G. Chapuy. A bijection for covered maps, or a shortcut between harer-zagier's and jackson's formulas. arXiv:1001.1592, 2010.

[Ber07] O. Bernardi. Bijective counting of tree-rooted maps and shuffles of parenthesis systems. *Electron. J. Combin.*, 14(1):R9, 2007.

[BFG04] J. Bouttier, P. Di Francesco, and E. Guitter. Planar maps as labeled mobiles. *Electron. J. Combin.*, 11(1):R69, 2004.

[BMJ06] M. Bousquet-Mélou and A. Jehanne. Polynomial equations with one catalytic variable, algebraic series and map enumeration. *J. Combin. Theory Ser. B*, 96(5):623 – 672, 2006.

[Bro64] W.G. Brown. Enumeration of triangulations of the disk,. *Proc. London Math. Soc.*, 14(3):746–768, 1964.

[Bro65] W.G. Brown. Enumeration of quadrangular dissections of the disk. *Canad. J. Math.*, 21:302–317, 1965.

[CV81] R. Cori and B. Vauquelin. Planar maps are well labeled trees. *Canad. J. Math.*, 33(5):1023–1042, 1981.

[dFOdM01] H. de Fraysseix and P. Ossona de Mendez. On topological aspects of orientations. *Discrete Math.*, 229:57–72, 2001.

[Fel04] S. Felsner. Lattice structures from planar graphs. *Electron. J. Combin.*, 11(1), 2004.

[FPS08] É. Fusy, D. Poulalhon, and G. Schaeffer. Dissections, orientations, and trees, with applications to optimal mesh encoding and to random sampling. *Transactions on Algorithms*, 4(2):Art. 19, April 2008.

[Fus07] É. Fusy. *Combinatoire des cartes planaires et applications algorithmiques*. PhD thesis, École Polytechnique, 2007.

[GJ83] I. P. Goulden and D. M. Jackson. *Combinatorial Enumeration*. John Wiley, New York, 1983.

[PS03] D. Poulalhon and G. Schaeffer. A bijection for triangulations of a polygon with interior points and multiple edges. *Theoret. Comput. Sci.*, 307(2):385–401, 2003.

[PS06] D. Poulalhon and G. Schaeffer. Optimal coding and sampling of triangulations. *Algorithmica*, 46(3-4):505–527, 2006.

[Sch89] W. Schnyder. Planar graphs and poset dimension. *Order*, 5(4):323–343, 1989.

[Sch98] G. Schaeffer. *Conjugaison d'arbres et cartes combinatoires aléatoires*. PhD thesis, Univ. Bordeaux I, 1998.

[Tut63] W. T. Tutte. A census of planar maps. *Canad. J. Math.*, 15:249–271, 1963.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 408–419

Combinatorial aspects of Escher tilings

A. Blondin Massé[1,2] and S. Brlek[1,†] and S. Labbé[1,3]

[1] *Laboratoire de Combinatoire et d'Informatique Mathématique - LaCIM - Université du Québec à Montréal Case Postale 8888, succursale Centre-ville Montréal (QC) CANADA H3C 3P8*

[2] *Laboratoire de Mathématiques - LAMA- CNRS UMR 5127, Université de Savoie 73376 Le Bourget du Lac FRANCE*

[3] *Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier - LIRMM - 161 rue Ada, 34392 Montpellier Cedex 5 France*

Abstract. In the late 30's, Maurits Cornelis Escher astonished the artistic world by producing some puzzling drawings. In particular, the tesselations of the plane obtained by using a single tile appear to be a major concern in his work, drawing attention from the mathematical community. Since a tile in the continuous world can be approximated by a path on a sufficiently small square grid - a widely used method in applications using computer displays - the natural combinatorial object that models the tiles is the polyomino. As polyominoes are encoded by paths on a four letter alphabet coding their contours, the use of combinatorics on words for the study of tiling properties becomes relevant. In this paper we present several results, ranging from recognition of these tiles to their generation, leading also to some surprising links with the well-known sequences of Fibonacci and Pell.

Résumé. Lorsque Maurits Cornelis Escher commença à la fin des années 30 à produire des pavages du plan avec des tuiles, il étonna le monde artistique par la singularité de ses dessins. En particulier, les pavages du plan obtenus avec des copies d'une seule tuile apparaissent souvent dans son oeuvre et ont attiré peu à peu l'attention de la communauté mathématique. Puisqu'une tuile dans le monde continu peut être approximée par un chemin sur un réseau carré suffisemment fin - une méthode universellement utilisée dans les applications utilisant des écrans graphiques - l'object combinatoire qui modèle adéquatement la tuile est le polyomino. Comme ceux-ci sont naturellement codés par des chemins sur un alphabet de quatre lettres, l'utilisation de la combinatoire des mots devient pertinente pour l'étude des propriétés des tuiles pavantes. Nous présentons dans ce papier plusieurs résultats, allant de la reconnaissance de ces tuiles à leur génération, conduisant à des liens surprenants avec les célèbres suites de Fibonacci et de Pell.

Keywords: Tesselations, tilings, polyomino, Fibonacci, Pell.

1 Introduction

We study here a special class of periodic tilings consisting of translated copies of a single tile, and we refer the reader to Grünbaum and Shephard (1987) for a more general presentation of tilings, and to Ardila and Stanley (2005) for an introduction to combinatorial problems related with tilings. For instance, consider the problem of tiling the plane with an infinite number of copies of a single tile. While it is not known whether it admits a periodic tiling of the plane, the situation is easier with translations of a polyomino.

[†] Corresponding author.

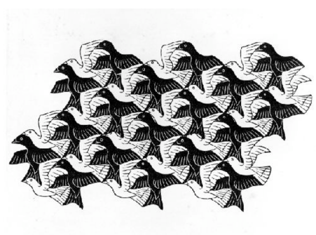
The problem of deciding if a given polyomino tiles the plane by translation goes back to Wijshoff and van Leeuven (1984) who coined the term *exact polyomino* for these. Up to our knowledge, Beauquier and Nivat (1991) were the first to provide a characterization stating that the boundary $b(P)$ of an exact polyomino P satisfies the following (not necessarily unique) *Beauquier-Nivat factorization*

$$b(P) = A \cdot B \cdot C \cdot \widehat{A} \cdot \widehat{B} \cdot \widehat{C} \tag{1}$$

where at most one variable may be empty. Hereafter, this condition is referred as the BN-factorization,

Polyominoes having a BN-factorization (where BN stands for Beauquier-Nivat) with A, B and C nonempty were called *pseudo-hexagons*. For sake of simplicity, we call them *hexagons*. The example on the right shows a hexagonal tiling. Indeed the basic tile is composed of 6 sides, each one corresponding to one side of a hexagon.

If one of the variables in Equation (1) is empty, they are called *squares*. A tiling may have both features as shown below.



Indeed, in the tiling on the left, we have two basic ways of decomposing the tesselation: into squares (a pair of white and black birds) and hegaxons (3 white and 3 black). Even more, a tile may have both hexagon and square factorizations, as a polyomino consisting of k contiguous unit squares shows.

In this paper we present a combinatorial approach for understanding the structure of the polyominoes that tile the plane by translation. The combinatorics on words point of view is powerful for a number of decision problems such as deciding if a polyomino tiles the plane by translation or checking if a tile is digitally convex. Enumeration of such tiles is a challenging problem, and we have exhibited new classes of polyominoes having surprising properties.

Indeed there are square tiles that can be assembled in exactly two different ways, defining two sets of distinct translations: we call them *double squares*. On the other hand we did not find a square having 3 distinct square factorizations, confirming a conjecture due to Provençal (2008). In particular, we describe two infinite families of squares linked to the Christoffel words and to the Fibonacci sequence.

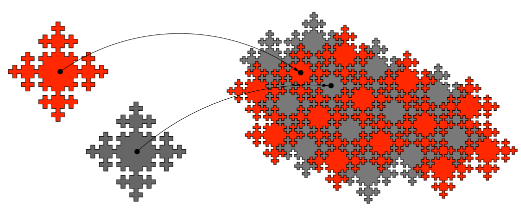


Fig. 1: Tiling with a Fibonacci polyomino.

Proofs of the results are based on combinatorics on words techniques and are omitted due to lack of space.

2 Preliminaries

The usual terminology and notation on words is from Lothaire (1997). An *alphabet* Σ is a finite set whose elements are called *letters*. A finite word w is a sequence of letters, that is, a function $w : \{1, 2, \ldots, n\} \to \Sigma$, where w_i is the i-th letter, $1 \leq i \leq n$. The *length* of w, denoted by $|w|$, is given by the integer n. The unique word of length 0 is denoted ε, and the set of all finite words over Σ is denoted Σ^*. The set of n-length word is Σ^n, and $\Sigma^{\geq k}$ denotes those of length at least k. The *reversal* $\widetilde{w}$ of $w = w_1 w_2 \cdots w_n$ is the word $\widetilde{w} = w_n w_{n-1} \cdots w_1$. Words p satisfying $p = \widetilde{p}$ are called *palindromes*. The set of all palindromes over Σ is denoted $\mathrm{Pal}(\Sigma^*)$. A word u is a *factor* of another word w if there exist $x, y \in \Sigma^*$ such that $w = xuy$. We denote by $|w|_u$ the number of occurrences of u in w. Two words u and v are *conjugate* if there are words x and y such that $u = xy$ and $v = yx$. In that case, we write $u \equiv v$. Clearly, $\equiv$ is an equivalence relation. Given two alphabets Σ_1 and Σ_2, a *morphism* is a function $\varphi : \Sigma_1^* \to \Sigma_2^*$ compatible with concatenation, that is, $\varphi(uv) = \varphi(u)\varphi(v)$ for any $u, v \in \Sigma_1^*$. It is clear that a morphism is completely defined by its action on the letters of Σ_1.

Paths on the square lattice. The notation of this section is partially adapted from Brlek et al. (2006b). A path in the square lattice, identified as $\mathbb{Z} \times \mathbb{Z}$, is a polygonal path made of the elementary unit translations

$$a = (1, 0), \overline{a} = (-1, 0), b = (0, 1), \overline{b} = (0, -1).$$

A finite path w is therefore a word on the alphabet $\mathcal{F} = \{a, \overline{a}, b, \overline{b}\}$. Furthermore, we say that a path w is *closed* if it satisfies $|w|_a = |w|_{\overline{a}}$ and $|w|_b = |w|_{\overline{b}}$. A *simple path* is a word w such that none of its proper factors is a closed path. A *boundary word* is a closed path such that none of its proper factors is closed. Finally, a *polyomino* is the subset of $\mathbb{Z}^2$ contained in some boundary word. On the square grid, a path can be encoded by a sequence of basic movements in the left (L), right (R), forward (F) and backward (B) directions, so that there is a map $\mathcal{D} : \mathcal{F}^2 \to \mathcal{R} = \{\mathrm{L}, \mathrm{R}, \mathrm{F}, \mathrm{B}\}$ defined by

$$\mathcal{D}(u) = \begin{cases} \mathrm{L} & \text{if } u \in V_{\mathrm{L}} = \{ab, b\overline{a}, \overline{a}\overline{b}, \overline{b}a\}, \\ \mathrm{R} & \text{if } u \in V_{\mathrm{R}} = \{ba, a\overline{b}, \overline{b}\overline{a}, \overline{a}b\}, \\ \mathrm{F} & \text{if } u \in V_{\mathrm{F}} = \{aa, \overline{a}\overline{a}, \overline{b}\overline{b}, bb\}, \\ \mathrm{B} & \text{if } u \in V_{\mathrm{B}} = \{a\overline{a}, \overline{a}a, b\overline{b}, \overline{b}b\}. \end{cases}$$

It is extended to a function on arbitrary words, denoted by the same letter $\mathcal{D} : \mathcal{F}^{\geq 1} \to \mathcal{R}^*$, by setting

$$\mathcal{D}(w) = \begin{cases} \varepsilon & \text{if } |w| = 1, \\ \prod_{i=2}^{n} \mathcal{D}(w_{i-1} w_i) & \text{if } |w| \geq 2, \end{cases}$$

where $|w| = n$ and the product is the concatenation. For example, $\mathcal{D}(bab\overline{a}baa\overline{b}) = \mathrm{RLLRRFR}$. Notice that each path $w \in \mathcal{F}^{\geq 1}$ is completely determined, up to translation, by its initial step $\alpha \in \mathcal{F}$ and a word y on the alphabet $\mathcal{R}$. Therefore, for each $\alpha \in \mathcal{F}$ there is a function $\mathcal{D}_\alpha^- : \mathcal{R}^* \to \mathcal{F}^{\geq 1}$ defined recursively by

$$\mathcal{D}_\alpha^{-1}(y) = \begin{cases} \alpha & \text{if } |y| = 0, \\ \alpha \mathcal{D}_\beta^{-1}(y') & \text{if } |y| \geq 1, \end{cases}$$

where $\beta \in \mathcal{F}$ is the letter such that $\alpha\beta \in V_x$ and $y = xy'$ with $x \in \mathcal{R}$. For example, if $y = \mathrm{RLLRRFR}$, then $\mathcal{D}_b^-(y) = bab\overline{a}baa\overline{b}$, while $\mathcal{D}_a^-(y) = a\overline{b}ababa\overline{b}\overline{b}\overline{a}$. The next lemma gives some easily established statements and shows how both functions $\mathcal{D}$ and $\mathcal{D}^-$ behave.

Lemma 1 *Let $w, w' \in \mathcal{F}^*$, $y, y' \in \mathcal{R}^*$, $\alpha \in \mathcal{F}$ and $x \in \mathcal{R}$. Then*

(i) $\mathcal{D}^{-}_{w_1}\mathcal{D}(w) = w$ *and* $\mathcal{D} \circ \mathcal{D}^{-}_{\alpha}(y) = y$, *where w_1 is the first letter of w;*

(ii) $\mathcal{D}(ww') = \mathcal{D}(w) \cdot \mathcal{D}(w_n w'_1) \cdot \mathcal{D}(w')$ *and* $\mathcal{D}^{-}_{\alpha}(yxy') = \mathcal{D}^{-}_{\alpha}(y)\mathcal{D}^{-}_{\beta}(y')$, *where w_n is the last letter of w, w'_1 is the first letter of w' and β is the last letter of $\mathcal{D}^{-}_{\alpha}(yx)$.* □

In Brlek et al. (2006b), the authors introduced the *winding number*, the valuation Δ defined on $\mathcal{R}^*$ by $\Delta(y) = |y|_{\text{L}} - |y|_{\text{R}} + 2|y|_{\text{B}}$ as well as on $\mathcal{F}^{\geq 1}$ by setting $\Delta(w) = \Delta(\mathcal{D}w)$.

Transformations. Some useful transformations on $\mathcal{F}^*$ are rotations by an angle $k\pi/2$ and reflections with respect to axes of angles $k\pi/4$, where $k \in \mathbb{N}$. The rotation of angle $\pi/2$ translates merely in $\mathcal{F}$ by the morphism $\rho : a \mapsto b, b \mapsto \overline{a}, \overline{a} \mapsto \overline{b}, \overline{b} \mapsto a$. We denote the other rotations by ρ^2 and ρ^3 according to the usual notation. The rotation ρ^2 is also noted $\overline{\cdot}$ since it can be seen as the *complement* morphism defined by the relations $\overline{\overline{a}} = a$ and $\overline{\overline{b}} = b$. Similarly, for $k \in \{0, 1, 2, 3\}$, σ_k is the reflection defined by the axis containing the origin and having an angle of $k\pi/4$ with the abscissa. It may be seen as a morphism on $\mathcal{F}^*$ as well:

$$\sigma_0 : a \mapsto \overline{a}, \overline{a} \mapsto a, b \mapsto b, \overline{b} \mapsto \overline{b} \;\text{ and }\; \sigma_1 : a \mapsto b, b \mapsto a, \overline{a} \mapsto \overline{b}, \overline{b} \mapsto \overline{a}.$$

The two other reflections are $\sigma_2 = \sigma_0 \circ \rho^2$ and $\sigma_3 = \sigma_1 \circ \rho^2$. Another useful map is the antimorphism $\widehat{\cdot} = \overline{\cdot} \circ \widetilde{\cdot}$ defined on $\mathcal{F}^*$: $\widehat{w}$ is the path traversed in the opposite direction. The behaviour of the operators $\widehat{\cdot}$, $\widetilde{\cdot}$ and $\overline{\cdot}$ is illustrated in Figure 2.

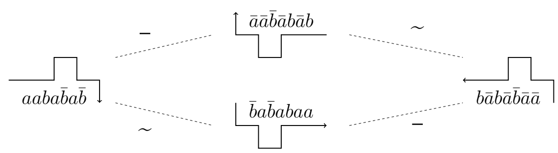


Fig. 2: Effect of the operators $\widehat{\cdot}$, $\widetilde{\cdot}$ and $\overline{\cdot}$ on $\mathcal{F}^*$.

On the alphabet $\mathcal{R}$, we define an involution $\imath : \text{L} \mapsto \text{R}, \text{R} \mapsto \text{L}, \text{F} \mapsto \text{F}, \text{B} \mapsto \text{B}$. This function $\imath$ extends to $\mathcal{R}^*$ as a morphism, so that the map $\widehat{\cdot}$ extends as well to $\widehat{\cdot} : \mathcal{R}^* \to \mathcal{R}^*$ by setting $\widehat{\cdot} = \imath \circ \widetilde{\cdot}$. All these operations are closely related as shown in the lemmas hereafter. The proofs are left to the reader.

Lemma 2 *Let $w \in \mathcal{F}^*$, $y \in \mathcal{R}^*$ and $\alpha \in \mathcal{F}$. The following properties hold:*

(i) $\mathcal{D}(w) = \mathcal{D}(\rho^i(w))$ *for all* $i \in \{1, 2, 3\}$,
(ii) $\imath(\mathcal{D}(w)) = \mathcal{D}(\sigma_i(w))$ *for all* $i \in \{0, 1, 2, 3\}$,
(iii) $\mathcal{D}(\widehat{w}) = \widehat{\mathcal{D}(w)} = \mathcal{D}(\widetilde{w})$,
(iv) $\rho^i(\mathcal{D}^{-}_{\alpha}(y)) = \mathcal{D}^{-}_{\rho^i(\alpha)}(y)$ *for all* $i \in \{1, 2, 3\}$,
(v) $\sigma_i(\mathcal{D}^{-}_{\alpha}(y)) = \mathcal{D}^{-}_{\sigma_i(\alpha)}\imath(y)$ *for all* $i \in \{0, 1, 2, 3\}$,
(vi) $\widetilde{\mathcal{D}^{-}_{\alpha}(y)} = \mathcal{D}^{-}_{\beta}\widehat{y}$ *where β is the last letter of* $\mathcal{D}^{-}_{\alpha}(y)$,
(vii) $\widehat{\mathcal{D}^{-}_{\alpha}(y)} = \mathcal{D}^{-}_{\overline{\beta}}(\widehat{y})$ *where β is the last letter of* $\mathcal{D}^{-}_{\alpha}(y)$,
(viii) *If β is the last letter of $\mathcal{D}^{-}_{\alpha}y$, then $\beta = \rho^i(\alpha)$ where $i = \Delta(y)$.* □

For the rest of the paper, the words w of $\mathcal{F}^*$ and $\mathcal{R}^*$ satisfying $\widehat{w} = w$ are called *antipalindromes*.

Lemma 3 *Let $w \in \mathcal{F}^*$. Then the following statements are equivalent.*

(i) $\widehat{w} = \rho^2(w)$
(ii) *w is a palindrome*
(iii) *$\mathcal{D}(w)$ is an antipalindrome.*

Finally, reflections on $\mathcal{F}^*$ are easily described on $\mathcal{R}^*$.

Lemma 4 *Let $w \in \mathcal{F}^*$. There exists $i \in \{0, 1, 2, 3\}$ such that $\widehat{w} = \sigma_i(w)$ if and only if $\mathcal{D}(w)$ is a palindrome.* $\square$

Square Tilings. Let P be a polyomino having W for boundary word, and Q a square having $V = AB\widehat{A}\widehat{B}$ as a BN-factorization. Then *the product of P and Q*, denoted by $P \circ Q$, is the polyomino whose boundary word is given by $\gamma(W)$, where $\gamma : \mathcal{F}^* \to \mathcal{F}^*$ is the morphism defined by

$$\gamma(a) = A, \gamma(\overline{a}) = \widehat{A}, \gamma(b) = B, \gamma(\overline{b}) = \widehat{B}.$$

The assumption for Q to be a square is essential in order to glue together the tiles. Here is an illustration of the composition where P is a tetramino, which is an hexagon but not a square.

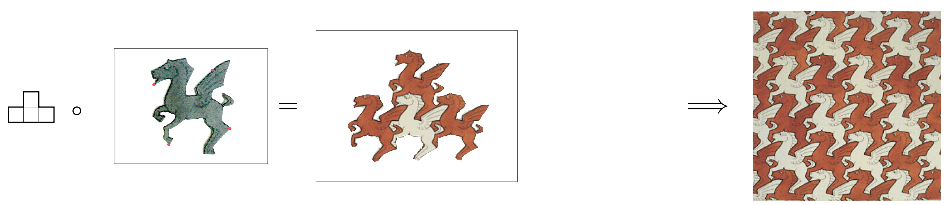


Fig. 3: Composition of tiles and the resulting tiling

Of particular interest are the constructions yielding double squares, discovered by Provençal (2008).

Proposition 5 *Let P be a double square and Q a square. Then the following properties hold:*

(i) *the BN-factorizations of P must overlap, i.e. no factor of a BN-factorization may be included in a factor of the other one.*

(ii) *$P \circ Q$ is a double square.* $\square$

The composition of tiles lead naturally to the notion of primality, and a polyomino R is called *prime* if the relation $R = P \circ Q$ implies that either $R = P$ or $R = Q$. Of course, every square with prime area is prime. In Figure 3, the winged horse is a prime square (!).

Lemma 6 *Let P be a square with boundary word W, A and B be words such that $W \equiv AB\widehat{A}\widehat{B}$. Then $A, B \in \mathrm{Pal}(\mathcal{F}^*)$ if and only if $W = w\overline{w}$ for some word w.*

For more reading on square tilings see Brlek and Provençal (2006); Brlek et al. (2009b).

3 Recognition of tiles

Wijshoff and van Leeuven (1984) provided a naive $\mathcal{O}(n^4)$ algorithm for recognizing exact polyominoes. Later, using the BN-factorization of Beauquier and Nivat, Gambini and Vuillon (2007) exhibited a general $\mathcal{O}(n^2)$ algorithm.

Brlek and Provençal (2006) designed a linear algorithm for recognizing squares. It uses all the power of combinatorics on words as accounted in Lothaire (2005). The main idea is to choose a position p in $W = AB\widehat{A}\widehat{B}$, and then, to list all the candidate factors A that overlap this fixed position p. To achieve this, the auxiliary functions Longest-Common-Right-Extension (LCRE) and Longest-Common-Left-Extension (LCLE) of W and $\widehat{W}$ at some respective positions i and j are essential : their computation is performed in constant time thanks to a pre-processing in linear time (!). (see Lothaire (2005) for more details)

Nevertheless, there is still a gap to close for completely solving the recognition problem. In the case of hexagons the solution is not complete: if the polyominoes do not have too long square factors then the algorithm is still linear (Brlek et al. (2009b)). A general algorithm in $\mathcal{O}(n(\log(n)^3)$ also appears in the thesis of Provençal (2008). Nevertheless, we conjecture that a linear algorithm exists.

It has been shown in Provençal (2008) that there exist polyominoes admitting a linear number of distinct non trivial factorizations as hexagons. The case is different for squares. Indeed, based on exhaustive computation of tiles of small length, showing no square with 3 distinct square factorizations the following result, conjectured in the thesis of Provençal (2008), holds.

Proposition 7 (Blondin Massé et al. (2010a)) *The number of distinct BN-factorizations of a square is at most* 2*.*

4 Generation of tiles

The greedy algorithm, consisting in computing for each even n all polyominoes of perimeter, does not allow to produce large size candidates. Therefore, we used the following approach, which is based on the generation of self avoiding walks.

1. Generate 2 self avoiding walks A, B of length n, m. Each self avoiding walk can be built in two steps:

   (a) generate randomly a word w of length n : this takes $\mathcal{O}(n)$;

   (b) check if w intersects itself, which amounts to check if a point appears twice in the walk. This step can be achieved in $\mathcal{O}(n)$ thanks to a sequential algorithm we provided in Brlek et al. (2009a)). The key point is that there is no need to sort the points (which requires a $\log n$ factor) thanks to the combination of two data structures: a radix-tree for storing the visited points on the square grid is enriched with a quad-tree structure encoding the neighborhood relation of points.

   Note : both steps can be combined in a single pass of Step 1(b). It suffices to substitute the sequential reading with the sequential random generation of a letter.

2. Check if the final word $AB\widehat{A}\widehat{B}$ does not intersect itself by extending the nonintersection verification in Step 1 (b) with the factor $\widehat{A}\widehat{B}$.

The resulting algorithm is clearly linear in the perimeter size.

Some double squares are displayed in Figure 4. In fact, there is an infinite number of these and we are able to describe some infinite families.

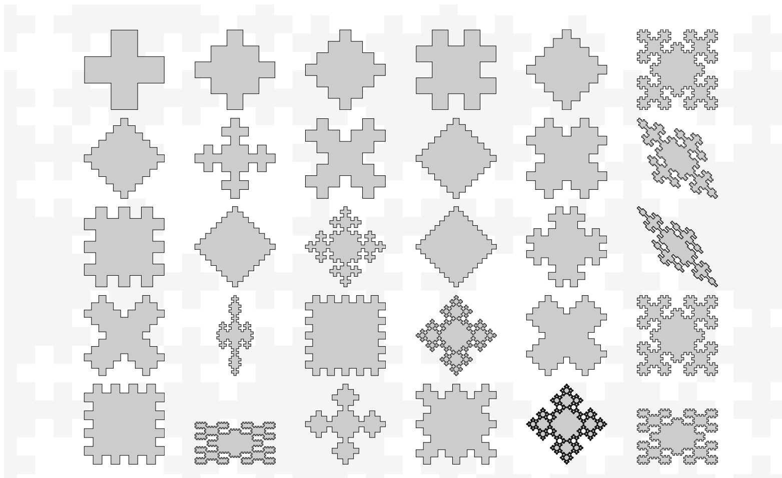


Fig. 4: Some double squares

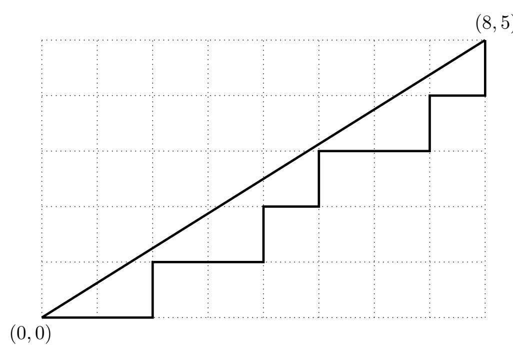
The first diagonal in Figure 4 contains what we call the Fibonacci tiles. Two special classes of double squares are described now.

4.1 Christoffel Tiles

Recall that Christoffel words are finite Sturmian words, that is, they are obtained by discretizing a segment in the plane. Let $(p, q) \in \mathbb{N}^2$ with $\gcd(p, q) = 1$, and let S be the segment with endpoints $(0, 0)$ and (p, q).

The word w is a *lower Christoffel word* if the path induced by w is under S and if they both delimit a polygon with no integral interior point. An *upper Christoffel word* is defined similarly. A *Christoffel word* is either a lower Christoffel word or an upper Christoffel word. On the right is illustrated the lower one corresponding to

$$w = aabaababaabab.$$



It is well known that if w and w' are respectively the lower and upper Christoffel words associated to (p, q), then $w' = \widetilde{w}$. Moreover, we have $w = amb$ and $w' = bma$, where m is a palindrome and a, b are letters. The word m is called *cutting word*. They have been widely studied in the literature (see e.g. Berstel et al. (2008)), where they are also called *central words*.

Let $\mathcal{B} = \{a, b\}$. Consider the morphism $\lambda : \mathcal{B}^* \to \mathcal{F}^*$ by $\lambda(a) = a\bar{b}ab$ and $\lambda(b) = ab$, which can be seen as a "crenelation" of the steps *east* and *north-east*.

Theorem 8 (Blondin Massé et al. (2009)) *Let $w = amb$ where a and b are letters.*

(i) *If m is a palindrome, then $\lambda(w\overline{w})$ is a square tile.*

(ii) *$\lambda(w\overline{w})$ is a double square if and only if w is a Christoffel word.*

We say that a crenelated tile $\lambda(w\overline{w})$ obtained from a lower Christoffel word w is a *basic Christoffel tile* while a *Christoffel tile* is a polyomino isometric to a basic Christoffel tile under some rotations ρ and symmetries σ_i (see Figure 5).

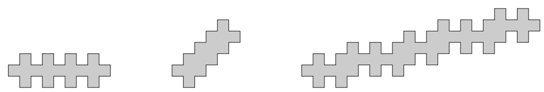


Fig. 5: Basic Christoffel tiles: (a) $w = aaaab$ (b) $w = abbbb$ and (c) $w = aabaababaabab$.

Theorem 9 (Blondin Massé et al. (2009)) *Let P be a crenelated tile. Then P is a double square if and only if it is obtained from a Christoffel word.* □

It can also be shown in view of Lemma 6, that each Christoffel tile is highly symmetrical.

Proposition 10 *If $AB\widehat{A}\widehat{B}$ is a BN-factorization of a Christoffel tile, then A and B are palindromes.*

Moreover, one verifies easily the following facts.

Proposition 11 *Let T be a Christoffel tile obtained from the (p, q) Christoffel word, where p and q are relatively prime. Then the perimeter and the area of T are given respectively by $\mathbf{P}(T) = 8p + 4q$ and $\mathbf{A}(T) = 4p + 3q - 2$.* □

4.2 Fibonacci Tiles

In this section, in order to simplify the notation, we redefine the operator $\overline{\cdot}$ on $\mathcal{R}^*$ by setting $\overline{y} = \iota(y)$, where ι is the involution $\iota : \mathtt{R} \leftrightarrow \mathtt{L}, \mathtt{F} \mapsto \mathtt{F}, \mathtt{B} \mapsto \mathtt{B}$. We define a sequence $(q_n)_{n\in\mathbb{N}}$ in $\mathcal{R}^*$ by setting $q_0 = \varepsilon$, $q_1 = \mathtt{R}$ and

$$q_n = \begin{cases} q_{n-1}q_{n-2} & \text{if } n \equiv 2 \bmod 3, \\ q_{n-1}\overline{q_{n-2}} & \text{if } n \equiv 0, 1 \bmod 3. \end{cases}$$

whenever $n \geq 2$. The first terms of $(q_n)_{n\in\mathbb{N}}$ are

q_0	=	ε	q_3	=	RL	q_6	=	RLLRLLRR
q_1	=	R	q_4	=	RLL	q_7	=	RLLRLLRRLRRLR
q_2	=	R	q_5	=	RLLRL	q_8	=	RLLRLLRRLRRLRRLLRLLRR

Note that $|q_n| = F_n$ is the n-th Fibonacci number. Moreover, given $\alpha \in \mathcal{F}$, the path $\mathcal{D}_\alpha^-(q_n)$ presents strong symmetric properties. The next two lemmas are from Blondin Massé et al. (2010b).

Lemma 12 *Let $n \in \mathbb{N}$. Then $q_{3n+1} = p\alpha$, $q_{3n+2} = q\alpha$ and $q_{3n+3} = r\overline{\alpha}$ for some antipalindrome p, and some palindromes q, r and some letter $\alpha \in \{\mathtt{L}, \mathtt{R}\}$.*

Lemma 13 *Let $n \in \mathbb{N}$ and $\alpha \in \mathcal{F}$. Then $\mathcal{D}_\alpha^-(q_n)$ is a simple path and $\mathcal{D}_\alpha^-(q_{3n+1})^4$ is the boundary word of a polyomino.*

A Fibonacci tile of order n is a polyomino having $\mathcal{D}_\alpha^-(q_{3n+1})^4$ as a boundary word, where $n \in \mathbb{N}$. They are somehow related to the Fibonacci fractals found in Monnerot-Dumaine. The first Fibonacci tiles are illustrated in Figure 6.

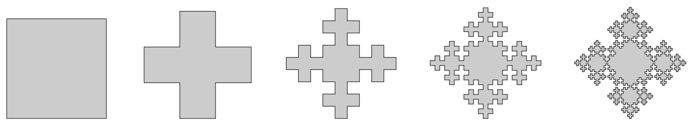


Fig. 6: Fibonacci tiles of order $n = 0, 1, 2, 3, 4$.

Theorem 14 *Fibonacci tiles are double squares.*

As for Christoffel tiles, Fibonacci tiles also suggest that the conjecture of Provençal (2008) for double squares is true as stated in the next result.

Corollary 15 *If $AB\widehat{A}\widehat{B}$ is a BN-factorisation of a Fibonacci tile, then A and B are palindromes.*

We have established in Blondin Massé et al. (2010b) that the perimeter of the Fibonacci tiles is given by $4F(3n+1)$ while their area $A(n)$ satisfies the recurrence formulas

$$\begin{aligned} A(0) &= 1, \quad A(1) = 5; \\ A(n) &= 6A(n-1) - A(n-2), \text{ for } n \geq 2, \end{aligned}$$

whose first terms are $1, 5, 29, 169, 985, 5741, 33461, 195025, 1136689, 6625109, 38613965, \ldots$ This sequence is the subsequence of odd index Pell numbers.

We end this section by presenting four families of double squares, a variant of the Fibonacci tiles whose areas satisfy the same recurrence. Indeed, consider the sequence $(r_{d,m,n})_{(d,m,n)\in\mathbb{N}^3}$ satisfying the following recurrence, for $d \geq 2$,

$$r_{d,m,n} = \begin{cases} r_{d-1,n,m}\overline{r_{d-2,n,m}} & \text{if } d \equiv 0 \bmod 3 \\ r_{d-1,n,m} r_{d-2,n,m} & \text{if } d \equiv 1 \bmod 3 \\ r_{d-1,m,n}\overline{r_{d-2,m,n}} & \text{if } d \equiv 2 \bmod 3 \end{cases}$$

Using similar arguments as in the Fibonacci tiles case, one shows that both families obtained respectively with seed values

$$\begin{aligned} r_{0,m,n} &= (\text{RLLR})^m\text{RLR}, & r_{1,m,n} &= (\text{RLLR})^n\text{R}, \\ r_{0,m,n} &= (\text{RL})^m\text{RLR}, & r_{1,m,n} &= (\text{RL})^n\text{RL} \end{aligned}$$

are such that $\mathcal{D}_\alpha^-(r_{3d,m,n}r_{3d,n,m})^2$, where $\alpha \in \mathcal{F}$, is a boundary word whose associated polyomino is a double square (see Figure 7). Their level of fractality increases with d so that one could say that they are crenelated versions of the Fibonacci Tiles.

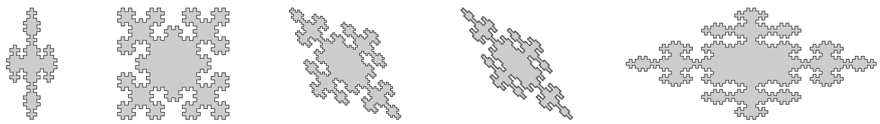


Fig. 7: Tiles obtained with different seeds: from $r_{2,0,1}$, from $r_{3,m,0}$ for $m = 0, 1, 2$, from $r_{3,1,0}$.

Similarly, let $(s_{d,m,n})_{(d,m,n)\in\mathbb{N}^3}$ be a sequence satisfying for $d \geq 2$ the recurrence

$$s_{d,m,n} = \begin{cases} s_{d-1,n,m}\overline{s_{d-2,n,m}} & \text{if } d \equiv 0, 2 \bmod 3, \\ s_{d-1,m,n}s_{d-2,m,n} & \text{if } d \equiv 1 \bmod 3. \end{cases}$$

Then the families obtained with seed values

$$\begin{array}{rclcrcl} s_{0,m,n} & = & (\mathrm{RLLR})^m\mathrm{RLR}, & & s_{1,m,n} & = & \mathrm{RL}, \\ s_{0,m,n} & = & (\mathrm{RL})^m\mathrm{RLR}, & & s_{1,m,n} & = & \mathrm{R} \end{array}$$

yield double squares $\mathcal{D}^-_\alpha(s_{3d,m,n}s_{3d,n,m})^2$ as well (see Figure 8). One may verify that $r_{d,0,0} = s_{d,0,0}$ for any $d \in \mathbb{N}$ for some conveniently chosen seed values.

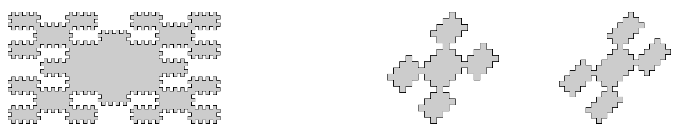


Fig. 8: Tile obtained from $s_{3,2,0}$, and from $s_{2,0,n}$ for $n = 1, 2$.

The area of the tiles $\mathcal{D}^-(r_{3d,m,n}r_{3d,n,m})^2$ and $\mathcal{D}^-(s_{3d,m,n}s_{3d,n,m})^2$ for each values of d, m and n share particular properties. In fact, all the sequences are obtained by the same recurrence (see the first values in Table 1), and we have the following proposition.

Proposition 16 *Let $m, n \in \mathbb{N}$ be fixed. The sequence of areas indexed by $d \in \mathbb{N}$ of the four families of generalized Fibonacci tiles satisfy the recurrence $A(d) = 6A(d-1) - A(d-2)$ for $d \geq 2$.* □

5 Concluding remarks

The study of double squares suggests some interesting and challenging problems. For instance, there is a conjecture of Provençal (2008) stating that if $AB\widehat{A}\widehat{B}$ is a BN-factorization of a prime double square, then A and B are palindromes, for which, despite a lot of computation time, we have not been able to provide any counter-example. Another problem is to prove that Christoffel and Fibonacci tiles are prime, that is, they are not obtained by composition of smaller squares. This leads to a number of questions on the "arithmetics" of tilings, such as the unique decomposition, distribution of prime tiles, and their enumeration. Partial results may be found in Brlek et al. (2006a)

It is also appealing to conjecture that a prime double square is either of Christoffel type or of Fibonacci type. However, that is not the case, as illustrated by Figure 9. This begs for a thorough study in order to exhibit a complete zoology of such tilings.

m	n	d							
		0	1	2	3	4	5	6	7
0	0	5	13	73	425	2477	14437	84145	490433
1	0	9	29	165	961	5601	32645	190269	1108969
0	1	9	29	165	961	5601	32645	190269	1108969
2	0	13	45	257	1497	8725	50853	296393	1727505
1	1	17	65	373	2173	12665	73817	430237	2507605
0	2	13	45	257	1497	8725	50853	296393	1727505
3	0	17	61	349	2033	11849	69061	402517	2346041
2	1	25	101	581	3385	19729	114989	670205	3906241
1	2	25	101	581	3385	19729	114989	670205	3906241
0	3	17	61	349	2033	11849	69061	402517	2346041
4	0	21	77	441	2569	14973	87269	508641	2964577
3	1	33	137	789	4597	26793	156161	910173	5304877
2	2	37	157	905	5273	30733	179125	1044017	6084977
1	3	33	137	789	4597	26793	156161	910173	5304877
0	4	21	77	441	2569	14973	87269	508641	2964577

Tab. 1: Area of the tile $\mathcal{D}^-(r_{3d,m,n}r_{3d,n,m})^2$ with seed values $r_{0,m,n} = (\text{RLLR})^m\text{RLR}$ and $r_{1,m,n} = (\text{RLLR})^n\text{R}$.

The fractal nature of the Fibonacci tiles strongly suggests that Lindemayer systems (L-systems) may be used for their construction Rozenberg and Salomaa (1980). The formal grammars used for describing them have been widely studied, and their impact in biology, computer graphics Rozenberg and Salomaa (2001) and modeling of plants is significant Prusinkiewicz and Lindenmayer (1990). A number of designs including snowflakes fall into this category.

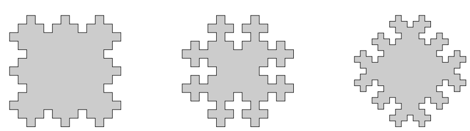


Fig. 9: Three double squares not in the Christoffel and Fibonacci tiles families.

References

F. Ardila and R. Stanley. Tilings. arXiv:math/0501170v3, 2005.

D. Beauquier and M. Nivat. On translating one polyomino to tile the plane. *Discrete Comput. Geom.*, 6: 575–592, 1991.

J. Berstel, A. Lauve, C. Reutenauer, and F. Saliola. *Combinatorics on Words: Christoffel Words and Repetition in Words*, volume 27 of *CRM monograph series*. American Mathematical Society, 2008.

A. Blondin Massé, S. Brlek, A. Garon, and S. Labbé. Christoffel and Fibonacci tiles. In S. Brlek, X. Provençal, and C. Reutenauer, editors, *DGCI 2009, 15th IAPR Int. Conf. on Discrete Geometry for Computer Imagery*, number 5810 in LNCS, pages 67–78. Springer-Verlag, 2009.

A. Blondin Massé, S. Brlek, A. Garon, and S. Labbé. Every polyomino yields at most two square tilings. In *7th Int. Conf. on Lattice Paths Combinatorics and Applications*, 2010a.

A. Blondin Massé, S. Brlek, S. Labbé, and M. Mendès France. Fibonacci snowflakes. *Annales des Sciences Mathématiques du Québec.*, 2010b. To appear.

S. Brlek and X. Provençal. An optimal algorithm for detecting pseudo-squares. In A. Kuba, L. G. Nyúl, and K. Palágyi, editors, *DGCI 2006, 13th Int. Conf. on Discrete Geometry for Computer Imagery*, number 4245 in LNCS, pages 403–412. Springer-Verlag, 2006.

S. Brlek, A. Frosini, S. Rinaldi, and L. Vuillon. Tilings by translation: Enumeration by a rational language approach. *Electr. J. Comb.*, 13(1), 2006a.

S. Brlek, G. Labelle, and A. Lacasse. Properties of the contour path of discrete sets. *Int. J. Found. Comput. Sci.*, 17(3):543–556, 2006b.

S. Brlek, M. Koskas, and X. Provençal. A linear time and space algorithm for detecting path intersection. In S. Brlek, X. Provençal, and C. Reutenauer, editors, *DGCI 2009, 15th IAPR Int. Conf. on Discrete Geometry for Computer Imagery*, number 5810 in LNCS, pages 398–409. Springer-Verlag, 2009a.

S. Brlek, X. Provençal, and J.-M. Fédou. On the tiling by translation problem. *Discrete Applied Mathematics*, 157(3):464–475, 2009b.

I. Gambini and L. Vuillon. An algorithm for deciding if a polyomino tiles the plane by translations. *Theoret. Informatics Appl.*, 41:147–155, 2007.

B. Grünbaum and G. C. Shephard. *Tilings and Patterns*. W. H. Freeman, New York, 1987.

M. Lothaire. *Combinatorics on Words*. Cambridge University Press, Cambridge, 1997.

M. Lothaire. *Applied Combinatorics on Words*. Cambridge University Press, Cambridge, 2005.

A. Monnerot-Dumaine. The Fibonacci word fractal. August 2009. Preprint available electronically at `http://hal.archives-ouvertes.fr/hal-00367972/en/`.

X. Provençal. *Combinatoire des mots, géométrie discrète et pavages*. PhD thesis, D1715, Université du Québec à Montréal, 2008.

P. Prusinkiewicz and A. Lindenmayer. *The algorithmic beauty of plants*. Springer-Verlag New York, Inc., New York, NY, USA, 1990. ISBN 0-387-97297-8.

G. Rozenberg and A. Salomaa. *Mathematical theory of L-systems*. Academic Press, Inc., Orlando, FL, USA, 1980. ISBN 0125971400.

G. Rozenberg and A. Salomaa, editors. *Lindenmayer Systems: Impacts on Theoretical Computer Science, Computer Graphics, and Developmental Biology*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2001. ISBN 0387553207.

H. A. G. Wijshoff and J. van Leeuven. Arbitrary versus periodic storage schemes and tesselations of the plane using one type of polyomino. *Inform. Control*, 62:1–25, 1984.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 420–431

Viewing counting polynomials as Hilbert functions via Ehrhart theory

Felix Breuer[1]†and Aaron Dall[2]

[1] *Freie Universität Berlin, Arnimallee 3, 14195 Berlin, Germany,* felix.breuer@fu-berlin.de
[2] adall1979@gmail.com

Abstract.

Steingrímsson (2001) showed that the chromatic polynomial of a graph is the Hilbert function of a relative Stanley-Reisner ideal. We approach this result from the point of view of Ehrhart theory and give a sufficient criterion for when the Ehrhart polynomial of a given relative polytopal complex is a Hilbert function in Steingrímsson's sense. We use this result to establish that the modular and integral flow and tension polynomials of a graph are Hilbert functions.

Résumé. Steingrímsson (2001) a montré que le polynôme chromatique d'un graphe est la fonction de Hilbert d'un idéal relatif de Stanley-Reisner. Nous abordons ce résultat du point de vue de la théorie d'Ehrhart et donnons un critère suffisant pour que le polynôme d'Ehrhart d'un complexe polytopal relatif donné soit une fonction de Hilbert au sens de Steingrímsson. Nous utilisons ce résultat pour établir que les polynômes de flux et de tension modulaires et intégraux d'un graphe sont des fonctions de Hilbert.

Keywords: Hilbert function, lattice polytope, relative Stanley-Reisner ring, tension polynomial, flow polynomial, relative polytopal complex

1 Introduction

Steingrímsson [Ste01] showed that the proper $k+1$-colorings of a graph G are in bijection with the monomials of degree k in a polynomial ring $\mathbb{K}[x_1, \ldots, x_n]$ that lie inside a square-free monomial ideal I_2, but outside a square-free monomial ideal I_1. In other words, he showed that the chromatic polynomial χ_G of G is the Hilbert function of a relative Stanley-Reisner ideal. To this end, he used a clever combinatorial construction to describe the ideals I_1 and I_2 explicitly.

In this article we approach the problem from the point of view of Ehrhart theory, which allows us to arrive quickly at a sufficient criterion for when the Ehrhart polynomial of a given relative polytopal complex is a Hilbert function in Steingrímsson's sense:

> The Ehrhart function of a relative polytopal complex in which all faces are compressed is the Hilbert function of a relative Stanley-Reisner ideal.

† Supported by the DFG research training group "Methods for Discrete Structures" (GrK 1408).

See Theorem 5. We then apply this general result to establish that four other counting polynomials defined in terms of graphs are Hilbert functions: the modular flow and tension polynomials and their integral variants. Also, we are able to improve Steingrímsson's result insofar as we are able to obtain the chromatic polynomial $\chi_G(k)$ itself as a Hilbert function, and not only the shifted chromatic polynomial $\chi_G(k+1)$. We conclude the paper by a giving another more algebraic proof of our geometric theorem, which allows us to generalize the result further.

These results have been developed in the authors' respective theses [Dal08] and [Bre09], to which we refer the interested reader for further details and additional material.

This paper is organized as follows. After some preliminary definitions in Section 2 we review Steingrímsson's theorem and related work in Section 3. In Section 4 our main result is derived. In Section 5 we apply this result to show that all five counting polynomials are Hilbert functions. We present a generalization of our main result in Section 6 along with a more algebraic proof. Finally, we give some constraints on the coefficients of the polynomials and discuss questions for further research in Section 7.

2 Preliminary Definitions

Before we begin, we gather some definitions. We recommend the textbooks [BR07, MS05, Sta96, Sch86] as references.

The **Ehrhart function** L_A of any set $A \subset \mathbb{R}^n$ is defined by $\mathsf{L}_A(k) = |\mathbb{Z}^n \cap k \cdot A|$ for $k \in \mathbb{N}$. A **lattice polytope** is a polytope in $\mathbb{R}^n$, such that all vertices are integer points. It is a theorem of Ehrhart that the Ehrhart function $\mathsf{L}_P(k)$ of a lattice polytope is a polynomial in k. Two polytopes P, Q are **lattice isomorphic**, $P \approx Q$, if there exists an affine isomorphism A such that $A|_{\mathbb{Z}^n}$ is a bijection onto $\mathbb{Z}^n$ and $AP = Q$. A d**-simplex** is the convex hull of $d+1$ affinely independent points. A d-simplex is **unimodular** if it is lattice isomorphic to the convex hull of $d+1$ standard unit vectors. A lattice polytope is **empty** if the only lattice points it contains are its vertices. A **hyperplane arrangement** is a finite collection $\mathcal{H}$ of affine hyperplanes and $\bigcup \mathcal{H}$ denotes the union of all of these.

A **polytopal complex** is a finite collection $\mathcal{C}$ of polytopes in some $\mathbb{R}^n$ with the following two properties: If $P \in \mathcal{C}$ and F is a face of P, then $F \in \mathcal{C}$; and if $P, Q \in \mathcal{C}$ then $F = P \cap Q \in \mathcal{C}$ and F is common face of both P and Q. The polytopes in $\mathcal{C}$ are also called **faces** and $\bigcup \mathcal{C}$ denotes the union of all faces of $\mathcal{C}$. A (geometric) **simplicial complex** is a polytopal complex in which all faces are simplices. An **abstract simplicial complex** is a set Δ of subsets of a finite set V, such that Δ is closed under taking subsets. A geometric simplicial complex Δ gives rise to an abstract simplicial complex $\mathsf{comb}(\Delta)$ via $\mathsf{comb}(\Delta) = \{\sigma | \sigma \text{ is the vertex set of some } F \in \mathcal{C}\}$. A polytopal complex $\mathcal{C}'$ that is a subset $\mathcal{C}' \subset \mathcal{C}$ of a polytopal complex $\mathcal{C}$ is called a **subcomplex** of $\mathcal{C}$. Subcomplexes of abstract simplical complexes are defined similarly. Given a collection S of polytopes in $\mathbb{R}^n$ such that for any $P, Q \in S$ the set $P \cap Q$ is a face of both P and Q, the polytopal complex $\mathcal{C}$ **generated** by S, is $\mathcal{C} = \{F | F \text{ a face of } P \in S\}$. A **subdivision** of a polytopal complex $\mathcal{C}$ is a polytopal complex $\mathcal{C}'$ such that $\bigcup \mathcal{C} = \bigcup \mathcal{C}'$ and every face of $\mathcal{C}'$ is contained in a face of $\mathcal{C}$. A **triangulation** is a subdivision in which all faces are simplicies. A **unimodular triangulation** is a triangulation in which all simplices are unimodular.

Let $\mathbb{K}[x] = \mathbb{K}[x_1, \ldots, x_n]$ denote the polynomial ring in n variables over some field $\mathbb{K}$ equipped with the standard grading by degree $\mathbb{K}[x] = \bigoplus_{k \geq 0} R_k$, where R_k is the $\mathbb{K}$-vector space generated by all monomials of degree k. A **graded** $\mathbb{K}[x]$**-module** is a module M that can be written as a direct sum of abelian groups $M = \bigoplus_{-\infty}^{\infty} M_k$ such that $R_i M_j \subset M_{i+j}$ for all i and j. The **Hilbert function** H_M of M is defined by $\mathsf{H}_M(k) = \dim_{\mathbb{K}} M_k$. Let I_1 be a monomial ideal in $\mathbb{K}[x]$ and consider the quotient ring

$\mathbb{K}[x]/I_1$ graded by degree. Then $\mathsf{H}_{\mathbb{K}[x]/I_1}(k) = |\{x^a \in \mathbb{K}[x] | x^a \notin I_1, \deg(x^a) = k\}|$. Furthermore let $I_2 \supset I_1$ be another monomial ideal. By abuse of notation we also denote by I_2 the ideal in $\mathbb{K}[x]/I_1$ generated by the same set of monomials. We can write $I_2 = \bigoplus_{k \geq 0} I_2^k$ where I_2^k is the vector space generated by the monomials x^a in I_2 with $\deg(x^a) = k$ that are non-zero in $\mathbb{K}[x]/I_1$. Then $R_i I_2^k \subset I_2^{i+k}$, where the product is taken in $\mathbb{K}[x]/I_1$, and $\mathsf{H}_{I_2}(k) = |\{x^a \in I_2 \setminus I_1 | \deg(x^a) = k\}|$. A **term order** on $\mathbb{K}[x]$ is a total order on the monomials $x^a \in \mathbb{K}[a]$ such that $1 \prec x^a$ for all $a \in \mathbb{Z}^n_{>0}$ and $x^a \prec x^b$ implies $x^{a+c} \prec x^{b+c}$ for all $a, b, c \in \mathbb{Z}^n_{\geq 0}$.

We consider oriented graphs that may have loops and multiple edges. Note, however, that the values of the five counting polynomials do not depend on the orientation of the graph and are thus invariants of the underlying unoriented graph. Formally, a **graph** is a tuple $(V, E, \mathsf{head}, \mathsf{tail})$, where V is a finite vertex set, E is a finite edge set and $\mathsf{head} : E \to V$ and $\mathsf{tail} : E \to V$ are maps. Graph theoretic concepts such as adjacency, paths, connectivity, etc. are defined in the usual way. We note that a cycle in the underlying unoriented graph can be coded as a map $c : E \to \{0, \pm 1\}$ where $c_e = +1$ if the direction in which e is traversed is consistent with the orientation of e in G, $c_e = -1$ if the direction of traversal is opposite to the orientation in G and $c_e = 0$ if e does not lie on the cycle. Here we view c both as a map and as a vector as we shall do with all maps defined in this article. Let $k \in \mathbb{Z}_{>0}$. A k**-coloring** of G is a map $x : V \to \{0, \ldots, k-1\}$ and it is called *proper* if $x_v \neq x_u$ whenever $u \sim v$. The **chromatic polynomial** χ_G is defined such that $\chi_G(k)$ is the number of proper k-colorings of G.

A k**-tension** of G is a map $t : E \to \{-k+1, \ldots, k-1\}$ such that

$$\sum_{e \in E} c_e t_e = 0 \qquad \text{for every cycle } c \text{ in } G. \tag{1}$$

Similarly, a $\mathbb{Z}_k$**-tension** of G is a map $t : E \to \mathbb{Z}_k$ such that (1) holds in $\mathbb{Z}_k$. A tension is **nowhere zero** if $t(e) \neq 0$ for all $e \in E$. Now we define functions θ_G and $\bar{\theta}_G$ as follows: $\theta_G(k)$ is the number of nowhere zero k-tensions of G and $\bar{\theta}_G(k)$ is the number of nowhere zero $\mathbb{Z}_k$-tensions of G. Both $\theta_G(k)$ and $\bar{\theta}_G(k)$ are polynomials in k, called the **integral** and the **modular tension polynomial**, respectively.

A k**-flow** of G is a map $f : E \to \{-k+1, \ldots, k-1\}$ such that

$$\sum_{\substack{e \in E \\ \mathsf{head}(e) = v}} f_e - \sum_{\substack{e \in E \\ \mathsf{tail}(e) = v}} f_e = 0 \qquad \text{for every vertex } v \text{ of } G. \tag{2}$$

Similarly, a $\mathbb{Z}_k$**-flow** of G is a map $f : E \to \mathbb{Z}_k$ such that (2) holds in $\mathbb{Z}_k$. The functions φ_G and $\bar{\varphi}_G$ are defined as follows: $\varphi_G(k)$ is the number of nowhere zero k-flows of G and $\bar{\varphi}_G(k)$ is the number of nowhere zero $\mathbb{Z}_k$-flows of G. Both $\varphi_G(k)$ and $\bar{\varphi}_G(k)$ are polynomials in k, called the **integral** and the **modular flow polynomial**, respectively. More about these polynomials can be found in [BZ06a, BZ06b, Koc02, Bre09].

3 Steingrímsson's theorem and related work

Steingrímsson [Ste01] showed that for any graph G the chromatic polynomial $\chi_G(k+1)$ shifted by one is the Hilbert function of a module with a particular structure.

Theorem 1 *(Steingrímsson [Ste01, Theorem 9]) For any graph G, there exists a number n, a square-free monomial ideal I_1 in the polynomial ring over n variables $\mathbb{K}[x] = \mathbb{K}[x_1, \ldots, x_n]$ and a square-free* *monomial ideal I_2 in $\mathbb{K}[X]/I_1$ such that*

$$\mathsf{H}_{I_2}(k) = \chi_G(k+1)$$

for all $k \in \mathbb{Z}_{>0}$, where H_{I_2} denotes the Hilbert function of I_2 with respect to the standard grading and χ_G denotes the chromatic polynomial of G.

In [Ste01] Steingrímsson went on to define the coloring complex of a graph to be the simplicial complex given by the square-free monomial ideal I_2. In the case of colorings the ideal I_1 has a simple description, so bounds on the f-vector of the coloring complex translate into bounds on the coefficients of the chromatic polynomial. The articles [Jon05, Hul07, HS08], building on Steingrímsson's work, have mainly dealt with showing various properties of the coloring complex. Steingrímsson himself gave a combinatorial description of the coloring complex and determined its Euler characteristic to be the number of acyclic orientations of G. To some extent this was already known: Welker observed that the coloring complex of a graph $G = (V, E)$ is the same as a complex appearing in the article [HRW98] by Herzog, Reiner and Welker, where this complex is shown to be homotopy equivalent to a wedge of spheres of dimension $|V| - 3$ and the number of spheres is the number of acyclic orientations of G minus one. Jonsson [Jon05] showed the coloring complex to be constructible and hence Cohen-Macaulay. This result was improved by Hultman [Hul07] who showed the coloring complex to be shellable and by Hersh and Swartz [HS08] who showed that the coloring complex has a convex ear decomposition. These results translate into bounds on the coefficients of χ_G.

In this article we concentrate on establishing the structural result that the four counting polynomials are Hilbert functions in Steingrímsson's sense. We deal with the question of obtaining bounds on the coefficients in the companion article [BD10]. In the present article, we only characterize the coefficient vectors of Hilbert polynomials of ideals of Steingrímsson's type in Section 7. In [BD10] we exploit geometric information to derive stronger constraints for these four polynomials.

4 Hilbert equals Ehrhart

In this section we relate Ehrhart functions of certain complexes to Hilbert functions of ideals defined in terms of these complexes. We begin with the well-known relation between simplicial complexes and the corresponding Stanley-Reisner ideals, move on to relative simplicial complexes and relative Stanley-Reisner ideals before we finally consider relative polytopal complexes.(i) As Ehrhart functions are defined in terms of geometric simplicial complexes while Stanley-Reisner ideals are defined in terms of abstract simplicial complexes, all the complexes we consider live in both worlds. A geometric simplicial complex Δ has an abstract simplicial complex $\mathsf{comb}(\Delta)$ associated with it, see Section 2.

Let Δ be an abstract simplicial complex on the ground set V. We identify the elements of the ground set of Δ with the variables in the polynomial ring $\mathbb{K}[x_v : v \in V] =: \mathbb{K}[x]$. Thus sets $S \subset V$ correspond to square-free monomials in $\mathbb{K}[x]$. The **Stanley-Reisner ideal** I_Δ of Δ is generated by the monomials corresponding to the minimal non-faces of Δ, more precisely

$$I_\Delta := \langle x^u \in K[x] | \mathsf{supp}(u) \notin \Delta \rangle.$$

Then, the **Stanley-Reisner ring** of Δ is the quotient $\mathbb{K}[\Delta] = \mathbb{K}[x]/I_\Delta$. We equip the ring $\mathbb{K}[x]$ with the standard grading, that is for any monomial $x^u \in \mathbb{K}[x]$ we have $\deg(x^u) = ||u||_1 = \sum_{i=1}^n u_i$. The fundamental result about Stanley-Reisner rings is this:

(i) We introduce the polytopal Stanley-Reisner ideals corresponding to polytopal complexes only later in Section 6.

Theorem 2 *[Sta96] Let Δ be a d-dimensional (abstract) simplicial complex with f_i faces of dimension i for $0 \leq i \leq d$. Then the Hilbert function $\mathsf{H}_{\mathbb{K}[\Delta]}$ of the Stanley-Reisner ring $\mathbb{K}[\Delta]$ satisfies*

$$\mathsf{H}_{\mathbb{K}[\Delta]}(k) \quad = \quad \sum_{i=0}^{d} f_i \binom{k-1}{i} \tag{3}$$

for $k \in \mathbb{Z}_{>0}$ and $\mathsf{H}_{\mathbb{K}[\Delta]}(0) = 1$.

We remark that the right-hand side of (3) evaluated at zero gives $\sum_{i=0}^{d} f_i \binom{-1}{i} = \chi(\Delta)$, the Euler characteristic of Δ.

If we are given a geometric simplicial complex Δ we will generally use $\mathsf{vert}(\Delta)$ as the ground set of the abstract simplicial complex $\mathsf{comb}(\Delta)$ and identify the variables of $\mathbb{K}[x]$ with the vertices of Δ. In this case we use I_Δ to refer to $I_{\mathsf{comb}(\Delta)}$ and similarly for $\mathbb{K}[\Delta]$.

Now the Ehrhart functions of a unimodular d-dimensional lattice simplex σ^d and its relative interior $\mathsf{relint}\, \sigma^d$ are, respectively,

$$\mathsf{L}_{\sigma^d}(k) = \binom{k+d}{d} \quad \text{and} \quad \mathsf{L}_{\mathsf{relint}\, \sigma^d}(k) = \binom{k-1}{d}. \tag{4}$$

Taken together, (3) and (4) tell us that for any (geometric) simplicial complex Δ in which all simplices are unimodular, the Ehrhart function $\mathsf{L}_\Delta(k) = |\mathbb{Z}^d \cap k \bigcup \Delta|$ of Δ satisfies

$$\mathsf{L}_\Delta(k) \quad = \quad \textstyle\sum_{\sigma \in \Delta} \mathsf{L}_{\mathsf{relint}\, \sigma}(k) \quad = \quad \sum_{i=0}^{d} f_i \binom{k-1}{i} \quad = \quad \mathsf{H}_{\mathbb{K}[\Delta]}(k) \tag{5}$$

for all $k \in \mathbb{Z}_{>0}$. Simply put: the Ehrhart function of a unimodular geometric simplicial complex and the Hilbert function of the corresponding Stanley-Reisner ring coincide. This fact is well-known, see for example [MS05]. Taking the above approach and calculating the Ehrhart functions of open simplices, however, allows us to do without Möbius inversion.

For our purpose we need a more general concept than that of a Stanley-Reisner ring. For an abstract simplicial complex Δ the Hilbert function $\mathsf{H}_{\mathbb{K}[\Delta]}(k)$ counts all those monomials x^u of degree k with $\mathsf{supp}(u) \in \Delta$. We are interested in a pair of simplicial complexes $\Delta' \subset \Delta$, the former being a subcomplex of the latter, and want to count those monomials x^u such that $\mathsf{supp}(u) \notin \Delta'$ but $\mathsf{supp}(u) \in \Delta$. To that end we follow Stanley [Sta96] in calling a pair of simplicial complexes $\Delta' \subset \Delta$ a **relative simplicial complex**. We denote by $I_{\Delta/\Delta'}$ the ideal in $\mathbb{K}[\Delta]$ generated by all monomials x^u with $\mathsf{supp}(u) \notin \Delta'$. We call this the **relative Stanley-Reisner ideal**. Its Hilbert function $\mathsf{H}_{I_{\Delta/\Delta'}}(k)$ counts the number of non-zero monomials x^u of degree k in $I_{\Delta'} \setminus I_\Delta$ or, equivalently, the number of non-zero monomials x^u in $\mathbb{K}[x]$ with $\mathsf{supp}(u) \in \Delta \setminus \Delta'$. (Notice how the roles of Δ and Δ' swap, depending on whether we formulate the condition using ideals or using complexes). Now, as Stanley remarks, Theorem 2 carries over to the relative case.

Theorem 3 *[Sta96] Let $\Delta' \subset \Delta$ be a relative d-dimensional abstract simplicial complex and let f_i denote the number of i-dimensional simplices in $\Delta \setminus \Delta'$. Then for all $k \in \mathbb{Z}_{>0}$*

$$\mathsf{H}_{I_{\Delta/\Delta'}}(k) \quad = \quad \sum_{i=0}^{d} f_i \binom{k-1}{i}. \tag{6}$$

If Δ is a geometric simplicial complex and $\Delta' \subset \Delta$ a subcomplex, we also call the pair $\Delta' \subset \Delta$ a relative geometric simplicial complex and define its relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$ to be $I_{\mathsf{comb}(\Delta)/\mathsf{comb}(\Delta')}$.

By the same argument as above, we conclude that for any relative d-dimensional geometric simplicial complex $\Delta' \subset \Delta$, all faces of which are unimodular,

$$\mathsf{L}_{\bigcup \Delta \setminus \bigcup \Delta'}(k) = \sum_{\sigma \in \Delta \setminus \Delta'} \mathsf{L}_{\mathsf{relint}\,(\sigma)}(k) = \sum_{i=0}^{d} f_i \binom{k-1}{i} = \mathsf{H}_{I_{\Delta/\Delta'}}(k) \tag{7}$$

for all $k \in \mathbb{Z}_{>0}$, i.e. the Ehrhart function of a relative simplicial complex with unimodular faces and the Hilbert function of the associated relative Stanley-Reisner ideal coincide. Moreover this function is a polynomial in k as

$$\binom{k-1}{i} = \frac{1}{i!} \prod_{i=1}^{i} (k-i)$$

is a polynomial for every $i \in \mathbb{Z}_{\geq 0}$ using the convention that $i! = \prod_{j=1}^{i} j$ and empty products are 1.

To be able to deal with the applications in Section 5 we need to go one step further. The complexes we will be dealing with, are not going to be simplicial. Their faces will be polytopes. So we define a **relative polytopal complex** to be a pair $\mathcal{C}' \subset \mathcal{C}$ of polytopal complexes, the former a subcomplex of the latter. Our goal is to realize the Ehrhart function $\mathsf{L}_{\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'}(k)$ as the Hilbert function of a relative Stanley-Reisner ideal.

By the above arguments, it would suffice to require that $\mathcal{C}$ has a unimodular triangulation. But for the sake of convenience we would like to impose a condition on $\mathcal{C}$ that can be checked one face at a time. Requiring that each face of $\mathcal{C}$ has a unimodular triangulation would not be sufficient. A unimodular triangulation Δ_F for each face $F \in \mathcal{C}$ does not guarantee that $\bigcup_{F \in \mathcal{C}} \Delta_F$ is a unimodular triangulation of $\mathcal{C}$: It may be that for faces $F_1, F_2 \in \mathcal{C}$ that share a common face $F = F_1 \cap F_2$ the unimodular triangulations Δ_{F_1} and Δ_{F_2} do not agree on F, i.e.

$$\{F \cap f_1 \mid f_1 \in \Delta_{F_1}\} \neq \{F \cap f_2 \mid f_2 \in \Delta_{F_2}\}.$$

Fortunately there is the notion of a compressed polytope: it suffices to require of each face $F \in \mathcal{C}$ individually that F is compressed, to guarantee that $\mathcal{C}$ as a whole has a unimodular triangulation.

Let $P \subset \mathbb{R}^d$ be a lattice polytope. Let $\prec$ be a total ordering of the lattice points in P. The **pulling triangulation** $\mathsf{pull}(P; \prec)$ of P with respect to the total ordering $\prec$ is defined recursively as follows. If P is an empty simplex, then $\mathsf{pull}(P; \prec)$ is the complex generated by P. Otherwise $\mathsf{pull}(P; \prec)$ is the complex generated by the set of polytopes

$$\bigcup_F \{\mathsf{conv}\{v, G\} : G \in \mathsf{pull}(F; \prec)\}$$

where v is the $\prec$-minimal lattice point in P and the union runs over all faces F of P that do not contain v. See also Sturmfels [Stu96]. This construction yields a triangulation and the vertices of $\mathsf{pull}(P, \prec)$ are lattice points in P. Pulling triangulations need not be unimodular, in fact the simplices in a pulling triangulation do not even have to be empty! Polytopes whose pulling triangulation is always unimodular get a special name. A polytope P is **compressed** if for any total ordering $\prec$ on the vertex set the pulling triangulation $\mathsf{pull}(P, \prec)$ is unimodular. These definitions have the following well-known properties.

Proposition 4 1. *Any* $\mathsf{dim}(P)$*-dimensional simplex in* $\mathsf{pull}(P, \prec)$ *contains the* $\prec$*-minimal lattice point* v *in* P *as a vertex.*

2. $\mathsf{pull}(F, \prec) = \mathsf{pull}(P, \prec) \cap F$ *for any total order* $\prec$ *on the lattice points in* P *and any face* F *of* P.

3. *All faces of a compressed polytopes are compressed.*

4. *Compressed polytopes are empty.*

A proof of this proposition can be found in [Bre09]. For more information on pulling triangulations and compressed polytopes we refer to [Stu96, OH01, Sul04, Stu91].

Now, if $\mathcal{C}$ is a polytopal complex with integral vertices such that every face $P \in \mathcal{C}$ is compressed, then we can fix an arbitrary total order $\prec$ on $\bigcup \mathcal{C} \cap \mathbb{Z}^d$ and construct the pulling triangulations $\mathsf{pull}(P, \prec)$ of all faces $P \in \mathcal{C}$ with respect to that one global order $\prec$. By Proposition 4 this means that the for any two $P_1, P_2 \in \mathcal{C}$ that share a common face $F = P_1 \cap P_2$ the triangulations induced on F agree: $\mathsf{pull}(P_1, \prec) \cap F = \mathsf{pull}(F, \prec) = \mathsf{pull}(P_2, \prec) \cap F$. Thus $\mathsf{pull}(\mathcal{C}, \prec) := \bigcup_{F \in \mathcal{C}} \mathsf{pull}(F, \prec)$ is a unimodular triangulation of $\mathcal{C}$ with $\mathsf{vert}(\mathcal{C}) = \mathsf{vert}(\mathsf{pull}(\mathcal{C}, \prec))$. We abbreviate $\Delta := \mathsf{pull}(\mathcal{C}, \prec)$.

If $\mathcal{C}'$ is any subcomplex of $\mathcal{C}$, we define Δ' to be the subcomplex of Δ consisting of those faces $F \in \Delta$ such that $F \subset \bigcup \mathcal{C}'$. So

$$\mathsf{L}_{\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'}(k) = \mathsf{L}_{\bigcup \Delta \setminus \bigcup \Delta'}(k) = \mathsf{H}_{I_{\Delta/\Delta'}}(k) \tag{8}$$

for $k \in \mathbb{Z}_{>0}$ which means that we have realized the Ehrhart function of $\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'$ as the Hilbert function of the relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$. Moreover, we have already seen that this function is a polynomial. We summarize these results in the following theorem.

Theorem 5 *Let* $\mathcal{C}$ *be a polytopal complex. If all faces of* $\mathcal{C}$ *are compressed lattice polytopes, then for any subcomplex* $\mathcal{C}' \subset \mathcal{C}$ *there exists a relative Stanley-Reisner ideal* $I_{\Delta,\Delta'}$ *such that for all* $k \in \mathbb{Z}_{>0}$

$$\mathsf{L}_{\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'}(k) = \mathsf{H}_{I_{\Delta/\Delta'}}(k)$$

and this function is a polynomial.

5 Counting Polynomials as Hilbert Functions

In this section we apply Theorem 5 to obtain analogues of Steingrímsson's Theorem 1 for all five counting polynomials. A useful tool in this context is going to be the following theorem by Ohsugi and Hibi which states that lattice polytopes that are slices of the unit cube are automatically compressed.

Theorem 6 *(Ohsugi and Hibi [OH01]) Let* P *be a lattice polytope in* $\mathbb{R}^n$. *If* P *is lattice isomorphic to the intersection of an affine subspace with the unit cube, i.e.* $P \approx [0,1]^n \cap L$ *for some affine subspace* L, *then* P *is compressed.*(ii)

(ii) Actually, Ohsugi and Hibi showed a more general result, but this will suffice for our purposes. Interestingly, Sullivant [Sul04] noted that this condition is also necessary for a lattice polytope to be compressed.

Integral Flow and Tension Polynomials

Let S be the linear subspace of $\mathbb{R}^E$ given by (2). The k-flows of G are in bijection with the lattice points in $k \cdot (-1,1)^E \cap S$. Furthermore let $\mathcal{H} = \{\{x | x_e = 0\} \mid e \in E\}$ denote the arrangement of all coordinate hyperplanes. Then the nowhere zero k-flows of G are in bijection with the lattice points in $k \cdot (-1,1)^E \cap S \setminus \bigcup \mathcal{H}$. The closures of the components of $(-1,1)^E \cap S \setminus \bigcup \mathcal{H}$ are of the form $(\prod_{e \in E} [a_e, a_e + 1]) \cap S$ for some $a \in \{-1, 0\}^E$. Let $\mathcal{C}$ be the polytopal complex generated by these and let $\mathcal{C}'$ be the subcomplex of all faces of $\mathcal{C}$ contained in the boundary of $[-1,1]^E$ or contained in one of the coordinate hyperplanes. Then $\varphi_G(k) = \mathsf{L}_{\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'}(k)$. Because (2) gives rise to a totally unimodular matrix, the maximal faces of $\mathcal{C}$ are lattice polytopes.[iii] Moreover by Theorem 6, they are compressed. Thus Theorem 5 can be applied to yield the following result.

Theorem 7 *For any graph G there exists a relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$ such that for all $k \in \mathbb{Z}_{>0}$*

$$\varphi_G(k) = \mathsf{H}_{I_{\Delta/\Delta'}}(k).$$

The above geometric construction can be found in [BZ06b]. A similar construction given in [Dal08] can be used to show an analogue of the above theorem for the integral tension polynomial.

Theorem 8 *For any graph G there exists a relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$ such that for all $k \in \mathbb{Z}_{>0}$*

$$\theta_G(k) = \mathsf{H}_{I_{\Delta/\Delta'}}(k).$$

Modular Flow and Tension Polynomials

Let $v \in V$ and define the vector $a^v \in \{0, \pm 1\}^E$ by $a^v_e = +1$ if $\mathsf{head}(e) = v \neq \mathsf{tail}(e)$, $a^v_e = -1$ if $\mathsf{head}(e) \neq v = \mathsf{tail}(e)$, and $a^v_e = 0$ otherwise. Let A denote the matrix with the vectors a^v for $v \in V$ as rows. Now, we identify the integers $\{0, \dots, k-1\}$ with their respective cosets in $\mathbb{Z}_k$ so that a function $f : E \to \mathbb{Z}_k$ can be viewed as an integer vector $f \in [0,k)^E$. Using this identification all equations (2) hold for a given f if and only if $Af = kb$ for some $b \in \mathbb{Z}^V$. Thus the set of nowhere zero $\mathbb{Z}_k$-flows on G can be identified with the set of lattice points in $k \cdot ((0,1)^E \cap \bigcup_b H_b)$ where $H_b = \{f | Af = b\}$ and b ranges over all integer vectors such that $(0,1)^E \cap H_b \neq \emptyset$. Let $\mathcal{C}$ be the complex generated by the respective closed polytopes $[0,1]^E \cap H_b$. Let $\mathcal{C}'$ be the subcomplex consisting of all those faces that are contained in the boundary of the cube $[0,1]^E$. Then $\bar{\varphi}_G(k) = \mathsf{L}_{\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'}(k)$. The faces of $\mathcal{C}$ are lattice polytopes because A is totally unimodular and by Theorem 6, they are compressed. Thus Theorem 5 can be applied to yield the following result.

Theorem 9 *For any graph G there exists a relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$ such that for all $k \in \mathbb{Z}_{>0}$*

$$\bar{\varphi}_G(k) = \mathsf{H}_{I_{\Delta/\Delta'}}(k).$$

The above geometric construction can be found in [BS09], which also contains a similar construction using which an analogue of the above theorem for the modular tension polynomial can be shown [Bre09].

Theorem 10 *For any graph G there exists a relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$ such that for all $k \in \mathbb{Z}_{>0}$*

$$\bar{\theta}_G(k) = \mathsf{H}_{I_{\Delta/\Delta'}}(k).$$

[iii] We refer to [Sch86] for the concept of a totally unimodular matrix and related results.

Chromatic Polynomial

Let G be a graph without loops.(iv) For each $e \in E$ define $H_e = \{x \mid x_{\mathsf{head}(e)} = x_{\mathsf{tail}(e)}\}$ and consider the graphic hyperplane arrangement $\mathcal{H} = \{H_e \mid e \in E\}$. Then the proper k-colorings of G are in bijection with the lattice points in $k \cdot ([0,1)^E \setminus \bigcup \mathcal{H})$. The closure of any component C of $[0,1)^E \setminus \bigcup \mathcal{H}$ is of the form $P_\sigma = \{x \in [0,1]^V \mid \sigma_e(x_{\mathsf{head}(e)} - x_{\mathsf{tail}(e)}) \geq 0\}$ where $\sigma \in \{\pm 1\}^E$ is a sign vector. Let $\mathcal{C}$ be the polytopal complex generated by the P_σ and let $\mathcal{C}'$ be the subcomplex consisting of all faces that are contained in some hyperplane H_e or in some hyperplane of the form $\{x \mid x_v = 1\}$ for some $v \in V$. Then $\chi_G(k) = \mathsf{L}_{\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'}(k)$. The faces of $\mathcal{C}$ are lattice polytopes and it can be shown that they are compressed. Thus Theorem 5 can be applied to yield the following result.

Theorem 11 *For any graph G there exists a relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$ such that for all $k \in \mathbb{Z}_{>0}$*

$$\chi_G(k) = \mathsf{H}_{I_{\Delta/\Delta'}}(k).$$

This is an improvement upon Steingrímsson's Theorem insofar as we obtain the chromatic polynomial $\chi_G(k)$ itself as a Hilbert function of a relative Stanley-Reisner ideal and not the shifted polynomial $\chi_G(k+1)$. To obtain the shifted chromatic polynomial using the above construction we would need to consider the closed cube $[0,1]^V$ instead of the half-open cube $[0,1)^V$. A geometric construction similar to the one given above can be found in [BZ06a].

6 Non-Square-Free Ideals

What if the relative polytopal complex $\mathcal{C}$ does not have a unimodular triangulation? It turns out that if the polytopes in $\mathcal{C}$ are normal lattice polytopes, the Ehrhart function of $\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'$ is still the Hilbert function of a an ideal I_2 in a ring $\mathbb{K}[x]/I_1$, however we cannot guarantee that the ideals I_1, I_2 are square-free. That is, we are dealing with a relative multicomplex instead of a relative simplicial complex.

A lattice polytope P is **normal** if for every $k \in \mathbb{N}$ every $z \in kP \cap \mathbb{Z}^d$ can be written as the sum of k points in $P \cap \mathbb{Z}^d$. Note that a compressed polytope is automatically normal.

With this notion we can generalize Theorem 5 to include another case where the polytopal complex in question satisfies a weaker condition. The conclusion we obtain in this case is not as strong, however.

Theorem 12 *Let $\mathcal{C}$ be a polytopal complex in which all faces are normal lattice polytopes. Then for any subcomplex $\mathcal{C}' \subset \mathcal{C}$ there exist a monomial ideal I_1 in a polynomial ring $\mathbb{K}[x]$ equipped with the standard grading and a monomial ideal I_2 in $\mathbb{K}[x]/I_1$ such that for all $k \in \mathbb{Z}_{>0}$*

$$\mathsf{L}_{\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}'}(k) = \mathsf{H}_{I_2}(k)$$

and this function is a polynomial. If the faces of $\mathcal{C}$ are compressed, then moreover the ideals I_1 and I_2 can be chosen to be square-free.

The case where the faces of $\mathcal{C}$ are compressed and the ideals are square-free is just Theorem 5. We are not going to prove this again. Instead we give a self-contained algebraic proof of the case where the faces are only normal and we do not conclude that the ideals are square-free.

(iv) If G contains loops, then $\chi_G(k) = 0$ which, trivially, is a Hilbert function.

First we define the **polytopal Stanley-Reisner ideal** $I_{\mathcal{C}}$ of the polytopal complex $\mathcal{C}$ by

$$I_{\mathcal{C}} \quad := \quad \langle x^a \mid \text{there is no } P \in \mathcal{C} \text{ such that } \mathsf{supp}(a) \subset P \rangle$$

where again $\mathsf{supp}(a)$ denotes the set of lattice points u such that $a_u \neq 0$.

Proof: By homogenization, that is by passing to the complex generated by $\{P \times \{1\} | P \in \mathcal{C}\}$, we can assume without loss of generality that for every lattice point z there is at most one integer k such that $z \in k \bigcup \mathcal{C}$.

We are going to construct ideals $I_2 \supset I_1$ in a polynomial ring $\mathbb{K}[x]$ such that the monomials in $I_2 \setminus I_1$ of degree k are in bijection with the lattice points in $k(\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}')$.

Now consider the polynomial ring $\mathbb{K}[x_u : u \in \bigcup \mathcal{C} \cap \mathbb{Z}^d]$ equipped with the standard grading. Let $\prec$ be a term order on this polynomial ring. Let U be the matrix that has the vectors u as columns. Let n be the number of columns of U and let d be the number of rows. We define

$$\begin{aligned} I_1 \quad &:= \quad I_{\mathcal{C}} + \langle x^b \mid x^b \notin I_{\mathcal{C}} \text{ and there is an } x^a \notin I_{\mathcal{C}} \text{ such that } Ua = Ub \text{ and } x^a \prec x^b \rangle, \\ I_2 \quad &:= \quad I_{\mathcal{C}'} \end{aligned}$$

where $I_{\mathcal{C}}$ and $I_{\mathcal{C}'}$ denote the polytopal Stanley-Reisner ideals of the complexes $\mathcal{C}$ and $\mathcal{C}'$ respectively. We call a monomial x^a *valid* if $x^a \notin I_1$ but $x^a \in I_2$. Now we claim that the map $\pi : x^a \mapsto Ua$ defines a bijection between the valid monomials of degree k and the lattice points in $k(\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}')$.

If x^a is valid and of degree k, then $Ua \in k(\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}') \cap \mathbb{Z}^d$. First, we notice that $Ua \in \mathbb{Z}^d$, because a and U are integral. Second, we argue that $Ua \in k \bigcup \mathcal{C}$. Because x^a is valid, there exists a polytope P such that $\mathsf{supp}(a) \subset P$ and thus $Ua \subset kP$. Finally, we show that $Ua \notin k \bigcup \mathcal{C}'$. Suppose $Ua \in kP'$ for an inclusion-minimal $P' \in \mathcal{C}'$. Because $\mathcal{C}'$ is a subcomplex of $\mathcal{C}$, this implies that P' is a face of P and $\mathsf{supp}(a) \subset P'$. Hence $x^a \notin I_{\mathcal{C}'}$, which is a contradiction to x^a being valid.

π is surjective. Let $v \in k(\bigcup \mathcal{C} \setminus \bigcup \mathcal{C}') \cap \mathbb{Z}^d$ for some k. Then there is a polytope $P \in \mathcal{C} \setminus \mathcal{C}'$ such that $v \in \mathsf{relint}\ (kP)$. By the assumption that P is normal, there exists a non-negative integral representation b of v in terms of lattice points in $P \cap \mathbb{Z}^d$: $v = \sum_{u \in P \cap \mathbb{Z}^d} b_u u = Ub$. So by construction $x^b \notin I_{\mathcal{C}}$ and $Ub = v$. Consider the $\prec$-minimal monomial $x^a \notin I_{\mathcal{C}}$ with $Ua = Ub$. For this monomial we have $x^a \notin I_1$. Moreover, as $Ua = v \in \mathsf{relint}\ (kP)$ we have $x^a \in I_2$. Finally we have to check that $\deg(x^a) = k$. All elements of $\mathsf{supp}(a)$ are lattice points in P, so $Ua \in \deg(x^a)P$. However by our assumption at the beginning there is at most one integer k' such that $v = Ua \in k'P$. Thus $\deg(x^a) = k$.

$\hat{\pi}$ is injective. By definition of I_1 and as $\prec$ is a total order on the set of monomials, for every $v \in k(\bigcup \mathcal{C}) \cap \mathbb{Z}^d$ there is at most one monomial $x^a \notin I_1$ such that $Ua = v$. $\square$

We remark that this approach can also be used to give another proof of Theorem 5 using a fundamental correspondence between compressed polytopes and lattice point sets such that the corresponding toric ideal has square-free initial ideals under any reverse-lexicographic term order (see [Stu96]). This approach is explored in [Dal08] and [Bre09]. [Bre09] also contains a variant of the above result, due to Breuer and Sanyal, in the case where $\mathbb{K}[x]$ is equipped with a non-standard grading.

7 Bounds on the Coefficients

$\binom{k-1}{d}$ is a polynomial of degree d in k. The polynomials $\binom{k-1}{i}$ for $0 \leq i \leq d$ form a basis of the $\mathbb{K}$-vector space of all polynomials in $\mathbb{K}[k]$ of degree at most d and the polynomials $\binom{k-1}{i}$ for $0 \leq i$ form a basis of

$\mathbb{K}[k]$ when seen as a $\mathbb{K}$-vector space. By Theorem 3 the coefficients of the Hilbert function of a relative Stanley-Reisner ideal expressed with respect to this basis must be non-negative and integral. It turns out that this characterizes which polynomials appear as Hilbert functions of relative Stanley-Reisner ideals.

Theorem 13 *A polynomial $f(k) = \sum_{i=0}^{d} f_i \binom{k-1}{i}$ is the Hilbert function of some relative Stanley-Reisner ideal $I_{\Delta/\Delta'}$ if and only if $f_i \in \mathbb{Z}_{\geq 0}$ for all $0 \leq i \leq d$.*

Proof: We have already seen that the coefficients of $\mathsf{H}_{I_{\Delta/\Delta'}}(k)$ with respect to the basis $\binom{k-1}{d}$, $d \in \mathbb{Z}_{\geq 0}$ are necessarily non-negative integers. To see that this is also sufficient, let $f(k) = \sum_{i=0}^{d} f_i \binom{k-1}{i}$ with $f_i \in \mathbb{Z}_{\geq 0}$ for all $0 \leq i \leq d$. For $0 \leq i \leq d$ and $1 \leq j \leq f_i$ let σ_j^i denote a closed unimodular lattice simplex of dimension i in $\mathbb{R}^d$ such that the σ_j^i are pairwise disjoint. Let Δ denote the (disjoint) union of all these σ_j^i and define Δ' to be the union of the respective boundaries $\partial\sigma_j^i$. Then the set $\bigcup\Delta \setminus \bigcup\Delta'$ is the disjoint union of f_d relatively open unimodular lattice simplices of dimension d, f_{d-1} relatively open unimodular lattice simplices of dimension $d-1$ and so on. Consequently $\mathsf{H}_{I_{\Delta/\Delta'}}(k) = \mathsf{L}_{\bigcup\Delta\setminus\Delta'}(k) = f(k)$ as desired. $\square$

This immediately implies that all the counting functions we considered have non-negative integral coefficients with respect to this basis.

Theorem 14 *The modular and integral flow and tension polynomials as well as the chromatic polynomial of a graph have non-negative integer coefficients with respect to the basis $\{\binom{k-1}{d} | 0 \leq d \in \mathbb{Z}\}$ of $\mathbb{K}[k]$.*

For the modular and integral flow polynomials and the integral tension polynomial this is a new result, whereas for the chromatic polynomial and thus for the modular tension polynomial this is implicit in the previous work on the coloring complex. However, these bounds are not very strong, as we have not used any of the geometric information particular to these polynomials.

In [BD10] we give stronger constraints on the coefficients of the modular and integral flow and tension polynomials. There, we do not build on the realization of these polynomials as Hilbert functions, we rather make direct use of our geometric realization of these polynomials as Ehrhart functions of inside-out polytopes. We show that a wide class of inside-out polytopes, including those arising in these constructions, have a convex ear decomposition and use this fact to derive bounds.

Acknowledgements

We thank Matthias Beck and Christian Haase for suggesting this topic to us and for pointing us to the notion of compressed polytopes.

References

[BD10] Felix Breuer and Aaron Dall. Bounds on the coefficients of tension and flow polynomials. arxiv:math/1004.3470v1, 2010.

[BR07] Matthias Beck and Sinai Robins. *Computing the continuous discretely*. Undergraduate Texts in Mathematics. Springer, New York, 2007.

[Bre09] Felix Breuer. *Ham Sandwiches, Staircases and Counting Polynomials*. PhD thesis, Freie Universität Berlin, 2009.

[BS09] Felix Breuer and Raman Sanyal. Ehrhart theory, modular flow reciprocity, and the Tutte polynomial. arXiv:math/0907.0845v1, 2009.

[BZ06a] Matthias Beck and Thomas Zaslavsky. Inside-out polytopes. *Adv. Math.*, 205(1):134–162, 2006.

[BZ06b] Matthias Beck and Thomas Zaslavsky. The number of nowhere-zero flows on graphs and signed graphs. *J. Combin. Theory Ser. B*, 96(6):901–918, 2006.

[Dal08] Aaron Dall. The flow and tension complexes. Master's thesis, San Francisco State University, 2008.

[HRW98] Jürgen Herzog, Vic Reiner, and Volkmar Welker. The Koszul property in affine semigroup rings. *Pacific Journal of Mathematics*, 186(1):39–65, 1998.

[HS08] Patricia Hersh and Ed Swartz. Coloring complexes and arrangements. *Journal of Algebraic Combinatorics*, 27(2):205–214, 2008.

[Hul07] A Hultman. Link complexes of subspace arrangements. *European Journal of Combinatorics*, 28(3):781–790, 2007.

[Jon05] Jakob Jonsson. The topology of the coloring complex. *Journal of Algebraic Combinatorics*, 21:311–329, 2005.

[Koc02] Martin Kochol. Polynomials associated with nowhere-zero flows. *J. Combin. Theory Ser. B*, 84(2):260–269, 2002.

[MS05] Ezra Miller and Bernd Sturmfels. *Combinatorial Commutative Algebra*, volume 227 of *Graduate Texts in Mathematics*. Springer-Verlag, 2005.

[OH01] Hidefumi Ohsugi and Takayuki Hibi. Convex polytopes all of whose reverse lexicographic initial ideals are squarefree. *Proceedings of the American Mathematical Society*, 129(9):2541–2546, 2001.

[Sch86] Alexander Schrijver. *Theory of linear and integer programming*. Wiley-Interscience Series in Discrete Mathematics. John Wiley & Sons Ltd., Chichester, 1986. A Wiley-Interscience Publication.

[Sta96] Richard P. Stanley. *Combinatorics and Commutative Algebra Second Edition*, volume 41 of *Progress in Mathematics*. Birkhäuser, 1996.

[Ste01] Einar Steingrímsson. The coloring ideal and coloring complex of a graph. *Journal of Algebraic Combinatorics*, 14:73–84, 2001.

[Stu91] Bernd Sturmfels. Gröbner bases of toric varieties. *Tôhoku Math. Journal*, 43:249–261, 1991.

[Stu96] Bernd Sturmfels. *Gröbner Bases and Convex Polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, 1996.

[Sul04] Seth Sullivant. Compressed polytopes and statistical disclosure limitation. arXiv:math/0412535v1, 2004.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 432–442

The stability of the Kronecker product of Schur functions

Emmanuel Briand[1], Rosa Orellana[2] and Mercedes Rosas[3]

[1] *Departamento de Matemática Aplicada I, Universidad de Sevilla, Escuela Técnica Superior de Ingeniería Informática, Avda. Reina Mercedes S/N, 41012 Sevilla, Spain.*
[2] *Dartmouth College, Mathematics Department, 6188 Kemeny Hall, Hanover, NH, 03755, USA.*
[3] *Departamento de Álgebra, Universidad de Sevilla, Aptdo. de Correos 1160, 41080 Sevilla, Spain.*

Abstract. In the late 1930's Murnaghan discovered the existence of a stabilization phenomenon for the Kronecker product of Schur functions. For n large enough, the values of the Kronecker coefficients appearing in the product of two Schur functions of degree n do not depend on the first part of the indexing partitions, but only on the values of their remaining parts. We compute the exact value of n when this stable expansion is reached. We also compute two new bounds for the stabilization of a particular coefficient of such a product. Given partitions α and β, we give bounds for all the parts of any partition γ such that the corresponding Kronecker coefficient is nonzero. Finally, we also show that the reduced Kronecker coefficients are structure coefficients for the Heisenberg product introduced by Aguiar, Ferrer and Moreira.

Résumé. Dans les années 30 Murnaghan a découvert une propriété de stabilité pour le produit de Kronecker de fonctions de Schur. En degré assez grand, les valeurs des coefficients qui aparaissent dans le produit de Kronecker de deux fonctions de Schur ne dépendent pas de la première part des partitions en indice, mais seulement des parts suivantes. Dans ce travail nous calculons la valeur exacte du degré partir duquel ce développement stable est atteint. Nous calculons aussi deux nouvelles bornes supérieures pour la stabilisation d'un coefficient particulier d'un tel produit. Nous donnons en outre, pour α et β fixés, des bornes supérieures pour toutes les parts des partition γ rendant le coefficient de Kronecker d'indices α, β, γ non–nul. Finalement, nous identifions les coefficients de Kronecker réduits comme des constantes de structures pour le produit de Heisenberg de fonctions symétriques défini par Aguiar, Ferrer et Moreira.

Resumen. Hace poco más de 80 años Murnaghan descubrió un fenómeno de estabilidad para el producto de Kronecker de dos funciones de Schur. En grado suficientemente grande, los valores de los coeficientes de Kronecker que aparecen en el producto de Kronecker de dos funciones de Schur, no dependen de las primeras partes de las particiones que las indexan, sino solamente de sus demás partes. En este trabajo calculamos exactemente cuando este desarrollo estable esta alcanzado. También calculamos dos nuevas cotas para que cualquier familia dada de coeficientes de Kronecker se estabilice. Dadas dos particiones α y β, proporcionamos cotas superiores para todas las partes de cualquier partición γ tal que el coeficiente de Kronecker correspondiente no sea nulo. Finalmente, identificamos los coeficientes de Kronecker reducidos como constantes de estructura del producto de Heisenberg de funciones simétricas, introducido por Aguiar, Ferrer y Moreira.

Keywords: Symmetric functions, Kronecker coefficients

Introduction

The understanding of the *Kronecker coefficients of the symmetric group* $g_{\alpha,\beta}^{\gamma}$ (the multiplicities appearing when the tensor product of two irreducible representations of the symmetric group is decomposed into irreducibles; equivalently, the structural constants for the Kronecker product $*$ of symmetric functions in the basis of Schur functions, s_λ) is a longstanding open problem. Richard Stanley writes "One of the main problems in the combinatorial representation theory of the symmetric group is to obtain a combinatorial interpretation for the Kronecker coefficients" [30]. It is also a source of new challenges such as the problem of describing the set of non–zero Kronecker coefficients [28], a problem inherited from quantum information theory [18, 10]. Or proving that the positivity of a Kronecker coefficient can be decided in polynomial time, a problem posed by Mulmuley at the heart of his Geometric Complexity Theory [24] (see also the introductory paper by Bürgisser, Landsberg, Manivel and Weyman [2]).

In our work we study in more detail a remarkable stability property for the Kronecker products of Schur functions discovered by Murnaghan [26, 27]. This property is best shown on an example. Consider the Kronecker products $s_{(n-2,2)} * s_{(n-2,2)}$:

$$
\begin{aligned}
s_{2,2} * s_{2,2} &= s_4 + s_{1,1,1,1} + s_{2,2}\\
s_{3,2} * s_{3,2} &= s_5 + s_{2,1,1,1} + s_{3,2} + s_{4,1} + s_{3,1,1} + s_{2,2,1}\\
s_{4,2} * s_{4,2} &= s_6 + s_{3,1,1,1} + 2s_{4,2} + s_{5,1} + s_{4,1,1} + 2s_{3,2,1} + s_{2,2,2}\\
s_{5,2} * s_{5,2} &= s_7 + s_{4,1,1,1} + 2s_{5,2} + s_{6,1} + s_{5,1,1} + 2s_{4,2,1} + s_{3,2,2} + s_{4,3} + s_{3,3,1}\\
s_{6,2} * s_{6,2} &= s_8 + s_{5,1,1,1} + 2s_{6,2} + s_{7,1} + s_{6,1,1} + 2s_{5,2,1} + s_{4,2,2} + s_{5,3} + s_{4,3,1} + s_{4,4}\\
s_{7,2} * s_{7,2} &= s_9 + s_{6,1,1,1} + 2s_{7,2} + s_{8,1} + s_{7,1,1} + 2s_{6,2,1} + s_{5,2,2} + s_{6,3} + s_{5,3,1} + s_{5,4}
\end{aligned}
$$

And, actually, in all degree $n \geq 8$ we have the expansion:

$$s_{\bullet,2} * s_{\bullet,2} = s_\bullet + s_{\bullet,1,1,1} + 2s_{\bullet,2} + s_{\bullet,1} + s_{\bullet,1,1} + 2s_{\bullet,2,1} + s_{\bullet,2,2} + s_{\bullet,3} + s_{\bullet,3,1} + s_{\bullet,4}$$

For α partition and n integer, set $\alpha[n]$ for $(n-|\alpha|, \alpha_1, \alpha_2, \ldots)$. Murnaghan's general result is that for any partitions α and β, the expansions of $s_{\alpha[n]} * s_{\beta[n]}$ in the Schur basis all coincide for n big enough, except for the first part of the indexing partitions (which is determined by the degree n). This implies in particular that given any three partitions α, β and γ, the sequence of Kronecker coefficients $g_{\alpha[n]\beta[n]}^{\gamma[n]}$ is eventually constant. The *reduced Kronecker coefficient* $\overline{g}_{\alpha,\beta}^{\gamma}$ is defined as the stable value of this sequence. In our example, we see that $\overline{g}_{(2),(2)}^{(2)} = 2$ and $\overline{g}_{(2),(2)}^{(4)} = 1$.

When does a Kronecker product $s_{\alpha[n]} * s_{\beta[n]}$ stabilizes? When does a sequence of Kronecker coefficients $g_{\alpha[n]\beta[n]}^{\gamma[n]}$ becomes constant? Interestingly, these questions lead to look for linear inequalities fulfilled by the sets of triples of partitions (α, β, γ) whose corresponding reduced Kronecker coefficient $\overline{g}_{\alpha,\beta}^{\gamma}$ is non–zero. The analogous problem for Kronecker coefficients is of major importance, see [18, 28].

In view of the difficulty of studying the Kronecker coefficients, it is surprising to obtain theorems that hold in general. Regardless of this, we present new results of a general nature.

We find an elegant expression for the precise degree $n = \operatorname{stab}(\alpha, \beta)$ at which the expansion of the Kronecker product $s_{\alpha[n]} * s_{\beta[n]}$ stabilizes:

$$\operatorname{stab}(\alpha, \beta) = |\alpha| + |\beta| + \alpha_1 + \beta_1$$

Using Weyl's inequalities [35] for eigenvalues of triples of hermitian matrices fulfilling $A+B=C$, we find the maximum of γ_1 and upper bounds for all parts γ_k, among all γ in Supp$(\alpha,\beta)=\{\gamma : \overline{g}_{\alpha,\beta}^{\gamma}>0\}$.

Finally, we find upper bounds for the index $n=\operatorname{stab}(\alpha,\beta,\gamma)$ at which the sequence $g_{\alpha[n]\beta[n]}^{\gamma[n]}$ becomes constant, improving previously known bounds due to Brion [9] and Vallejo [34].

Detailed proofs for the results presented in this extended abstract can be found in [7].

1 Preliminaries

We assume that the reader is familar with the basic definitions in the theory of symmetric funcion, see [21] or [30].

Let λ be a partition of n. Let V_λ the irreducible representation of the symmetric group $\mathfrak{S}_n$ indexed by λ. The Kronecker coefficient $g_{\mu,\nu}^{\lambda}$ is the multiplicity of V_λ in the decomposition into irreducible representations of the tensor product $V_\mu \otimes V_\nu$. The Frobenius map identify V_λ with the Schur function s_λ. In doing so, it allows us to lift the tensor product of representations of the symmetric group to the setting of symmetric functions. Accordingly, the Kronecker coefficients $g_{\mu,\nu}^{\lambda}$ define the Kronecker product on symmetric functions by setting

$$s_\mu * s_\nu = \sum_\lambda g_{\mu,\nu}^{\lambda} s_\lambda.$$

We use the Jacobi-Trudi determinant to extend the definition of s_μ to the case where μ is any finite sequence of n integers :

$$s_\mu = \det\left(h_{\mu_j+i-j}\right)_{1\le i,j\le n}, \tag{1}$$

where h_k is the complete homogeneous symmetric function of degree k. In particular, $h_k=0$ if k is negative, and $h_0=1$. It is not hard to see that such a Jacobi–Trudi determinant s_μ is either zero or ± 1 times a Schur function.

The starting point of our investigations is a beautiful theorem of Murnaghan. Given a partition $\lambda=(\lambda_1,\lambda_2,\ldots)$ and an integer n, we denote by $\lambda[n]$ the sequence $(n-|\lambda|,\lambda_1,\lambda_2,\ldots)$. Notice that $\lambda[n]$ is a partition only if $n-|\lambda|\ge\lambda_1$.

Murnaghan Theorem (Murnaghan, [26, 27])**.** *There exists a family of non-negative integers $(\overline{g}_{\alpha\beta}^{\gamma})$ indexed by triples of partitions (α,β,γ) such that, for α and β fixed, only finitely many terms $\overline{g}_{\alpha\beta}^{\gamma}$ are nonzero, and for all $n\ge 0$,*

$$s_{\alpha[n]} * s_{\beta[n]} = \sum_\gamma \overline{g}_{\alpha\beta}^{\gamma} s_{\gamma[n]} \tag{2}$$

Moreover, the coefficient $\overline{g}_{\alpha\beta}^{\gamma}$ vanishes unless the weights of the three partitions fulfill the inequalities:

$$|\alpha|\le|\beta|+|\gamma|, \qquad |\beta|\le|\alpha|+|\gamma|, \qquad |\gamma|\le|\alpha|+|\beta|.$$

In what follows, we refer to these inequalities as *Murnaghan's inequalities*. We follow Klyachko [18] and call the coefficients $\overline{g}_{\alpha\beta}^{\gamma}$ the *reduced Kronecker coefficients*. An elegant proof of Murnaghan's Theorem, using vertex operators on symmetric functions, is given in [33].

Example 1. According to Murnaghan's theorem the reduced Kronecker coefficients determine the Kronecker product of two Schur functions, even for small values of n. For instance,

$$s_{2,2} * s_{2,2} = s_4 + s_{1,1,1,1} + 2s_{2,2} + s_{3,1} + s_{2,1,1} + 2s_{1,2,1} + s_{0,2,2} + s_{1,3} + s_{0,3,1} + s_{0,4}$$

The Jacobi-Trudi determinants corresponding to $s_{1,2,1}$ and $s_{0,2,2}$ have a repeated column, hence they are zero. On the other hand, it is easy to see that $s_{1,3} = -s_{2,2}$, $s_{0,3,1} = -s_{2,1,1}$, and $s_{0,4} = -s_{3,1}$. After taking into account the resulting cancellations, we recover the expression of the Kronecker product $s_{2,2} * s_{2,2}$ in the Schur basis: $s_4 + s_{1,1,1,1} + s_{2,2}$.

The reduced Kronecker coefficients contain the Littlewood–Richardson coefficients as special cases, as it was observed already by Murnaghan [27] and Littlewood [20]. Precisely, if $|\gamma| = |\alpha| + |\beta|$, then the reduced Kronecker coefficient $\overline{g}_{\alpha,\beta}^{\gamma}$ is equal to the Littlewood–Richardson coefficient $c_{\alpha,\beta}^{\gamma}$.

2 Recovering the Kronecker coefficients from reduced Kronecker coefficients

By definition, the reduced Kronecker coefficients are particular instances of Kronecker coefficients. We show that the reduced Kronecker coefficients contain enough information to recover exact value of the Kronecker coefficients. Let $u = (u_1, u_2, \ldots)$ be an infinite sequence and i a positive integer. Define $u^{\dagger i}$ as the sequence obtained from u by adding 1 to its $i-1$ first terms and erasing its i–th term:

$$u^{\dagger i} = (1 + u_1, 1 + u_2, \ldots, 1 + u_{i-1} + 1, u_{i+1}, u_{i+2}, \ldots)$$

Partitions are identified with infinite sequences by appending trailing zeros. Under this identification, when λ is a partition then so is $\lambda^{\dagger i}$ for all positive i.

Theorem 2.1 (Computing the Kronecker coefficients from the reduced Kronecker coefficients)**.** *Let n be a nonnegative integer and λ, μ, and ν be partitions of n. Then*

$$g_{\mu\nu}^{\lambda} = \sum_{i=1}^{\ell(\mu)\ell(\nu)} (-1)^{i+1} \overline{g}_{\bar{\mu}\bar{\nu}}^{\lambda^{\dagger i}} \tag{3}$$

3 The stabilization of the Kronecker products

Let us define here formally $\mathrm{stab}(\alpha, \beta)$. Let V be the linear operator on symmetric functions defined on the Schur basis by $V(s_\lambda) = s_{\lambda+(1)}$ for all partitions λ.

Definition ($\mathrm{stab}(\alpha,\beta)$)**.** Let α and β be partitions. Then $\mathrm{stab}(\alpha,\beta)$ is defined as the smallest integer n such that

$$s_{\alpha[n+k]} * s_{\beta[n+k]} = V^k \left(s_{\alpha[n]} * s_{\beta[n]} \right)$$

for all $k > 0$.

As an illustration see the example in the introduction where $\alpha = \beta = (2)$ and the Kronecker product is stable starting at $s_{(6,2)} * s_{(6,2)}$. Since $(6,2)$ is a partition of 8, we get that $\mathrm{stab}(\alpha,\beta) = 8$.

Theorem 3.1. *Let α and β be two partitions. Then*

$$\mathrm{stab}(\alpha, \beta) = |\alpha| + |\beta| + \alpha_1 + \beta_1.$$

To show that this theorem holds, we first reduce the calculation of $\mathrm{stab}(\alpha,\beta)$ to maximizing the linear form $|\gamma| + \gamma_1$ on $\mathrm{Supp}(\alpha, \beta)$

$$\mathrm{stab}(\alpha, \beta) = \max \left\{ |\gamma| + \gamma_1 \mid \gamma \text{ partition, } \overline{g}_{\alpha,\beta}^{\gamma} > 0 \right\}.$$

Then, we use the following formula that gives a decomposition of $\overline{g}_{\alpha\beta}^{\gamma}$ as a sum of nonnegative summands obtained from a formula due to Littlewood to show that $\max\left\{|\gamma| + \gamma_1 \mid \gamma \text{ partition}, \overline{g}_{\alpha,\beta}^{\gamma} > 0\right\} = |\alpha| + |\beta| + \alpha_1 + \beta_1$.

Let $c_{\alpha,\beta,\gamma}^{\delta}$ be the coefficient of s_δ in the product $s_\alpha s_\beta s_\gamma$.

Lemma 3.2. *Let α, β, γ be partitions. Then,*

$$\overline{g}_{\alpha,\beta}^{\gamma} = \sum g_{\delta,\epsilon}^{\zeta} c_{\delta,\sigma,\tau}^{\alpha} c_{\epsilon,\rho,\tau}^{\beta} c_{\zeta,\rho,\sigma}^{\gamma} \tag{4}$$

4 Row lengths for partitions indexing nonzero Kronecker coefficients.

In this section we give bounds for row lengths of partitions indexing nonzero Kronecker coefficients. We begin by reminding the reader about the powerful result:

Proposition 4.1 (Klemm [17], Dvir [13] Theorem 1.6, Clausen and Meier [11] Satz 1.1.)**.** *Let α and β be partitions of the same weight. Then,*

$$\max\left\{\gamma_1 \mid \gamma \text{ partition s. t. } g_{\alpha,\beta}^{\gamma} > 0\right\} = |\alpha \cap \beta|$$

where $\alpha \cap \beta = (\min(\alpha_1, \beta_1), \min(\alpha_2, \beta_2), \ldots)$.

Proposition 4.1 is the inspiration for some of the results in this section. Two closely related questions come to mind: First, can we prove an analogous result for the reduced Kronecker coefficients? Second, what can be said about the remaining parts of a partition γ such that $g_{\alpha,\beta}^{\gamma} > 0$ (or similarly, such that $\overline{g}_{\alpha,\beta}^{\gamma} > 0$)?

We answer the first question in the affirmative by showing that

Theorem 4.2. *Let α and β be partitions. Then,*

$$\max\left\{\gamma_1 \mid \gamma \text{ partition}, \overline{g}_{\alpha,\beta}^{\gamma} > 0\right\} = |\alpha \cap \beta| + \max(\alpha_1, \beta_1) \tag{5}$$

We also obtained a set of bounds for the remaing parts of such a γ using Weyl's inequalities triples of spectra of hermitian matrices fulfilling $A + B = C$ [35]. This bounds are known to hold as well for the indices of the non–zero Littlewood–Richardson coefficients (see for instance [14]).

Theorem 4.3. *Let α and β be partitions. If $\overline{g}_{\alpha,\beta}^{\gamma} > 0$, then, for all positive integers i, j, we have that*

$$\gamma_{i+j-1} \leq |E_i\alpha \cap E_j\beta| + \alpha_i + \beta_j \tag{6}$$

where $E_k\lambda$ stands for the partition obtained from λ by erasing its k–th part.

Finally, combining Murnaghan's inequalities with Proposition 4.1 we obtain

$$\max\left\{|\gamma| \mid \gamma \text{ partition}, \overline{g}_{\alpha,\beta}^{\gamma} > 0\right\} = |\alpha| + |\beta|,$$
$$\min\left\{|\gamma| \mid \gamma \text{ partition}, \overline{g}_{\alpha,\beta}^{\gamma} > 0\right\} = \max(|\alpha|, |\beta|) - |\alpha \cap \beta|.$$

The first equality readily implies that

$$\gamma_k \leq \frac{|\alpha| + |\beta|}{k}. \tag{7}$$

Example 2. Let $\alpha = (2)$ and $\beta = (4,3,2)$, then the first row of the table are the nonzero values of γ_k and the second row are the values predicted by equations (6) and (7):

k	1	2	3	4	5
max values for γ_k	6	4	3	2	1
bound for γ_k	6	5	3	2	2

In the case that $\alpha = (3,1)$ and $\beta = (2,2)$ we get

k	1	2	3	4	5	6
max values for γ_k	6	3	2	1	1	1
bound for γ_k	6	4	2	2	1	1

These bounds also provide bounds for the non–zero Kronecker coefficients. Indeed, Michel Brion [9] showed that for any given α, β and γ, the sequence of the Kronecker coefficients $g_{\alpha[n],\beta[n]}^{\gamma[n]}$ is weakly increasing. As a consequence, $\overline{g}_{\alpha,\beta}^{\gamma}$ is non–zero whenever $g_{\alpha[n],\beta[n]}^{\gamma[n]}$ is non–zero for some n.

5 The stabilization of the Kronecker coefficients

In this section we study of a weaker version of the stabilization problem. One consequence of Murnaghan's Theorem is that each particular sequence of Kronecker coefficients $g_{\alpha[n],\beta[n]}^{\gamma[n]}$ stabilizes to $\overline{g}_{\alpha,\beta}^{\gamma}$, possibly before n reaches $\mathrm{stab}(\alpha,\beta)$.

Definition ($\mathrm{stab}(\alpha,\beta,\gamma)$)**.** Let α, β, γ be partitions. Then $\mathrm{stab}(\alpha,\beta,\gamma)$ is defined as the the smallest integer N such that the sequences $\alpha[N]$, $\beta[N]$ and $\gamma[N]$ are partitions and $g_{\alpha[n],\beta[n]}^{\gamma[n]} = \overline{g}_{\alpha,\beta}^{\gamma}$ for all $n \geq N$.

Two bounds have already been found for $stab(\alpha,\beta,\gamma)$ by Brion [9] and Vallejo [34]. Brions' and Vallejo's bounds, respectively, are

$$M_B(\alpha,\beta;\gamma) = |\alpha| + |\beta| + \gamma_1,$$
$$M_V(\alpha,\beta;\gamma) = |\gamma| + \begin{cases} \max\{|\alpha| + \alpha_1 - 1, |\beta| + \beta_1 - 1, |\gamma|\} & \text{if } \alpha \neq \beta \\ \max\{|\alpha| + \alpha_1, |\gamma|\} & \text{if } \alpha = \beta \end{cases}$$

Our first contribution is the following Lemma which describes a general technique for producing linear upper bounds for $\mathrm{stab}(\alpha,\beta,\gamma)$.

Lemma 5.1. *Let f be a function on triples of partitions such that for all i,*

$$f(\alpha,\beta,\gamma^{\dagger 1}) \geq f(\alpha,\beta,\gamma^{\dagger i}).$$

Set $\mathcal{M}_f(\alpha,\beta,\gamma) = |\gamma| + f(\alpha,\beta,\bar{\gamma})$ and assume also that whenever $\overline{g}_{\alpha,\beta}^{\gamma} > 0$,

$$\mathcal{M}_f(\alpha,\beta,\gamma) \geq \max\left(|\alpha| + \alpha_1, |\beta| + \beta_1, |\gamma| + \gamma_1\right). \tag{8}$$

Then whenever $\overline{g}_{\alpha,\beta}^{\gamma} > 0$,

$$\mathrm{stab}(\alpha,\beta,\gamma) \leq \mathcal{M}_f(\alpha,\beta,\gamma).$$

Three functions f such that (8) holds have already appeared in this paper. Each one gives a bound for $\mathrm{stab}(\alpha, \beta, \gamma)$.

1. Murnaghan's triangle inequalities and Theorem 4.2 imply that (8) holds for $f(\alpha, \beta, \tau) = |\alpha| + |\beta| - |\tau|$. Using our lemma, we recover Brion's bound.

2. From Theorem 4.2 we obtain that (8) holds for $f(\alpha, \beta, \tau) = |\bar{\alpha} \cap \bar{\beta}| + \alpha_1 + \beta_1$. In this situation, we obtain that $M_1(\alpha, \beta, \gamma) = |\gamma| + |\bar{\alpha} \cap \bar{\beta}| + \alpha_1 + \beta_1$. From the symmetry of the Kronecker coefficients, we conclude that $\mathrm{stab}(\alpha, \beta, \gamma) \leq N_1(\alpha, \beta, \gamma)$ where

$$N_1(\alpha, \beta, \gamma) = \min(M_1(\alpha, \beta, \gamma), M_1(\beta, \gamma, \alpha), M_1(\gamma, \alpha, \beta))$$

   We have shown that N_1 improves both the bounds of Vallejo and of Brion.

3. Finally, Theorem 3.1 shows that (8) holds for $f(\alpha, \beta, \tau) = 1/2\,(|\alpha| + |\beta| + \alpha_1 + \beta_1 - |\tau|)$. Then $\mathrm{stab}(\alpha, \beta, \gamma) \leq N_2(\alpha, \beta, \gamma)$, where

$$N_2(\alpha, \beta, \gamma) = \left[\frac{|\alpha| + |\beta| + |\gamma| + \alpha_1 + \beta_1 + \gamma_1}{2} \right]$$

   where $[x]$ denotes the integer part of x.

We conclude this section by applying our bounds to some interesting examples of Kronecker coefficients appearing in the literature.

Example 3 (The Kronecker coefficients indexed by three hooks)**.** Our first example looks at the elegant situation where the three indexing partitions are hooks. Note that after deleting the first part of a hook we always obtain a one column shape. Let $\alpha = (1^e)$, $\beta = (1^f)$ and $\gamma = (1^d)$ be the reduced partitions, with d, e and f positive. In Theorem 3 of [29], it was shown that Murnaghan's inequalities describe the stable value of the Kronecker coefficient $g_{\alpha[n], \beta[n]}^{\gamma[n]}$,

$$\overline{g}_{\alpha,\beta}^{\gamma} = ((e \leq d + f))((d \leq e + f))((f \leq e + d))$$

where $((P))$ equals 1 if the proposition is true, and 0 if not.

Moreover, $\mathrm{stab}(\alpha, \beta, \gamma)$ was actually computed in the proof of Theorem 3 [29]. It was shown that the Kronecker coefficient equals 1 if and only if Murnaghan's inequalities hold, as well as the additional inequality $e + f \leq d + 2(n - d) - 2$. This last inequality says that:

$$\mathrm{stab}(\alpha, \beta, \gamma) = \left[\frac{d + e + f + 3}{2} \right] = N_2(\alpha, \beta, \gamma)$$

To summarize, for triples of hooks, Murnaghan's inequalities govern the value of the reduced Kronecker coefficients, and N_2 is a sharp bound. On the other hand, the bounds provided by N_1, N_B, and N_V are not in general sharp.

□

Example 4 (The Kronecker coefficients indexed by two two-row shapes)**.** After deleting the first part of a two-row partition we obtain a partition of length 1. Let α and β be one-row partitions. We have:

$$N_1(\alpha,\beta,\gamma) = \alpha_1 + \beta_1 + \gamma_1$$
$$N_2(\alpha,\beta,\gamma) = \alpha_1 + \beta_1 + \gamma_1 + \left[\frac{\gamma_2+\gamma_3}{2}\right]$$

It follows from [8] that when $\overline{g}^{\gamma}_{\alpha,\beta} > 0$,

$$\mathrm{stab}(\alpha,\beta,\gamma) = \gamma_1 - \gamma_3 + \alpha_1 + \beta_1.$$

Neither N_1 nor N_2 are sharp bounds. Indeed, for $\overline{g}^{\gamma}_{\alpha,\beta} > 0$ we have $\mathrm{stab}(\alpha,\beta,\gamma) < N_1$ if $\gamma_3 > 0$, and $\mathrm{stab}(\alpha,\beta,\gamma) < N_2$ if $\gamma_2 > 0$.

Moreover, $N_1 < N_2$ when $\gamma_2 + \gamma_3 > 1$.

□

Example 5 (The Kronecker coefficients: One of the partitions is a two-row shape)**.** The case when γ has only one row, $\gamma = (p)$, was studied in [4]. It was shown there (Theorem 5.1) that

$$\mathrm{stab}(\alpha,\beta,(p)) \leq |\alpha| + \alpha_1 + 2\,p.$$

6 Further remarks on the reduced Kronecker coefficients

There is strong evidence to believe that the reduced Kronecker coefficients are better behaved than the Kronecker coefficients and in some sense easier to study.

The saturation theorem of Terence Tao and Allen Knutson, imply that deciding whether a Littlewood-Richardson coefficient is positive can be done in polynomial time [23, 19, 12]. On the other hand, it is known that the Kronecker coefficients do not satisfy the saturation property. For example,

$$g^{(n,n)}_{(n,n),(n,n)} = 0 \text{ if } n \text{ is odd, but } g^{(n,n)}_{(n,n),(n,n)} = 1 \text{ if } n \text{ is even.}$$

This suggests that the Kronecker coefficients are harder to compute.

On the other hand, the reduced Kronecker coefficients are conjectured to satisfy the saturation property by Klyachko and Kirillov, [18, 16], and in a stronger form by King [15].

We believe that the study of the reduced Kronecker coefficients $\overline{g}^{\lambda}_{\mu,\nu}$ will lead to a better understanding of the Kronecker coefficients.

This paper is part of a series [6, 8] that studies the reduced Kronecker coefficients. Theorem 2.1 first appeared in [5], where it was used to compute the first explicit piecewise quasipolynomial description for the Kronecker coefficients indexed by two two–row shapes. That description was then used in [8] to test several conjectures of Mulmuley. As a result, we found counterexamples [6] for the strong version of his SH conjecture [24] on the behavior of the Kronecker coefficients under stretching of its indices. As pointed out by Ron King [15], our counterexample also implies that $Q^{\nu}_{\lambda,\mu}(t) = g^{t\nu}_{t\lambda,t\mu}$ is not an Ehrhart quasipolynomial. Therefore $Q^{\nu}_{\lambda,\mu}(t)$ can not count the number of integral points in any rational complex polytope.

We have also found a very interesting connection to another product $\#$ on symmetric functions, introduced by Aguiar, Ferrer and Moreira [1, 25] under the names *smash* or *Heisenberg product*, and independently, yet less explicitely, by Scharf, Thibon and Wybourne [32]. It fulfills

$$f\#g = \sum f_1 \cdot (f_2 * g_1) \cdot g_2$$

where $\Delta(f) = \sum f_1 \otimes f_2$ and $\Delta(g) = \sum g_1 \otimes g_2$ are in Sweedler's notation.

We have shown that the reduced Kronecker coefficients are the structure constants for this product in the basis $\{s_\lambda[\mathbb{X}-1]\}$ (the Schur functions at the alphabet $\mathbb{X}-1$, in the λ–ring notation). That is,

$$s_\alpha[\mathbb{X}-1]\#s_\beta[\mathbb{X}-1] = \sum_\gamma \overline{g}_{\alpha,\beta}^{\gamma} s_\gamma[\mathbb{X}-1].$$

At this point, we hope that the reader is convinced that the reduced Kronecker coefficients are interesting objects on their own.

Acknowledgments

We thank Ernesto Vallejo for pointing to us the reference [9], Ron King for pointing to us Littlewood's formula, and Richard Stanley for suggesting to look at [16]. We also thank John Stembridge for making freely available his Maple package SF [31].

References

[1] Marcelo Aguiar, Walter Ferrer, and Walter Moreira. The smash product of symmetric functions. Extended abstract. ArXiv:math.CO/0412016, 2004.

[2] Peter Bürgisser, Joseph Landsberg, Laurent Manivel, and Jerzy Weyman. An overview of mathematical issues arising in the Geometric complexity theory approach to VP v.s. VNP. ArXiv:0907.2850, 2009.

[3] P. H. Butler and R. C. King. The symmetric group: characters, products and plethysms. *J. Mathematical Phys.*, 14:1176–1183, 1973.

[4] Cristina M. Ballantine and Rosa C. Orellana. On the Kronecker product $s_{(n-p,p)} * s_\lambda$. *Electronic Journal of Combinatorics*, 12:#R28, 26 pp. (electronic), 2005.

[5] Emmanuel Briand, Rosa Orellana, and Mercedes Rosas. Quasipolynomial formulas for the Kronecker coefficients indexed by two two–row shapes (extended abstract). ArXiv:0812.0861v1, 2008.

[6] Emmanuel Briand, Rosa Orellana, and Mercedes Rosas. Reduced Kronecker coefficients and counter–examples to Mulmuley's conjecture SH. ArXiv:0810.3163, 2008.

[7] Emmanuel Briand, Rosa Orellana, and Mercedes Rosas. On the stability of the Kronecker products of Schur functions. ArXiv:0907.4652, 2009.

[8] Emmanuel Briand, Rosa Orellana, and Mercedes Rosas. Quasipolynomial formulas for the Kronecker coefficients indexed by two two–row shapes. In preparation.

[9] Michel Brion. Stable properties of plethysm: on two conjectures of Foulkes. *Manuscripta Mathematica*, 80:347–371, 1993.

[10] Matthias Christandl, Aram W. Harrow, and Graeme Mitchison. Nonzero Kronecker coefficients and what they tell us about spectra. *Comm. Math. Phys.*, 270(3):575–585, 2007.

[11] Michael Clausen and Helga Meier. Extreme irreduzible Konstituenten in Tensordarstellungen symmetrischer Gruppen. *Bayreuth. Math. Schr.*, 45:1–17, 1993.

[12] Jesús A. De Loera and Tyrrell B. McAllister. On the computation of Clebsch-Gordan coefficients and the dilation effect. *Experiment. Math.*, 15(1):7–19, 2006.

[13] Yoav Dvir. On the Kronecker product of S_n characters. *J. Algebra*, 154(1):125–140, 1993.

[14] William Fulton. Eigenvalues, invariant factors, highest weights, and Schubert calculus. *Bull. Amer. Math. Soc. (N.S.)*, 37(3):209–249 (electronic), 2000.

[15] Ron King Some remarks on characters of symmetric groups, Schur functions, Littlewood-Richardson and Kronecker coefficients *Workshop on the Mathematical Foundations of Quantum Information, Universidad de Sevilla* , November 23-27 2009. `http://congreso.us.es/enredo2009/Workshop.html`

[16] Anatol N. Kirillov. An invitation to the generalized saturation conjecture. *Publ. Res. Inst. Math. Sci.*, 40(4):1147–1239, 2004.

[17] Michael Klemm. Tensorprodukte von Charakteren der symmetrischen Gruppe. *Arch. Math. (Basel)*, 28(5):455–459, 1977.

[18] Alexander Klyachko. Quantum marginal problem and representations of the symmetric group. ArXiv:quant-ph/0409113, september 2004.

[19] Allen Knutson and Terence Tao. Honeycombs and sums of Hermitian matrices. *Notices Amer. Math. Soc.*, 48(2):175–186, 2001.

[20] D. E. Littlewood. Products and plethysms of characters with orthogonal, symplectic and symmetric groups. *Canad. J. Math.*, 10:17–32, 1958.

[21] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[22] Laurent Manivel. On rectangular Kronecker coefficients. arXiv:0907.3351v1, July 2009.

[23] Ketan D. Mulmuley and Milind Sohoni. Geometric complexity theory III: on deciding positivity of Littlewood–Richardson coefficients. ArXiv:cs.CC/0501076, January 2005.

[24] Ketan D. Mulmuley. Geometric complexity theory VI: the flip via saturated and positive integer programming in representation theory and algebraic geometry. Technical Report TR–2007–04, Computer Science Department, The University of Chicago, may 2007. (GCT6). Available as ArXiv:cs/0704.0229 and at `http://ramakrishnadas.cs.uchicago.edu`. Revised version to be available here.

[25] Walter Moreira Products of representations of the symmetric group and non-commutative versions. Ph.D. Thesis Texas A&M University, 2008.

[26] Francis D. Murnaghan. The Analysis of the Kronecker Product of Irreducible Representations of the Symmetric Group. *Amer. J. Math.*, 60(3):761–784, 1938.

[27] Francis D. Murnaghan. On the analysis of the Kronecker product of irreducible representations of S_n. *Proc. Nat. Acad. Sci. U.S.A.*, 41:515–518, 1955.

[28] Luke Oeding. Report on "Geometry and representation theory of tensors for computer science, statistics and other areas.". ArXiv:0810.3940, 2008.

[29] Mercedes H. Rosas. The Kronecker product of Schur functions indexed by two–row shapes or hook shapes. *Journal of algebraic combinatorics*, 14(2):153–173, 2001. Previous version at ArXiv:math/0001084.

[30] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[31] John Stembridge. A Maple package for symmetric functions. *Journal of Symbolic Computation*, 20:755–768, 1995. The Maple package SF is available at `http://www.math.lsa.umich.edu/~jrs/maple.html#SF`.

[32] Thomas Scharf, Jean-Yves Thibon, and Brian G. Wybourne. Reduced notation, inner plethysms and the symmetric group. *J. Phys. A*, 26(24):7461–7478, 1993.

[33] Jean-Yves Thibon. Hopf algebras of symmetric functions and tensor products of symmetric group representations. *Internat. J. Algebra Comput.*, 1(2):207–221, 1991.

[34] Ernesto Vallejo. Stability of Kronecker products of irreducible characters of the symmetric group. *Electronic journal of combinatorics*, 6(1):1–7, 1999.

[35] Hermann Weyl. Das asymtotische Verteilungsgesetz der Eigenwerte lineare partialler Differentialgleichungen. *Mathematische Annalen*, 71:441–479, 1912.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 443–454

The Geometry of Lecture Hall Partitions and Quadratic Permutation Statistics

Katie L. Bright[1] and Carla D. Savage[1†]

[1] *Computer Science, North Carolina State University, Box 8206, Raleigh, NC 27695, USA. (*`klbrigh2@ncsu.edu`, `savage@csc.ncsu.edu`*)*

Abstract. We take a geometric view of lecture hall partitions and anti-lecture hall compositions in order to settle some open questions about their enumeration. In the process, we discover an intrinsic connection between these families of partitions and certain quadratic permutation statistics. We define some unusual quadratic permutation statistics and derive results about their joint distributions with linear statistics. We show that certain specializations are equivalent to the lecture hall and anti-lecture hall theorems and another leads back to a special case of a Weyl group generating function that "ought to be better known."

Résumé. Nous regardons géométriquement les partitions amphithéâtre et les compositions planétarium afin de résoudre quelques questions énumératives ouvertes. Nous découvrons un lien intrinsèque entre ces familles des partitions et certaines statistiques quadratiques de permutation. Nous définissons quelques statistiques quadratiques peu communes des permutations et dérivons des résultats sur leurs distributions jointes avec des statistiques linéaires. Nous démontrons que certaines spécialisations sont équivalentes aux théorèmes amphithéâtre et planétarium. Une autre spécialisation mène à un cas spécial de la série génératrice d'un groupe de Weyl qui "devrait être mieux connue".

Keywords: lecture hall partitions, anti-lecture hall compositions, permutation statistics, lattice point enumeration, generating functions

1 Introduction

A *lecture hall partition* of length n is an integer sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ [BME97] satisfying

$$0 \leq \frac{\lambda_1}{1} \leq \frac{\lambda_2}{2} \leq \ldots \leq \frac{\lambda_n}{n}.$$

An *anti-lecture hall composition* of length n is an integer sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ [CS03] satisfying

$$\frac{\lambda_1}{1} \geq \frac{\lambda_2}{2} \geq \ldots \geq \frac{\lambda_n}{n} \geq 0.$$

[†]This work was supported in part by NSA

These intriguing combinatorial objects and their various generalizations have been the subject of several papers and they have been shown to be related to Bott's formula in the theory of affine Coxeter groups [BME97, BME99], Euler's partition theorem [BME97, Yee01, SY08], the Gaussian polynomials [CLS07, CS04], the q-Chu-Vandermonde Identities [CLS07, CS04], the q-Gauss summation [ACS09], and the little Göllnitz partition theorems [CSS09]. In this paper we regard them from the point of view of lattice point enumeration and uncover several new results and connections.

The set $\mathbb{Z}^n \subset \mathbb{R}^n$ is the n-dimensional integer lattice and its elements are called lattice points. $\mathbb{Z}^n_{\geq 0}$ denotes the set of lattice points with all coordinates nonnegative. So, lecture hall partitions and anti-lecture hall compositions of length n can be viewed as lattice points in $\mathbb{Z}^n_{\geq 0}$.

Let L_n be the set of lecture hall partitions of length n and A_n, the set of anti-lecture hall compositions of length n. Define the subsets $L_n^{(t)}$ and $A_n^{(t)}$ by the constraints:

$$L_n^{(t)}: \quad 0 \leq \frac{\lambda_1}{1} \leq \frac{\lambda_2}{2} \leq \ldots \leq \frac{\lambda_n}{n} \leq t$$

and

$$A_n^{(t)}: \quad t \geq \frac{\lambda_1}{1} \geq \frac{\lambda_2}{2} \geq \ldots \geq \frac{\lambda_n}{n} \geq 0.$$

The following was shown in [CLS07]

Theorem 1.1 *For integer $t \geq 0$,*

$$\left| L_n^{(t)} \right| = (t+1)^n = \left| A_n^{(t)} \right|.$$

Let Q_t^n denote the lattice points in the n-dimensional cube of width t:

$$Q_t^n = \{(x_1, x_2, \ldots, x_n) \in \mathbb{Z}^n_{\geq 0} \mid \ 0 \leq x_i \leq t, \ \ 1 \leq i \leq n\}.$$

Matthias Beck observed [Bec09] that since also $|Q_t^n| = (t+1)^n$, there should be some natural bijections with $L_n^{(t)}$ and $A_n^{(t)}$.

In Section 2, we prove two simple bijections

$$\Theta : \mathbb{Z}^n_{\geq 0} \to L_n$$

and

$$\Phi : \mathbb{Z}^n_{\geq 0} \to A_n$$

with the property that for every $t \geq 0$,

$$\Theta^{-1}(L_n^{(t)}) = Q_n^t = \Phi^{-1}(A_n^{(t)}).$$

Previously, a bijection between $L_n^{(t)}$ and $A_n^{(t)}$ was known [CLS07], but it depended on t, it did not extend to L_n and A_n, and it did not explain the relationship between their generating functions. In contrast, a new bijection $L_n \to A_n$ reveals the functional relationship between their generating functions and restricts to a bijection between $L_n^{(t)}$ and $A_n^{(t)}$. What emerges is a characterization of L_n and A_n in terms of (new) permutation statistics.

In Section 3, we use the bijections Θ and Φ to derive generating functions for L_n and A_n in terms of permutation statistics and show how to derive one from the other. Similar ideas underlie the computation of the refined generating function for L_n in [BME99] and A_n in [CS04], but the connection with permutation statistics, a key ingredient in the relationship between L_n and A_n, was missed.

In Section 4, we show how the generating functions derived in Section 3 imply new results about distributions of quadratic permutation statistics and connections with affine Coxeter groups.

2 The bijections

The bijections between points in the cube and the lecture hall partitions and anti-lecture hall compositions have simple descriptions in terms of permutations and their inversion sequences, so we first review some notation and results.

2.1 Permutation statistics and stable sorting

Let S_n be the set of permutations of $\{1, 2, \ldots, n\}$. For $\pi \in S_n$, an *inversion* of π is a pair (i, j) such that $i < j$, but $\pi_i > \pi_j$. The number of inversions of π is denoted $\mathrm{inv}(\pi)$. A *descent* of π is a position i such that $1 \leq i < n$ and $\pi_i > \pi_{i+1}$. The set of all descents of π is denoted $\mathrm{Des}(\pi)$ and its size is $\mathrm{des}(\pi) = |\mathrm{Des}(\pi)|$.

Define the *inversion sequence* of π as the sequence $\epsilon(\pi) = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$, where ϵ_i is the number of elements of $\{1, \ldots, n\}$ to the right of i, in π, which are smaller than i. Then $\mathrm{inv}(\pi) = \epsilon_1 + \epsilon_2 + \cdots + \epsilon_n$.

It is well-known ([Knu73], p. 12) that the mapping $\pi \to \epsilon(\pi)$ is a bijection between S_n and integer sequences I_n, where

$$I_n = \{(\epsilon_1, \epsilon_2, \ldots, \epsilon_n) \mid 0 \leq \epsilon_i < i\}.$$

For $\pi \in S_n$, although in general $\epsilon(\pi) \neq \epsilon(\pi^{-1})$, it is known that ([Knu73], p. 14-15):

$$\mathrm{inv}(\pi) = \mathrm{inv}(\pi^{-1}). \tag{1}$$

A permutation $\pi \in S_n$ *stably sorts* a sequence $s = (s_1, \ldots, s_n)$ into weakly increasing order if

$$s_{\pi_1} \leq s_{\pi_2} \leq \ldots \leq s_{\pi_n}$$

and if $i \in \mathrm{Des}(\pi)$ then $s_{\pi_i} < s_{\pi_{i+1}}$, that is, equal elements of s retain their relative order. For every sequence s of length n there is a unique $\pi \in S_n$ such that π stably sorts s into weakly increasing order.

Let $(w_1 \leq w_2 \leq \ldots \leq w_n)$ denote a weakly increasing sequence and $(w_1 \geq w_2 \geq \ldots \geq w_n)$ a weakly decreasing sequence. For a sequence $s = (s_1, \ldots, s_n)$ and $\pi \in S_n$, define $\pi(s)$ by $\pi(s) = (s_{\pi_1}, s_{\pi_2}, \ldots, s_{\pi_n})$.

Define

$$S_n/(w_1 \leq w_2 \leq \ldots \leq w_n) = \{\pi \in S_n \mid \text{if } i \in \mathrm{Des}(\pi^{-1}) \text{ then } w_i < w_{i+1}\}$$

and

$$S_n/(w_1 \geq w_2 \geq \ldots \geq w_n) = \{\pi \in S_n \mid \text{if } i \in \mathrm{Des}(\pi^{-1}) \text{ then } w_i > w_{i+1}\}$$

Informally, $\pi \in S_n/w$ iff π^{-1} is the unique permutation in S_n that stably sorts $\pi(w)$ into w.

Define

$$I_n/w = \{\epsilon \in I_n \mid \text{if } w_i = w_{i+1} \text{ then } \epsilon_i \geq \epsilon_{i+1}\}.$$

It is straightforward to prove the following lemma, which characterizes the multiset permutations of $\{w_1, \ldots, w_n\}$ in terms of their inversion sequences.

Lemma 2.1 *Given $w = (w_1 \leq w_2 \leq \ldots \leq w_n)$ or $w = (w_1 \geq w_2 \geq \ldots \geq w_n)$, the mapping $\pi \to \epsilon(\pi)$ on S_n restricts to a bijection between S_n/w and I_n/w. The mapping $\pi \to \pi(w)$ is a bijection between S_n/w and (distinguishable) permutations of w. In particular, there is a bijection between permutations of w and inversion sequences in I_n/w.*

2.2 The bijection for lecture hall partitions

Bijection $\Theta : \mathbb{Z}^n_{\geq 0} \to L_n$:

For $p \in \mathbb{Z}^n_{\geq 0}$, define $\Theta(p)$ as follows:

1. Let π^{-1} be the unique permutation that stably sorts p into weakly increasing order $(w_1 \leq w_2 \leq ... \leq w_n) = \pi^{-1}(p)$
2. Let $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ be the inversion sequence of π

Then $\Theta(p) = \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ where $\lambda_i = iw_i - \epsilon_i, \quad i = 1, 2, \ldots, n.$

Example 2.1 Let $p = (9, 0, 3, 3, 5, 4, 3, 8, 1, 8, 2, 9) \in \mathbb{Z}^{12}_{\geq 0}$. Then $w = (0, 1, 2, 3, 3, 3, 4, 5, 8, 8, 9, 9)$ and $\pi = (11, 1, 4, 5, 8, 7, 6, 9, 2, 10, 3, 12)$ and $\epsilon(\pi) = (0, 0, 0, 2, 2, 2, 3, 4, 2, 1, 10, 0)$. So

$$\begin{aligned}
\Theta(p) &= \lambda \\
&= (0-0, 2-0, 6-0, 12-2, 15-2, 18-2, 28-3, 40-4, 72-2, 80-1, 99-10, 108-0) \\
&= (0, 2, 6, 10, 13, 16, 25, 36, 70, 79, 89, 108).
\end{aligned}$$

To check that $\Theta(p) = \lambda \in L_n$, verify that

$$0 \leq \frac{0}{1} \leq \frac{2}{2} \leq \frac{6}{3} \leq \frac{10}{4} \leq \frac{13}{5} \leq \frac{16}{6} \leq \frac{25}{7} \leq \frac{36}{8} \leq \frac{70}{9} \leq \frac{79}{10} \leq \frac{89}{11} \leq \frac{108}{12}.$$

Also, note that $p \in Q^{12}_9$, since its largest coordinate is 9 and that $\Theta(p) = \lambda \in L^{(9)}_n$, since $\lambda_{12}/12 = (108)/(12) \leq 9$.

Theorem 2.2 *Θ is a bijection between lattice points in $\mathbb{Z}^n_{\geq 0}$ and lecture hall partitions of length n. In fact, $\Theta(Q^n_t) = L^{(t)}_n$.*

Proof: First, to prove $\Theta(Q^n_t) \subseteq L^{(t)}_n$, let $p \in Q^n_t$ and $\lambda = \Theta(p)$. From the definition of Θ, $\lambda_i = iw_i - \epsilon_i$, where $w = (w_1 \leq w_2 \leq \ldots \leq w_n)$ is the sorted sequence of coordinates of p and $(\epsilon_1, \epsilon_2, \ldots, \epsilon_n) = \epsilon(\pi)$ for the unique $\pi \in S_n/w$ with $\pi(w) = p$. By Lemma 2.1, $\epsilon(\pi) \in I_n/w$, so $0 \leq \epsilon_i < i$ and if $w_i = w_{i+1}$, then $\epsilon_i \geq \epsilon_{i+1}$.

To show that $\lambda \in L_n^{(t)}$, we must show that

$$0 \leq \frac{1w_1 - \epsilon_1}{1} \leq \ldots \leq \frac{iw_i - \epsilon_i}{i} \leq \frac{(i+1)w_{i+1} - \epsilon_{i+1}}{i+1} \leq \ldots \leq \frac{nw_n - \epsilon_n}{n} \leq t.$$

Clearly, the first inequality holds, since $w_1 \geq 0$ and $\epsilon_1 = 0$. Also, since $p \in Q_t^n$, $w_n \leq t$, so the last inequality holds.

To show $\frac{iw_i - \epsilon_i}{i} \leq \frac{(i+1)w_{i+1} - \epsilon_{i+1}}{i+1}$, consider the relationship between w_i and w_{i+1}. If $w_i = w_{i+1}$, then since $\epsilon_i \geq \epsilon_{i+1}$,

$$\frac{iw_i - \epsilon_i}{i} = w_{i+1} - \frac{\epsilon_i}{i} \leq w_{i+1} - \frac{\epsilon_{i+1}}{i} \leq w_{i+1} - \frac{\epsilon_{i+1}}{i+1} = \frac{(i+1)w_{i+1} - \epsilon_{i+1}}{i+1}.$$

Otherwise, $w_{i+1} \geq w_i + 1$, so since $0 \leq \epsilon_{i+1} < i + 1$,

$$\frac{(i+1)w_{i+1} - \epsilon_{i+1}}{i+1} \geq w_i + 1 - \frac{\epsilon_{i+1}}{i+1} \geq w_i + 1 - \frac{i}{i+1} > w_i \geq \frac{iw_i - \epsilon_i}{i}.$$

To complete the proof that $\Theta(Q_t^n) = L_n^{(t)}$, since by Theorem 1.1, $|L_n^{(t)}| = |Q_t^n|$, it suffices to show that Θ is one-to-one. Suppose $\Theta(p) = \lambda = \Theta(r)$ for $p, r \in Q_t^n$. Then $\lambda = (w_1 - \epsilon_1, \ldots, iw_i - \epsilon_i, \ldots, nw_n - \epsilon_n)$ for some $w = (w_1 \leq w_2 \leq \ldots \leq w_n)$ and ϵ satisfying $\epsilon \in I_n/w$, and in particular, $0 \leq \epsilon_i < i$. But this determines w uniquely as

$$w = (\lceil \lambda_1/1 \rceil, \lceil \lambda_2/2 \rceil, \ldots, \lceil \lambda_n/n \rceil) \tag{2}$$

and thus ϵ uniquely as

$$\epsilon_i = iw_i - \lambda_i.$$

There is a unique permutation $\pi \in S_n$ with inversion sequence ϵ and by Lemma 2.1, $\pi \in S_n/w$. Then, by the definition of Θ, $\pi(w) = p$ and $\pi(w) = r$. Thus $p = r$ and therefore Θ is a bijection. $\square$

2.3 The bijection for anti-lecture hall compositions

Bijection $\Phi : \mathbb{Z}_{\geq 0}^n \to L_n$:

For $p \in \mathbb{Z}_{\geq 0}^n$, define $\Phi(p)$ as follows:

1. Let π^{-1} be the unique permutation that stably sorts p into weakly decreasing order $(w_1 \geq w_2 \geq ... \geq w_n) = \pi^{-1}(p)$
2. Let $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ be the inversion sequence of π

Then $\Phi(p) = \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ where $\lambda_i = iw_i + \epsilon_i, \quad i = 1, 2, \ldots, n.$

Example 2.2 Let $p = (9, 0, 3, 3, 5, 4, 3, 8, 1, 8, 2, 9) \in \mathbb{Z}_{\geq 0}^{12}$. Then $w = (9, 9, 8, 8, 5, 4, 3, 3, 3, 2, 1, 0)$ and $\pi = (1, 12, 7, 8, 5, 6, 9, 3, 11, 4, 10, 2)$ and $\epsilon(\pi) = (0, 0, 1, 1, 3, 3, 5, 5, 3, 1, 3, 10)$. So

$$\Phi(p) = \lambda = (9, 18, 25, 33, 28, 27, 26, 29, 30, 21, 14, 10).$$

Theorem 2.3 *Φ is a bijection between lattice points in $\mathbb{Z}^n_{\geq 0}$ and anti-lecture hall compositions of length n. In fact, $\Phi(Q^n_t) = A^{(t)}_n$.*

Proof: In the same spirit as the proof of Theorem 2.2, first show that λ resulting from $\Phi(p)$ is, in fact, an anti-lecture hall composition. Then, cite Theorem 1.1 to show $\left|A^{(t)}_n\right| = |Q^n_t|$. To complete the bijective proof, show that Φ is one-to-one by observing that if $\Phi(p) = \lambda = \Phi(r)$, then because

$$w = (\lfloor \lambda_1/1 \rfloor, \lfloor \lambda_2/2 \rfloor, \ldots, \lfloor \lambda_n/n \rfloor), \tag{3}$$

we know, from Lemma 2.1 that $p = r$. □

3 Generating Functions

In this section we will derive generating functions for lecture hall partitions and anti-lecture hall compositions via the bijections Θ and Φ. We need the following additional observations about permutations.

Lemma 3.1 *If $\pi \in S_n$ stably sorts $(p_1, \ldots, p_n)$ into weakly increasing order and $\sigma \in S_n$ stably sorts $(p_n, \ldots, p_1)$ into weakly decreasing order then $\sigma_i = n + 1 - \pi_{n+1-i}$.*

Lemma 3.2 *Let $\sigma, \pi \in S_n$ be related by $\sigma_i = n + 1 - \pi_{n+1-i}$. Then their inverses are similarly related: $\sigma^{-1}_i = n + 1 - \pi^{-1}_{n+1-i}$.*

Lemma 3.3 *If $\sigma, \pi \in S_n$ are related by $\sigma_i = n+1-\pi_{n+1-i}$, then $\mathrm{des}(\sigma) = \mathrm{des}(\pi)$ and $\mathrm{inv}(\sigma) = \mathrm{inv}(\pi)$.*

For a point $p \in \mathbb{Z}^n_{\geq 0}$, the *weight* of p is $|p| = p_1 + \ldots + p_n$. For $\lambda \in A_n$, let

$$\lfloor \lambda \rfloor = (\lfloor \lambda_1/1 \rfloor, \lfloor \lambda_2/2 \rfloor, \ldots, \lfloor \lambda_n/n \rfloor).$$

Note from (3) that, for $\lambda \in A_n$,

$$|\lfloor \lambda \rfloor| = |\Phi^{-1}(\lambda)|.$$

Similarly, for $\lambda \in L_n$, let

$$\lceil \lambda \rceil = (\lceil \lambda_1/1 \rceil, \lceil \lambda_2/2 \rceil, \ldots, \lceil \lambda_n/n \rceil).$$

Then from (2) for $\lambda \in L_n$,

$$|\lceil \lambda \rceil| = |\Theta^{-1}(\lambda)|.$$

Define

$$L_n(u, q) = \sum_{\lambda \in L_n} u^{|\lceil \lambda \rceil|} q^{|\lambda|} \quad \text{and} \quad A_n(u, q) = \sum_{\lambda \in A_n} u^{|\lfloor \lambda \rfloor|} q^{|\lambda|}.$$

It was shown in [BME99] that

$$L_n(u, q) = \prod_{i=1}^{n} \frac{1 + uq^i}{1 - u^2 q^{n+i}} \tag{4}$$

and in [CS04] that

$$A_n(u,q) = \prod_{i=1}^{n} \frac{1+uq^i}{1-u^2q^{1+i}}. \tag{5}$$

Although a relationship between $A_n(u,q)$ and $L_n(u,q)$ can be deduced from (4) and (5), each generating function was derived independently and until now the relationship could not be explained combinatorially.

From Theorem 2.2, the mapping $p \to \Theta(p)$ is a bijection $\mathbb{Z}^n_{\geq 0} \to L_n$ and if $\lambda = \Theta(p)$ then $|p| = |\lceil \lambda \rceil|$. Thus

$$L_n(u,q) = \sum_{p \in \mathbb{Z}^n_{\geq 0}} u^{|p|} q^{|\Theta(p)|}.$$

Similarly, from Theorem 2.3, the mapping $p \to \Phi(p)$ is a bijection $\mathbb{Z}^n_{\geq 0} \to A_n$ and if $\lambda = \Phi(p)$ then $|p| = |\lfloor \lambda \rfloor|$. So,

$$A_n(u,q) = \sum_{p \in \mathbb{Z}^n_{\geq 0}} u^{|p|} q^{|\Phi(p)|}.$$

For the first time we are able to show that $L_n(u,q)$ can be derived from $A_n(u,q)$. Define the reverse of a sequence $s = (s_1, s_2, \ldots, s_n)$ by $s^{\mathrm{rev}} = (s_n, s_{n-1}, \ldots, s_1)$.

Theorem 3.4

$$L_n(u,q) = A_n(uq^{n+1}, q^{-1}).$$

Proof: For $p \in \mathbb{Z}^n_{\geq 0}$, we compare the contribution of $\Phi(p)$ to $A_n(u,q)$ with the contribution of $\Theta(p^{\mathrm{rev}})$ to $L_n(u,q)$. Let π^{-1} be the permutation that stably sorts p into weakly decreasing order $w = (w_1 \geq \ldots \geq w_n)$ and let σ^{-1} be the permutation that stably sorts p^{rev} into weakly increasing order w^{rev}. Then

$$\begin{aligned} \Phi(p) &= (w_1 + \epsilon_1(\pi), 2w_2 + \epsilon_2(\pi), \ldots, nw_n + \epsilon_n(\pi)) \\ \Theta(p^{\mathrm{rev}}) &= (w_n - \epsilon_1(\sigma), 2w_{n-1} - \epsilon_2(\sigma), \ldots, nw_1 - \epsilon_n(\sigma)). \end{aligned}$$

So

$$u^{|p|} q^{|\Phi(p)|} = u^{|w|} q^{\sum_i^n i w_i} q^{\mathrm{inv}(\pi)}$$

and

$$\begin{aligned} u^{|p^{\mathrm{rev}}|} q^{|\Theta(p^{\mathrm{rev}})|} &= u^{|w^{\mathrm{rev}}|} q^{\sum_i^n (n+1-i) w_i} q^{-\mathrm{inv}(\sigma)} \\ &= (uq^{n+1})^{|w|} q^{-\sum_i^n i w_i} q^{-\mathrm{inv}(\sigma)} = (uq^{n+1})^{|p|} q^{-|\Phi(p)|}, \end{aligned}$$

where we have used $\mathrm{inv}(\sigma) = \mathrm{inv}(\pi)$ from Lemma 3.3. Note finally that summing over all $p \in \mathbb{Z}^n_{\geq 0}$ is equivalent to summing over all $p^{\mathrm{rev}} \in \mathbb{Z}^n_{\geq 0}$, so

$$\begin{aligned} L_n(u,q) = \sum_{p \in \mathbb{Z}^n_{\geq 0}} u^{|p|} q^{|\Theta(p)|} &= \sum_{p^{\mathrm{rev}} \in \mathbb{Z}^n_{\geq 0}} u^{|p^{rev}|} q^{|\Theta(p^{rev})|} \\ &= \sum_{p \in \mathbb{Z}^n_{\geq 0}} (uq^{n+1})^{|p|} q^{-|\Phi(p)|} = A_n(uq^{n+1}, q^{-1}). \end{aligned}$$

□

Now we derive the generating function for A_n in terms of permutation statistics.

Theorem 3.5

$$A_n(u,q) = \sum_{\pi \in S_n} \frac{q^{\mathrm{inv}(\pi)} \prod_{i \in \mathrm{Des}(\pi)} u^i q^{i(i+1)/2}}{(1-uq)(1-u^2q^{1+2})\cdots(1-u^nq^{1+2+\ldots+n})}$$

Proof: For $T \subseteq \mathbb{Z}^n_{\geq 0}$, define F_T by

$$F_T(u,z_1,\ldots,z_n) = \sum_{\lambda \in T} u^{|\lambda|} z_1^{\lambda_1} z_2^{\lambda_2} \ldots z_n^{\lambda_n}.$$

Given $D \subseteq \{1,2,\ldots,n-1\}$, define

$$S_D = \{(w_1 \geq w_2 \geq \ldots \geq w_n) \in \mathbb{Z}^n_{\geq 0} \mid w_i > w_{i+1} \text{ if } i \in D\}.$$

Then

$$F_{S_D}(u,z_1,\ldots,z_n) = \frac{\prod_{i \in D} u^i z_1 z_2 \cdots z_i}{(1-uz_1)(1-u^2z_1z_2)\cdots(1-u^nz_1z_2\cdots z_n)}.$$

We count A_n from $\mathbb{Z}^n_{\geq 0}$ via Φ. Use the permutations $\pi \in S_n$ to partition the points $p \in \mathbb{Z}^n_{\geq 0}$ into sets T_π defined by

$$T_\pi = \{p \mid p = \pi(w_1 \geq w_2 \geq \ldots \geq w_n) \text{ such that } i \in \mathrm{Des}(\pi^{-1}) \ \rightarrow\ w_i > w_{i+1}\}.$$

So, we are partitioning the points according to the permutation π such that π^{-1} stably sorts p into weakly decreasing order. The bijection $\Phi : \mathbb{Z}^n_{\geq 0} \to A_n$ does the following to the points in T_π: They are first mapped onto the points in $S_{\mathrm{Des}(\pi^{-1})}$. Then for each i, the ith coordinate is multiplied by i and added to $\epsilon_i(\pi)$. So in the generating function

$$z_1^{\epsilon_1(\pi)} \cdots z_n^{\epsilon_n(\pi)} F_{S_{\mathrm{Des}(\pi^{-1})}}(u,z_1,z_2^2,\ldots,z_n^n),$$

u keeps track of the weight of $p \in T_\pi$ and the variables z_i track the weight of $\Phi(p)$. Putting this together,

$$\begin{aligned}
A_n(u,z_1,\ldots,z_n) &= \sum_{\lambda \in A_n} u^{|\lfloor\lambda\rfloor|} z_1^{\lambda_1} \cdots z_n^{\lambda_n} = \sum_{\pi \in S_n} \sum_{p \in T_\pi} u^{|p|} z_1^{\Theta(p)_1} \cdots z_n^{\Theta(p)_n} \\
&= \sum_{\pi \in S_n} \sum_{p \in T_\pi} z_1^{\epsilon_1(\pi)} \cdots z_n^{\epsilon_n(\pi)} F_{S_{\mathrm{Des}(\pi^{-1})}}(u,z_1,z_2^2,\ldots,z_n^n) \\
&= \sum_{\pi \in S_n} \frac{z_1^{\epsilon_1(\pi)} \cdots z_n^{\epsilon_n(\pi)} \prod_{i \in \mathrm{Des}(\pi^{-1})} u^i z_1 z_2^2 \cdots z_i^i}{(1-uz_1)(1-u^2z_1z_2^2)\cdots(1-u^nz_1z_2^2\cdots z_n^n)}.
\end{aligned}$$

Setting all $z_i = q$, and using (1), which states that $\mathrm{inv}(\pi) = \mathrm{inv}(\pi^{-1})$,

$$A_n(u,q) = \sum_{\pi \in S_n} \frac{q^{\mathrm{inv}(\pi)} \prod_{i \in \mathrm{Des}(\pi)} u^i q^{i(i+1)/2}}{(1-uq)(1-u^2q^{1+2})\cdots(1-u^nq^{1+2+\ldots+n})}.$$

□

Theorem 3.6

$$L_n(u,q) = \sum_{\pi \in S_n} \frac{q^{-\mathrm{inv}(\pi)} \prod_{i \in \mathrm{Des}(\pi)} (uq^{(n+1)})^i q^{-i(i+1)/2}}{\prod_{i=1}^{n} (1 - u^i q^{i(n+1) - i(i+1)/2})}.$$

Proof: From Theorem 3.4, $L_n(u,q) = A_n(uq^{n+1}, q^{-1})$, so apply Theorem 3.5. □

Combining Theorems 3.5 and 3.6 with equations (4) and (5) will have implications about the distribution of certain permutation statistics, discussed in the next section.

4 Quadratic Permutation Statistics

Define the q-integer $[\ n\ ]_q$ by $[\ n\ ]_1 = 1$ and for $q \neq 1$, by $[\ n\ ]_q = (1-q^n)/(1-q)$. In Section 2.1 we defined the permutation statistics inv and des. The *major index* of $\pi \in S_n$ is the sum of the descent positions: $\mathrm{maj}(\pi) = \sum_{i \in \mathrm{Des}(\pi)} i$. It is known that

$$\sum_{\pi \in S_n} q^{\mathrm{maj}(\pi)} = \prod_{i=1}^{n} [\ i\ ]_q, \tag{6}$$

and that inv and maj are equally distributed over all permutations [Mac60].

Motivated by Theorems 3.5 and 3.6, we introduce quadratic permutation statistics sq and bin:

$$\mathrm{sq}(\pi) \ = \sum_{i \in \mathrm{Des}(\pi)} i^2 \qquad \text{and} \qquad \mathrm{bin}(\pi) \ = \sum_{i \in \mathrm{Des}(\pi)} \binom{i+1}{2}.$$

Because of the way "inv" is involved with the distribution of these quadratic statistics, we also define

$$\begin{aligned} \mathrm{sqin}(\pi) &= \mathrm{sq}(\pi) + \mathrm{inv}(\pi) \\ \mathrm{binv}(\pi) &= \mathrm{bin}(\pi) + \mathrm{inv}(\pi) \end{aligned}$$

and prove two distribution theorems that refine (6). The first comes from the enumeration of anti-lecture hall compositions.

Theorem 4.1

$$\sum_{\pi \in S_n} u^{\mathrm{maj}(\pi)} q^{\mathrm{binv}(\pi)} = \prod_{i=1}^{n} (1 - u^i q^{\binom{i+1}{2}}) \frac{1 + uq^i}{1 - u^2 q^{1+i}}.$$

Proof: Restate the generating function for $A_n(u,q)$ in Theorem 3.5 in terms of the new permutation statistics and apply equation (5). □

Setting $q = 1$ in Theorem 4.1 gives (6). Setting $u = 1$ in Theorem 4.1 gives the following interesting generating function for the symmetric group.

Corollary 4.2

$$\sum_{\pi \in S_n} q^{\mathrm{binv}(\pi)} = \prod_{i=1}^{n} [\,2\,]_{q^i} \frac{[\,i(i+1)/2\,]_q}{[\,i+1\,]_q}$$

In Theorem 4.1, setting $q = q^{-1}$ and then $u = q^{n+1}$ gives an unusual variation of (6).

Corollary 4.3

$$\sum_{\pi \in S_n} q^{(n+1)\mathrm{maj}(\pi)-\mathrm{binv}(\pi)} = \prod_{i=1}^{n} [\,i\,]_{q^{2(n-i)+1}}.$$

Proof: By Theorem 3.4, $A_n(u, q^{n+1}, q^{-1}) = L_n(u, q)$. Equate $L_n(u, q)$ in Theorem 3.6 and equation (4), setting $u = 1$, and simplify:

$$\sum_{\pi \in S_n} q^{(n+1)\mathrm{maj}(\pi)-\mathrm{binv}(\pi)} = \prod_{i=1}^{n}(1 - q^{i(2n-i+1)/2}) \prod_{i=1}^{n} \frac{1}{1 - q^{2i-1}}.$$

The result follows by observing that

$$1 - q^{i(2n-i+1)/2} = \begin{cases} 1 - q^{k(2(n-k)+1)} = [\,k\,]_{q^{2(n-k)+1}}(1 - q^{2(n-k)+1}) & \text{if } i = 2k \\ 1 - q^{(2k+1)(n-k)} = [\,n-k\,]_{q^{2k+1}}(1 - q^{2k+1}) & \text{if } i = 2k+1. \end{cases}$$

□

The second distribution theorem has the following form.

Theorem 4.4

$$\sum_{\pi \in S_n} q^{\mathrm{maj}(\pi)} t^{\mathrm{sqin}(\pi)} = \prod_{i=1}^{n} [\,i\,]_{qt^i}$$

Before proving Theorem 4.4, we observe that it has the following specializations. Setting $t = 1$ in Theorem 4.4 gives (6). Setting $q = 1$ in Theorem 4.4 gives the following, which appears to be a new observation:

Corollary 4.5

$$\sum_{\pi \in S_n} t^{\mathrm{sqin}(\pi)} = \prod_{i=1}^{n} [\,i\,]_{t^i}.$$

Setting $q = q^n$ and $t = 1/q$ in Theorem 4.4 gives:

Corollary 4.6

$$\sum_{\pi \in S_n} q^{nmaj(\pi)-\mathrm{sqin}(\pi)} = \prod_{i=1}^{n} [\,i\,]_{q^{n-i}}.$$

In [SW98], Stembridge and Waugh derive a Weyl group generating function which, especially in the case of the symmetric group, they felt "ought to be better known". In [Zab03], Zabrocki gave a simple combinatorial proof of that special case, which was exactly Corollary 4.6 above. It appears that we have come full circle, since lecture hall partitions originally arose in Eriksson's work on affine Coxeter groups [BME97].

Setting $q = q^{2n+1}$ and $t = 1/q^2$ in Theorem 4.4 gives:

Corollary 4.7

$$\sum_{\pi \in S_n} q^{(2n+1)maj(\pi) - 2\mathrm{sqin}(\pi)} = \prod_{i=1}^{n} [\, i \,]_{q^{2(n-i)+1}},$$

We finish this section with a proof of Theorem 4.4, which follows the same strategy as Zabrocki's proof of Corollary 4.6 above. In contrast to the proof of Theorem 4.3, this is a direct and elementary proof which does not rely on the theory of lecture hall partitions or affine Coxeter groups. We have not as yet found a similar approach to Theorem 4.3.

Proof: (of Theorem 4.4) Expand the product in Theorem 4.4 as

$$\begin{aligned} \prod_{i=1}^{n} [\, i \,]_{qt^i} &= (1)(1+qt^2)(1+qt^3+(qt^3)^2)\ldots && (7) \\ &= \sum_{(r_1,\ldots,r_n)} q^{r_1+\ldots+r_n} t^{1r_1+2r_2+\ldots+nr_n} && (8) \end{aligned}$$

where the sum is over the $n!$ sequences $(r_1, \ldots, r_n)$ satisfying $0 \le r < i$. So, we will establish a bijection from S_n to these sequences with the property that if π maps to $(r_1, \ldots, r_n)$, then $\mathrm{maj}(\pi) = r_1 + \ldots + r_n$ and $\mathrm{sqin}(\pi) = 1r_1 + 2r_2 + \ldots + nr_n$.

Given π, let $\epsilon = \epsilon(\pi^{-1})$ be the inversion sequence of π^{-1}. Define r by $r_i = \epsilon_i - \epsilon_{i+1} + i$ if $i \in \mathrm{Des}(\pi)$ and $r_i = \epsilon_i - \epsilon_{i+1}$, otherwise. Observe that $\epsilon_i < \epsilon_{i+1}$ if and only if $i \in \mathrm{Des}(\pi)$. By definition, $0 \le \epsilon_i < i$ for every i Thus $0 \le r_i < i$ for every i. Clearly $r_1 + \ldots + r_n = \mathrm{maj}(\pi)$ and

$$\sum_{i=1}^{n} ir_i = \sum_{i \in \mathrm{Des}(\pi)} i^2 + \sum_{i=1}^{n} i(\epsilon_i - \epsilon_{i+1}) = \mathrm{sq}(\pi) + \mathrm{inv}(\pi^{-1}) = \mathrm{sq}(\pi) + \mathrm{inv}(\pi).$$

Finally, observe that ϵ, and therefore π^{-1} and π, can be recovered from r: $\epsilon_n = r_n$ and for $i < n$, given r_i and ϵ_{i+1}, it must be that $\epsilon_i = r_i + \epsilon_{i+1}$ if $r_i + \epsilon_{i+1} < i$ and otherwise, $\epsilon_i = r_i + \epsilon_{i+1} - i$. □

5 Further directions

We mention a few questions suggested by this work. Are there other areas where quadratic permutation statistics arise naturally? Other joint distribution results? Can we give a direct and elementary proof of Theorem 4.4 on the joint distribution of maj and binv that is independent of the theory of lecture hall partitions and Weyl groups? The lecture hall theorem came from the theory of affine Coxeter groups and Bott's formula; do anti-lecture hall compositions have any place in this theory?

Acknowledgements

We are grateful to Matthias Beck for inspiring us to think geometrically. Thanks also to Ira Gessel for pointing us to the Zabrocki paper, when we were searching for examples of quadratic permutation statistics.

References

[ACS09] George E. Andrews, Sylvie Corteel, and Carla D. Savage. On q-series identities arising from lecture hall partitions. *Int. J. Number Theory*, 5(2):327–337, 2009.

[Bec09] Matthias Beck. Personal communication, March, 2009.

[BME97] Mireille Bousquet-Mélou and Kimmo Eriksson. Lecture hall partitions. *Ramanujan J.*, 1(1):101–111, 1997.

[BME99] Mireille Bousquet-Mélou and Kimmo Eriksson. A refinement of the lecture hall theorem. *J. Combin. Theory Ser. A*, 86(1):63–84, 1999.

[CLS07] Sylvie Corteel, Sunyoung Lee, and Carla D. Savage. Enumeration of sequences constrained by the ratio of consecutive parts. *Sém. Lothar. Combin.*, 54A:Art. B54Aa, 12 pp. (electronic), 2005/07.

[CS03] Sylvie Corteel and Carla D. Savage. Anti-lecture hall compositions. *Discrete Math.*, 263(1-3):275–280, 2003.

[CS04] Sylvie Corteel and Carla D. Savage. Lecture hall theorems, q-series and truncated objects. *J. Combin. Theory Ser. A*, 108(2):217–245, 2004.

[CSS09] Sylvie Corteel, Carla D. Savage, and Andrew V. Sills. Lecture hall sequences, q-series, and asymmetric partition identities. *Developments in Mathematics*, 2009. to appear.

[Knu73] Donald E. Knuth. *The art of computer programming. Volume 3*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. Sorting and searching, Addison-Wesley Series in Computer Science and Information Processing.

[Mac60] Percy A. MacMahon. *Combinatory analysis*. Two volumes (bound as one). Chelsea Publishing Co., New York, 1960.

[SW98] John R. Stembridge and Debra J. Waugh. A Weyl group generating function that ought to be better known. *Indag. Math. (N.S.)*, 9(3):451–457, 1998.

[SY08] Carla D. Savage and Ae Ja Yee. Euler's partition theorem and the combinatorics of ℓ-sequences. *J. Combin. Theory Ser. A*, 115(6):967–996, 2008.

[Yee01] Ae Ja Yee. On the combinatorics of lecture hall partitions. *Ramanujan J.*, 5(3):247–262, 2001.

[Zab03] Mike Zabrocki. A bijective proof of an unusual symmetric group generating function. 2003. preprint, `arXiv:math/0910301v1`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 455–466

A preorder-free construction of the Kazhdan-Lusztig representations of Hecke algebras $H_n(q)$ of symmetric groups

Charles Buehrle[1] and Mark Skandera[2]

Authors' address: Dept. of Mathematics, Lehigh University, Bethlehem, PA 18015, USA
[1] `ceb405@lehigh.edu` [2] `mas906@lehigh.edu`

Abstract. We use a quantum analog of the polynomial ring $\mathbb{Z}[x_{1,1}, \ldots, x_{n,n}]$ to modify the Kazhdan-Lusztig construction of irreducible $H_n(q)$-modules. This modified construction produces exactly the same matrices as the original construction in [*Invent. Math.* **53** (1979)], but does not employ the Kazhdan-Lusztig preorders. Our main result is dependent on new vanishing results for immanants in the quantum polynomial ring.

Résumé. Nous utilisons un analogue quantique de l'anneau $\mathbb{Z}[x_{1,1}, \ldots, x_{n,n}]$ pour modifier la construction Kazhdan-Lusztig des modules-$H_n(q)$ irreductibles. Cette construction modifiée produit exactement les mêmes matrices que la construction originale dans [*Invent. Math.* **53** (1979)], mais sans employer les préordres de Kazhdan-Lusztig. Notre résultat principal dépend de nouveaux résultats de disparaition pour des immanants dans l'anneau polynôme de quantique.

Resumen. Utilizamos un analog cuántico del anillo $\mathbb{Z}[x_{1,1}, \ldots, x_{n,n}]$ para modificar la construcción de Kazhdan-Lusztig de módulos-$H_n(q)$ irreducibles. Esta construcción modificada produce exactamente las mismas matrices que la construcción original en [*Invent. Math.* **53** (1979)], pero sin emplear los preórdenes de Kazhdan-Lusztig. Nuestro resultado principal es depende en los nuevos resultados de desaparición para los imanantes en el anillo polinómico del cuántico.

Keywords: Kazhdan-Lusztig, immanants, irreducible representations, Hecke algebra

1 Introduction

In 1979, Kazhdan and Lusztig introduced [8] a family of modules for Coxeter groups and related Hecke algebras. These modules, which happen to be irreducible for Coxeter groups of type-A and have many fascinating properties, also aid in the understanding of modules for quantum groups and other algebras. Important ingredients in the construction of the Kazhdan-Lusztig modules are the computation of certain polynomials in $\mathbb{Z}[q]$ known as *Kazhdan-Lusztig polynomials*, and the description of preorders on Coxeter group elements known as the *Kazhdan-Lusztig preorders*. These two tasks, which present something of an obstacle to one wishing to construct the modules, have become fascinating research topics in their own right. Even in the simplest case of a Coxeter group, the symmetric group S_n, the Kazhdan-Lusztig polynomials and preorders are somewhat poorly understood, see [2, Chp. 6], [13].

As an alternative to the "traditional" Kazhdan-Lusztig construction of type-A modules in terms of subspaces of the type-A Hecke algebra $H_n(q)$, one may construct modules in terms of subspaces of a non-commutative "quantum polynomial ring". Theoretically, this alternative offers no special advantage over the original construction. On the other hand, a simple modification of this alternative completely eliminates the need for the Kazhdan-Lusztig preorders in a new construction of $H_n(q)$-modules.

In Sections 2-3, we review essential definitions for the symmetric group, Hecke algebra, and Kazhdan-Lusztig modules. In Section 4 we review definitions related to a quantum analog of the polynomial ring $\mathbb{Z}[x_{1,1}, \ldots, x_{n,n}]$ and a particular $n!$-dimensional subspace called the *quantum immanant space*. In Section 4, we use the basis of Kazhdan-Lusztig immanants studied in [10] to transfer the traditional Kazhdan-Lusztig representations of $H_n(q)$ to the immanant space.

Results of Clausen [4] will then motivate us to modify the above representations in Section 5 and to apply vanishing properties of Kazhdan-Lusztig immanants similar to those obtained in [11]. This leads to our main result that the resulting representations, which do not rely upon the Kazhdan-Lusztig preorders, have matrices equal to those corresonding to the original Kazhdan-Lusztig representations in [8].

2 Tableaux and the symmetric group

We call a weakly decreasing sequence $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ of positive integers with $\sum_{i=1}^{\ell} \lambda_i = r$ an *integer partition* of r, and we denote this by $\lambda \vdash r$ or $|\lambda| = r$. A partial ordering on integer partitions of r called *dominance order* is given by $\lambda \succeq \mu$ if and only if

$$\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i, \text{ for all } i = 1, \ldots, \ell. \tag{1}$$

From an integer partition λ we can construct a *Ferrers diagram* which has λ_i left justified dots in row i. When we replace the dots in a diagram with $1, \ldots, r$ we have a *Young tableau* where the *shape* of the tableau is λ. An *injective tableau* is merely one in which the replacing is performed injectively, i.e. the $1, \ldots, r$ appear exactly once in the tableau. We call a tableau *column-(semi)strict* if its entries are (weakly) increasing downward in columns. A tableau is *row-(semi)strict* if entries (weakly) increase from left to right in rows. We call a tableau *semistandard* if it is column-strict and row-semistrict, and *standard* if it is semistandard and injective. We define transposition of partitions $\lambda \mapsto \lambda^\top$ (also known as conjugation) and tableaux $T \mapsto T^\top$ in a manner analogous to matrix transposition. We define a *bitableau* to be a pair of tableaux of the same shape, and say that it posesses a certain tableau property if both of its tableaux posess this property.

For each partition λ we define the *superstandard tableau* of shape λ to be the tableau $T(\lambda)$ having entries in reading order. For example,

$$T((4,2,1)) = \begin{matrix} 1 & 2 & 3 & 4 \\ 5 & 6 & & \\ 7 & & & \end{matrix} \;. \tag{2}$$

The standard presentation of S_n is given by generators $s_1, \ldots, s_{n-1}$ and relations

$$\begin{aligned} s_i^2 &= 1, && \text{for } i = 1, \ldots, n-1, \\ s_i s_j s_i &= s_j s_i s_j, && \text{if } |i-j| = 1, \\ s_i s_j &= s_j s_i, && \text{if } |i-j| \geq 2. \end{aligned} \tag{3}$$

Let S_n act on rearrangements of the letters $[n] = \{1, \ldots, n\}$ by

$$s_i \circ v_1 \cdots v_n \underset{\text{def}}{=} v_1 \cdots v_{i-1} v_{i+1} v_i v_{i+2} \cdots v_n. \tag{4}$$

For each permutation $w = s_{i_1} \cdots s_{i_\ell} \in S_n$ we define the *one-line notation* of w to be the word

$$w_1 \cdots w_n \underset{\text{def}}{=} s_{i_1} \circ (\cdots (s_{i_\ell} \circ (1 \cdots n)) \cdots). \tag{5}$$

For each $w \in S_n$ we define two tableaux, $P(w), Q(w)$ which are obtained from the *Robinson-Schensted correspondence* using column insertion to the one-line notation of w. (See, e.g., [12, Sec. 3.1].) It is well known that these tableaux satisfy $P(w^{-1}) = Q(w)$. Since $\mathrm{sh}(P(w)) = \mathrm{sh}(Q(w))$ we can define the *shape of a permutation* as $\mathrm{sh}(w) = \mathrm{sh}(P(w))$.

Given a permutation $w \in S_n$ expressed in terms of generators $w = s_{i_1} \cdots s_{i_\ell}$ we say this expression is *reduced* if w cannot be expressed as a shorter product of generators of S_n. We call the *length* of a permutation $w \in S_n$ $\ell(w) = \ell$, in the previous equation. We define the *Bruhat order* on S_n by $v \leq w$ if some (equivalently every) reduced expression for w contains a reduced expression for v as a subword (The reader is referred to [2] for more on this topic). Throughout this paper we will use w_0 to denote the unique maximal element in the Bruhat order. Multiplying a permutation on the right by w_0 also changes the bitableau of the Robinson-Schensted correspondence for that permutation. Specifically, this change can be described in terms of transposition. (See [2, Appendix].)

Lemma 2.1 *If $v \in S_n$, then $Q(v) = (Q(vw_0))^\top$.*

3 Kazhdan-Lusztig representations

Given an indeterminate q we define the *Hecke algebra*, $H_n(q)$, to be the $\mathbb{Z}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-algebra with multiplicative identity $\widetilde{T}_e$ generated by $\{\widetilde{T}_{s_i}\}_{i=1}^{n-1}$ with relations

$$\widetilde{T}_{s_i}^2 = (q^{\frac{1}{2}} - q^{\frac{-1}{2}})\widetilde{T}_{s_i} + \widetilde{T}_e, \qquad \text{for } i = 1, \ldots, n-1, \tag{6}$$

$$\widetilde{T}_{s_i}\widetilde{T}_{s_j}\widetilde{T}_{s_i} = \widetilde{T}_{s_j}\widetilde{T}_{s_i}\widetilde{T}_{s_j}, \qquad \text{if } |i-j| = 1, \tag{7}$$

$$\widetilde{T}_{s_i}\widetilde{T}_{s_j} = \widetilde{T}_{s_j}\widetilde{T}_{s_i}, \qquad \text{if } |i-j| \geq 2. \tag{8}$$

We then can define $\widetilde{T}_w$ for any $w \in S_n$ by $\widetilde{T}_w = \widetilde{T}_{s_{i_1}} \cdots \widetilde{T}_{s_{i_l}}$ where $w = s_{i_1} \cdots s_{i_l}$ is any reduced expression. Inverses of generators are given by

$$\widetilde{T}_{s_i}^{-1} = \widetilde{T}_{s_i} - (q^{\frac{1}{2}} - q^{\frac{-1}{2}})\widetilde{T}_e = \widetilde{T}_{s_i} - q^{\frac{-1}{2}}(q-1)\widetilde{T}_e. \tag{9}$$

When $q = 1$ we see that this presentation is simply that of the symmetric group algebra $\mathbb{Z}[S_n]$.

An important involution of the Hecke algebra is the so called *bar involution*. The involution is defined as

$$\sum_w a_w \widetilde{T}_w \mapsto \overline{\sum_w a_w \widetilde{T}_w} = \sum_w \overline{a_w}\,\overline{\widetilde{T}_w} \tag{10}$$

where

$$\overline{q} = q^{-1}, \qquad \overline{\widetilde{T}_w} = \left(\widetilde{T}_{w^{-1}}\right)^{-1}. \tag{11}$$

The *Kazhdan-Lusztig basis*, $\{C'_w(q) \,|\, w \in S_n\}$, is the unique basis of $H_n(q)$ such that the basis elements are invariant under the bar involution, $\overline{C'}_w(q) = C'_w(q)$ for all $w \in S_n$, and that $C'_w(q)$ in terms of the $\{\widetilde{T}_v \,|\, v \in S_n\}$ is given by

$$C'_w(q) = \sum_{v \leq w} q_{v,w}^{-1} P_{v,w}(q)\,\widetilde{T}_v, \tag{12}$$

where $P_{v,w}(q)$ are polynomials in q of degree at most $\frac{\ell(w)-\ell(v)-1}{2}$ and where we define the convenient notation $\epsilon_{v,w} = (-1)^{\ell(w)-\ell(v)}$, $q_{v,w} = (q^{\frac{1}{2}})^{\ell(w)-\ell(v)}$. These polynomials are known as the *Kazhdan-Lusztig polynomials* and in fact belong to $\mathbb{N}[q]$.

Kazhdan and Lusztig also introduced another basis $\{C_w(q)\,|\,w \in S_n\}$ with similar properties which is traditionally known as the Kazhdan-Lusztig basis, but for our purposes the $\{C'_w(q)\,|\,w \in S_n\}$ basis is more convenient. $C_w(q)$ and $C'_w(q)$ are related by $C'_w(q) = \psi(C_w(q))$, where ψ is the semilinear map defined by

$$\psi : q^{\frac{1}{2}} \mapsto q^{-\frac{1}{2}} \text{ and } \widetilde{T}_w \mapsto \epsilon_{e,w}\widetilde{T}_w. \tag{13}$$

Thus $C_w(q)$ is also bar invariant and its expression in terms of $\{\widetilde{T}_v\,|\,v \in S_n\}$ is

$$C_w(q) = \sum_{v \leq w} \epsilon_{v,w} q_{v,w} \overline{P_{v,w}(q)}\widetilde{T}_v. \tag{14}$$

As a preliminary to the proof of the existence and uniqueness of their bases Kazhdan and Lusztig also defined the following function

$$\mu(u,v) \underset{\text{def}}{=} \begin{cases} \text{coefficient of } q^{(\ell(v)-\ell(u)-1)/2} \text{ in } P_{u,v}(q) & \text{if } u < v, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Note that $\mu(u,v) = 0$ if $\ell(v) - \ell(u)$ is even since $P_{u,v}(q)$ has only integer powers of q. Also, it is well known that $P_{u,v}(q) = P_{w_0uw_0,w_0vw_0}(q)$, and therefore that $\mu(u,v) = \mu(w_0uw_0, w_0vw_0)$. Kazhdan and Lusztig showed further [8, Cor. 3.2] $\mu(u,v) = \mu(w_0v, w_0u)$, even though $P_{u,v}(q)$ and $P_{w_0v,w_0u}(q)$ are not equal in general.

In the existence proof of the Kazhdan-Lusztig basis in [8, Pf. of Thm. 1.1] an expression for the action of $\widetilde{T}_s$, s a basic transposition, on the basis element $C'_w(q)$ is given by

$$C'_w(q)\widetilde{T}_s = \begin{cases} -q^{-\frac{1}{2}}C'_w(q) + C'_{ws}(q) + \sum\limits_{\substack{v<w \\ vs<v}} \mu(v,w)C'_v(q) & \text{if } ws > w, \\ q^{\frac{1}{2}}C'_w(q) & \text{if } ws < w. \end{cases} \tag{16}$$

Along with these bases Kazhdan and Lusztig defined a preorder on S_n in order to construct representations of $H_n(q)$. This preorder, called the *right preorder*, is denoted by $\leq_R$ and is defined as the transitive closure of $\lessdot_R$ where $u \lessdot_R v$ if $C'_u(q)$ has nonzero coefficient in the expression of $C'_v(q)\widetilde{T}_w$ for some $w \in S_n$. It follows from a result in [1] that $w \leq_R v$ implies $\mathrm{sh}(v) \preceq \mathrm{sh}(w)$.

We follow the description in [7, Appendix] of the Kazhdan-Lusztig construction of an irreducible $H_n(q)$-module indexed by partition $\lambda \vdash n$. Here and henceforth the span will be over the Laurent polynomial ring $\mathbb{Z}[q^{\frac{1}{2}}, q^{-\frac{1}{2}}]$. Choosing tableau T of shape λ, we allow $H_n(q)$ to act by right multiplication on

$$K^\lambda \underset{\text{def}}{=} \mathrm{span}\{C'_w(q)\,|\,Q(w) = T\}, \tag{17}$$

regarded as the quotient

$$\mathrm{span}\{C'_v(q)\,|\,v \leq_R w\}/\mathrm{span}\{C'_v(q)\,|\,v \leq_R w, v \not\geq_R w\}. \tag{18}$$

The quotient is necessary because K^λ is not in general closed under the action of $H_n(q)$. In particular, for $\lambda \neq (1^n)$ we have the containments $K^\lambda \subset H_n(q)K^\lambda \subseteq K^\lambda \oplus \mathrm{span}\{C'_v(q)\,|\,v \leq_R w, v \not\geq_R w\}$.

4 The quantum polynomial ring and Kazhdan-Lusztig immanants

Let $x = (x_{i,j})$ be an $n \times n$-matrix of variables. The polynomial ring $\mathbb{Z}[x]$ has a natural grading $\mathbb{Z}[x] = \oplus_{r \geq 0} \mathcal{A}_r$, where $\mathcal{A}_r$ is the span of all monomials of total degree r. Further decomposing each space $\mathcal{A}_r$, we define a multigrading

$$\mathbb{Z}[x] = \bigoplus_{r \geq 0} \mathcal{A}_r = \bigoplus_{r \geq 0} \bigoplus_{L,M} \mathcal{A}_{L,M}, \tag{19}$$

where $L = \{\ell(1) \leq \ldots \leq \ell(r)\}$ and $M = \{m(1) \leq \ldots \leq m(r)\}$ are r-element multisets of $[n]$, written as weakly increasing sequences, and where $\mathcal{A}_{L,M}$ is the span of monomials whose row and column indices are given by L and M, respectively. We define the *generalized submatrix* of x with respect to (L, M) by

$$x_{L,M} = \begin{bmatrix} x_{\ell(1),m(1)} & \cdots & x_{\ell(1),m(r)} \\ x_{\ell(2),m(1)} & \cdots & x_{\ell(2),m(r)} \\ \vdots & & \vdots \\ x_{\ell(r),m(1)} & \cdots & x_{\ell(r),m(r)} \end{bmatrix}. \tag{20}$$

We refer to the space

$$\mathcal{A}_{[n],[n]} = \mathrm{span}\{x_{1,w_1} \cdots x_{n,w_n} \,|\, w \in S_n\}, \tag{21}$$

as the *immanant space*, and define the notation $x^{u,v} = x_{u_1,v_1} \cdots x_{u_n,v_n}$ for permutations $u, v \in S_n$. Immanants are a generalization of the determinant and permanent of a matrix introduced in [9].

A natural S_n-action on $\mathbb{Z}[x]$ is given by

$$g(x) \circ s_i \underset{\mathrm{def}}{=} g(xs_i), \tag{22}$$

where $g \in \mathbb{Z}[x]$ and xs_i is interpreted as the product of x and the permutation matrix of s_i.

We now define a generalization of the polynomial ring $\mathbb{Z}[x]$ called the *quantum polynomial ring*, $\mathcal{A}(n;q)$. The ring $\mathcal{A}(n;q)$ is a noncommutative $\mathbb{Z}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-algebra on n^2 generators $x = (x_{1,1} \ldots, x_{n,n})$ with relations (assuming $i < j$ and $k < \ell$),

$$\begin{aligned} x_{i,\ell}x_{i,k} &= q^{\frac{1}{2}} x_{i,k}x_{i,\ell}, \\ x_{j,k}x_{i,k} &= q^{\frac{1}{2}} x_{i,k}x_{j,k}, \\ x_{j,k}x_{i,\ell} &= x_{i,\ell}x_{j,k}, \\ x_{j,\ell}x_{i,k} &= x_{i,k}x_{j,\ell} + (q^{\frac{1}{2}} - q^{\frac{-1}{2}})x_{i,\ell}x_{j,k}. \end{aligned} \tag{23}$$

A natural basis for the quantum polynomial ring consists of the set of monomials in lexicographic order. Analogous to the multigrading of $\mathbb{Z}[x]$ is the multigrading

$$\mathcal{A}(n;q) = \bigoplus_{r \geq 0} \mathcal{A}_r(n;q) = \bigoplus_{r \geq 0} \bigoplus_{L,M} \mathcal{A}_{L,M}(n;q), \tag{24}$$

where $\mathcal{A}_r(n;q)$ is the span of all monomials of total degree r, and where $\mathcal{A}_{L,M}(n;q)$ is the span of monomials whose row and column indices are given by r-element multisets L and M of $[n]$. We again call the space $\mathcal{A}_{[n],[n]}(n;q) = \mathrm{span}\{x^{e,w} \,|\, w \in S_n\}$ the *immanant space* of $\mathcal{A}(n;q)$ or the *quantum immanant space*.

Define a right action of the Hecke algebra on $\mathcal{A}_{[n],[n]}(n;q)$ by

$$x^{e,v} \circ \widetilde{T}_{s_i} = \begin{cases} x^{e,vs_i} & \text{if } vs_i > v, \\ x^{e,vs_i} + (q^{\frac{1}{2}} - q^{\frac{-1}{2}})x^{e,v} & \text{if } vs_i < v. \end{cases} \tag{25}$$

Related to the bar involution on $H_n(q)$ is another *bar involution* on $\mathcal{A}_{[n],[n]}(n;q)$ defined by

$$\sum_w a_w x^{e,w} \mapsto \overline{\sum_w a_w x^{e,w}} = \sum_w \overline{a_w}\ \overline{x^{e,w}} \tag{26}$$

where

$$\overline{q} = q^{-1}, \qquad \overline{x^{e,w}} = x^{w_0,w_0 w} = x_{n,w_n} \cdots x_{1,w_1}. \tag{27}$$

Lemma 4.1 *The bar involutions of (10) and (26) are compatible with the action of $H_n(q)$ on $\mathcal{A}_{[n],[n]}(n;q)$. That is,*

$$\overline{x^{e,v} \circ \widetilde{T}_{s_i}} = \overline{x^{e,v}} \circ \overline{\widetilde{T}_{s_i}} \tag{28}$$

for all $v \in S_n$.

Proof: Omitted. □

It is known that there is a unique, bar-invariant basis of $\mathcal{A}_{[n],[n]}(n;q)$ closely related to the Kazhdan-Lusztig basis of the Hecke algebra. We call the elements of this basis the *Kazhdan-Lusztig immanants* $\{\mathrm{Imm}_v(x;q) \,|\, v \in S_n\}$. First appearing in Du [5],[6], this basis has the following theorem-definition. (See, e.g., [3, Thm. 5.3])

Theorem 4.2 *For any $v \in S_n$, there is a unique element $\mathrm{Imm}_v(x;q) \in \mathcal{A}_{[n],[n]}(n;q)$ such that*

$$\overline{\mathrm{Imm}_v(x;q)} = \mathrm{Imm}_v(x;q) \tag{29}$$

$$\mathrm{Imm}_v(x;q) = \sum_{w \geq v} \epsilon_{v,w} q_{v,w}^{-1} Q_{v,w}(q) x^{e,w}, \tag{30}$$

where $Q_{v,w}(q)$ are polynomials in q of degree $\leq \frac{\ell(w)-\ell(v)-1}{2}$ if $v < w$ and $Q_{v,v}(q) = 1$.

The polynomials $Q_{u,v}(q)$ above are actually the *inverse Kazhdan-Lusztig polynomials*, found in [8, Sec. 3]. They are related to the Kazhdan-Lusztig polynomials by

$$Q_{u,v}(q) = P_{w_0 v, w_0 u}(q) = P_{v w_0, u w_0}(q). \tag{31}$$

We can now describe a right action of $H_n(q)$ on the immanant space by its action on the Kazhdan-Lusztig immanants.

Corollary 4.3 *The right action of the Hecke algebra on $\mathcal{A}_{[n],[n]}(n;q)$ is described by*

$$\mathrm{Imm}_v(x;q) \circ \widetilde{T}_{s_i} = \begin{cases} q^{\frac{1}{2}} \mathrm{Imm}_v(x;q) + \mathrm{Imm}_{vs_i}(x;q) + \displaystyle\sum_{\substack{w > v \\ ws_i > w}} \mu(v,w) \mathrm{Imm}_w(x;q) & \text{if } vs_i < v, \\ -q^{\frac{-1}{2}} \mathrm{Imm}_v(x;q) & \text{if } vs_i > v. \end{cases} \tag{32}$$

Proof: Omitted. □

A deeper connection between the Kazhdan-Lusztig immanants and the Kazhdan-Lusztig basis is evident in the $\mathbb{Z}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-bilinear form on $\mathcal{A}_{[n],[n]}(n;q) \times H_n(q)$ defined by by $\langle x^{e,v}, \widetilde{T}_w \rangle = \delta_{v,w}$. Specifically, we have $\langle \mathrm{Imm}_v(x;q), C'_w(q)\rangle = \delta_{v,w}$, so the Kazhdan-Lusztig basis is dual to the basis of Kazhdan-Lusztig immanants.

In the following lemma we relate the definition of the right preorder in the Hecke algebra with these Kazhdan-Lusztig immanants. The results in the proof will also be essential in describing the relationship of the $H_n(q)$-representations associated with the Kazhdan-Lusztig basis and immanants.

Lemma 4.4 *Let $v, v' \in S_n$. Then $v \lessdot_R v'$ if $\mathrm{Imm}_{v'}(x;q)$ appears with nonzero coefficient in the Kazhdan-Lusztig immanant expansion of $\mathrm{Imm}_v(x;q) \circ \widetilde{T}_u$ for some $u \in S_n$.*

Proof: Omitted. □

With Lemma 4.4 we can now express the preorder in terms of the Kazhdan-Lusztig immanants. We can now construct $H_n(q)$-modules indexed by $\lambda \vdash n$, as in [7, Appendix], with the Kazhdan-Lusztig immanants. We choose a tableau T of shape λ and allow $H_n(q)$ to act by right multiplication on

$$V^\lambda \underset{\mathrm{def}}{=} \mathrm{span}\{\mathrm{Imm}_w(x;q) \,|\, Q(w) = T\}, \tag{33}$$

regarded as the quotient

$$\mathrm{span}\{\mathrm{Imm}_v(x;q) \,|\, v \geq_R w\}/\mathrm{span}\{\mathrm{Imm}_v(x;q) \,|\, v \geq_R w, v \not\leq_R w\}. \tag{34}$$

The quotient is necessary because like K^λ, V^λ is not in general closed under the action of $H_n(q)$. In particular, whenever $\lambda \neq (1^n)$ we have the containments

$$V^\lambda \subset H_n(q)V^\lambda \subseteq V^\lambda \oplus \mathrm{span}\{\mathrm{Imm}_v(x;q) \,|\, v \geq_R w, v \not\leq_R w\}. \tag{35}$$

5 Generalized submatrices and vanishing properties of immanants

In [11] Rhoades and Skandera stated conditions on immanants $\mathrm{Imm}_w(x)$ in $\mathbb{Z}[x]$ and on $n \times n$-matrices A which imply that $\mathrm{Imm}_w(A) = 0$. Here we present new, analogous vanishing results for immanants in $\mathcal{A}_{[n],[n]}(n;q)$. Specifically we will state conditions on quantum immanants $\mathrm{Imm}_w(x;q)$ in $\mathcal{A}(n;q)$ and on generalized submatrices $x_{L,M}$ of the quantum matrix x, which imply that $\mathrm{Imm}_w(x_{L,M};q) = 0$. Using these results we can eliminate the quotient needed in the construction (34) of the $H_n(q)$-modules. This provides a quantum analog of the authors' results in [3].

To express the vanishing results we need to define the *row repetition partition* of an $n \times n$-matrix A by

$$\mu_{[j]}(A) \underset{\mathrm{def}}{=} (\mu_1, \ldots, \mu_k), \tag{36}$$

where k is the number of distinct rows in the $n \times j$-submatrix $A_{[n],[j]}$, and $\mu_1, \ldots, \mu_k$ are the multiplicities with which distinct rows appear, written in weakly decreasing order. Also we define the permutation $w_{[j]} \in S_j$ from $w \in S_n$ by arranging $1, \ldots, j$ in the same relative order of the first j terms in the one line notation of w.

Lemma 5.1 *Fix a permutation $w \in S_n$ and indices $1 \leq j \leq n$. If $\mathrm{sh}(w_{[j]}) \not\geq \mu_{[j]}(x_{L,[n]})$, then*

$$\mathrm{Imm}_w(x_{L,[n]};q) = 0. \tag{37}$$

An immediate consequence of this vanishing result follows after defining a partial order found in [11]. A partial order on standard tableaux is the *iterated dominance of tableaux*. Given two standard tableau T, U both having n boxes, we define $U \lhd_I T$ if for $j = 1, \ldots, n$ we have

$$\mathrm{sh}(U_{[j]}) \prec \mathrm{sh}(T_{[j]}), \tag{38}$$

where $U_{[j]}$ is the subtableau of U consisting of all entries less than or equal to j.

Corollary 5.2 *Fix a partition $\lambda \vdash n$ and define the multiset $L = 1^{\lambda_1} \cdots n^{\lambda_n}$, where n^k is shorthand for n appearing k times. For each permutation w satisfying $\mathrm{sh}(w) \not\preceq \lambda$ or satisfying $\mathrm{sh}(w) = \lambda$ and $Q(w) \neq T(\lambda)$, we have that $\mathrm{Imm}_w(x_{L,[n]}) = 0$.*

Proof: If w satisfies $\mathrm{sh}(w) \not\preceq \lambda$ then the case with $j = n$ of Lemma 5.1 implies that $\mathrm{Imm}_w(x_{L,[n]}; q) = 0$. Suppose that $\mathrm{sh}(w) = \lambda$ and $Q(w) \neq T(\lambda)$. Since the tableau $T(\lambda)$ is greatest in iterated dominance among all tableaux of shape λ, we have that $Q(w) \lhd_I T(\lambda)$ and there exists an index j such that

$$\mathrm{sh}(Q(w)_{[j]}) \prec \mathrm{sh}(T(\lambda)_{[j]}) = \mu_{[j]}(x_{L,[n]}). \tag{39}$$

Then by the fact that $\mathrm{sh}(w_{[j]}) = \mathrm{sh}(Q(w)_{[j]})$ we see that $\mathrm{sh}(w_{[j]}) \prec \mu_{[j]}(x_{L,[n]})$, which by Lemma 5.1 implies that $\mathrm{Imm}_w(x_{L,[n]}; q) = 0$. $\square$

We can define a right action of $H_n(q)$ on $\mathcal{A}_{L,[n]}(n; q)$ by the formula

$$(x_{L,[n]})^{e,w} \circ \widetilde{T}_s = \begin{cases} (x_{L,[n]})^{e,ws}, & ws > w \\ (x_{L,[n]})^{e,ws} + (q^{\frac{1}{2}} - q^{\frac{-1}{2}})(x_{L,[n]})^{e,w}, & ws < w. \end{cases} \tag{40}$$

We can then extend this action for the Kazhdan-Lusztig immanants evaluated at generalized submatrices.

Corollary 5.3 *Fix $u \in S_n$ and an n-element multiset L of $[n]$. For a basic transposition s, the right action of $H_n(q)$ on the element $\mathrm{Imm}_u(x_{L,[n]}; q)$ of $\mathcal{A}_{L,[n]}(n; q)$ is given by*

$$\mathrm{Imm}_u(x_{L,[n]}; q) \circ \widetilde{T}_s =$$
$$\begin{cases} q^{\frac{1}{2}} \mathrm{Imm}_u(x_{L,[n]}; q) + \mathrm{Imm}_{us}(x_{L,[n]}; q) + \displaystyle\sum_{\substack{w > u \\ ws > w}} \mu(u, w) \mathrm{Imm}_w(x_{L,[n]}; q), & us < u \\ -q^{\frac{-1}{2}} \mathrm{Imm}_u(x_{L,[n]}; q), & us > u. \end{cases} \tag{41}$$

Proof: For $u \in S_n$ the Kazhdan-Lusztig immanant indexed by u evaluated at the matrix $x_{L,[n]}$ is given by

$$\mathrm{Imm}_u(x_{L,[n]}; q) = \sum_{w \geq u} \epsilon_{u,w} q_{u,w}^{-1} Q_{u,w}(q) (x_{L,[n]})^{e,w}. \tag{42}$$

Now we have an action of $H_n(q)$ on the immanants by (40),

$$\begin{aligned}
\mathrm{Imm}_u(x_{L,[n]};q)\circ\widetilde{T}_s &= \sum_{w\geq u}\epsilon_{u,w}q_{u,w}^{-1}Q_{u,w}(q)\,(x_{L,[n]})^{e,w}\circ\widetilde{T}_s\\
&= \sum_{\substack{w\geq u\\ ws>w}}\epsilon_{u,w}q_{u,w}^{-1}Q_{u,w}(q)(x_{L,[n]})^{e,ws}\\
&\quad+\sum_{\substack{w\geq u\\ ws<w}}\epsilon_{u,w}q_{u,w}^{-1}Q_{u,w}(q)\left((x_{L,[n]})^{e,ws}+(q^{\frac{1}{2}}-q^{\frac{-1}{2}})(x_{L,[n]})^{e,w}\right)\\
&= -q^{\frac{-1}{2}}\sum_{\substack{w\geq u\\ ws>w}}\epsilon_{u,w}q_{u,w}^{-1}Q_{u,ws}(q)(x_{L,[n]})^{e,w}\\
&\quad+\sum_{\substack{w\geq u\\ ws<w}}\epsilon_{u,w}q_{u,w}^{-1}\left((q^{\frac{1}{2}}-q^{\frac{-1}{2}})Q_{u,w}(q)-q^{\frac{1}{2}}Q_{u,ws}(q)\right)(x_{L,[n]})^{e,w}.
\end{aligned}\tag{43}$$

If $us>u$ we know that $Q_{u,w}(q)=Q_{u,ws}(q)$ for any permutation w. Thus we have from (43) the action of $\widetilde{T}_s$ is

$$\begin{aligned}
\mathrm{Imm}_u(x_{L,[n]};q)\circ\widetilde{T}_s &= -q^{\frac{-1}{2}}\sum_{\substack{w\geq u\\ ws>w}}\epsilon_{u,w}q_{u,w}^{-1}Q_{u,w}(q)(x_{L,[n]})^{e,w}\\
&\quad+\sum_{\substack{w\geq u\\ ws<w}}\epsilon_{u,w}q_{u,w}^{-1}\left((q^{\frac{1}{2}}-q^{\frac{-1}{2}})Q_{u,w}(q)-q^{\frac{1}{2}}Q_{u,w}(q)\right)(x_{L,[n]})^{e,w}\\
&= -q^{\frac{-1}{2}}\mathrm{Imm}_u(x_{L,[n]};q),
\end{aligned}\tag{44}$$

as we expected.

If $us<u$ we know that $Q_{us,w}(q)=Q_{us,ws}(q)$ for any permutation w. By careful application of the recursive formula for the inverse Kazhdan-Lusztig polynomials we can also observe the following relationships. If $ws>w$ then we see that

$$Q_{u,ws}(q)=Q_{us,w}(q)-qQ_{u,w}(q)+\sum_{\substack{u<v\leq w\\ v<vs}}q_{v,w}\mu(u,v)Q_{v,w}(q).\tag{45}$$

If $ws<w$ then we see that

$$Q_{u,w}(q)+qQ_{u,ws}(q) = Q_{us,ws}(q)+\sum_{\substack{u<v\leq ws\\ v<vs}}q_{v,ws}\mu(u,v)Q_{v,ws}(q)\tag{46}$$

$$= Q_{us,w}(q)+\sum_{\substack{u<v\leq ws\\ v<vs}}q_{v,ws}\mu(u,v)Q_{v,ws}(q).\tag{47}$$

Thus we have from (43) the action of $\widetilde{T}_s$ is

$$
\begin{aligned}
&\mathrm{Imm}_u(x_{L,[n]};q)\circ\widetilde{T}_s\\
&\quad= -q^{\frac{-1}{2}}\sum_{\substack{w\geq u\\ ws>w}}\epsilon_{u,w}q_{u,w}^{-1}\left(Q_{us,w}(q)-qQ_{u,w}(q)+\sum_{\substack{u<v\leq w\\ v<vs}}q_{v,w}\mu(u,v)Q_{v,w}(q)\right)(x_{L,[n]})^{e,w}\\
&\quad+\sum_{\substack{w\geq u\\ ws<w}}\epsilon_{u,w}q_{u,w}^{-1}\left(q^{\frac{1}{2}}Q_{u,w}(q)-q^{\frac{-1}{2}}\left(Q_{us,w}(q)+\sum_{\substack{u<v\leq ws\\ v<vs}}q_{v,ws}\mu(u,v)Q_{v,ws}(q)\right)\right)(x_{L,[n]})^{e,w}\\
&\quad=\sum_{\substack{w\geq u\\ ws>w}}\epsilon_{u,w}q_{u,w}^{-1}\left(q^{\frac{1}{2}}Q_{u,w}(q)-q^{\frac{-1}{2}}Q_{us,w}(q)-q^{\frac{-1}{2}}\sum_{\substack{u<v\leq w\\ v<vs}}q_{v,w}\mu(u,v)Q_{v,w}(q)\right)(x_{L,[n]})^{e,w}\\
&\quad+\sum_{\substack{w\geq u\\ ws<w}}\epsilon_{u,w}q_{u,w}^{-1}\left(q^{\frac{1}{2}}Q_{u,w}(q)-q^{\frac{-1}{2}}Q_{us,w}(q)-q^{\frac{-1}{2}}\sum_{\substack{u<v\leq ws\\ v<vs}}q_{v,ws}\mu(u,v)Q_{v,ws}(q)\right)(x_{L,[n]})^{e,w}\\
&\quad= q^{\frac{1}{2}}\mathrm{Imm}_u(x_{L,[n]};q)+\mathrm{Imm}_{us}(x_{L,[n]};q)+\sum_{\substack{v>u\\ v<vs}}\mu(u,v)\mathrm{Imm}_v(x_{L,[n]};q),
\end{aligned}
\tag{48}
$$

as we expected. □

We can now see that the right $H_n(q)$-action defined in Corollary 5.3 actually describes an $H_n(q)$-module if we evaluate the immanants at generalized submatrices.

Theorem 5.4 *Let $\lambda\vdash n$ and set $L=1^{\lambda_1}\cdots n^{\lambda_n}$. Define*

$$W^\lambda \underset{\mathrm{def}}{=} \mathrm{span}\{\mathrm{Imm}_w(x_{L,[n]};q)\,|\,Q(w)=T(\lambda)\},\tag{49}$$

where $T(\lambda)$ is the superstandard tableau of shape λ. Then W^λ is an $H_n(q)$-module.

Proof: By (35) we know that it suffices to show that $\mathrm{Imm}_v(x_{L,[n]};q)=0$ for $v>_R w$ where $Q(w)=T(\lambda)$. Since $v>_R w$ then we know that $\mathrm{sh}(w)\succ\mathrm{sh}(v)$. The row multiplicity partition of $x_{L,[n]}$ is $\mu(x_{L,[n]})=\lambda$. So $\mathrm{sh}(v)\prec\mathrm{sh}(w)=\mu(x_{L,[n]})$. Thus $\mathrm{sh}(v)\not\succeq\mu(x_{L,[n]})$. Therefore, by Lemma 5.1, $\mathrm{Imm}_v(x_{L,[n]};q)=0$ for all $v>_R w$. □

The condition for inclusion in the basis of this module is $Q(w)^\top=T(\lambda)$ unlike the condition, $Q(w)=T$ where $\mathrm{sh}(T)=\lambda$, used in the definition of V^λ above. The need for the change in conditions is due to the result Corollary 5.2.

We would now like to show that these modules, W^λ, are isomorphic to the modules constructed by the action $H_n(q)$ on the Kazhdan-Lusztig basis. We shall then show that the action of $\widetilde{T}_{s_i}$ on either basis yields equal matrices, up to ordering of the basis elements. Let $\rho_1: H_n(q)\to\mathrm{End}(K^\lambda)$ and $\rho_2: H_n(q)\to\mathrm{End}(W^\lambda)$ be the representations of $H_n(q)$ defined by the right actions described in (16) and Corollary 4.3, respectively.

Theorem 5.5 *Let $X_1(h)$, $X_2(h)$ be the matrices of $\rho_1(h)$, $\rho_2(h)$ with respect to the Kazhdan-Lusztig basis and the Kazhdan-Lusztig immanant basis. Then $X_1(h)=X_2(h)$.*

Proof: First, we construct K^λ as in (17) with $T = T(\lambda)$. Let $B = \{v \in S_n \,|\, Q(v) = T(\lambda)\}$. From Lemma 2.1 we see that if $C_w(q)$ is a basis element of K^λ, i.e. $w \in B$, then $Q(ww_0)^\top = Q(w) = T(\lambda)$. Thus if $w \in Bw_0$, then $\mathrm{Imm}_w(x_{L,[n]}; q)$ is a basis element of W^λ, as in (49). Define coefficients $a^{s_i}_{v,w}$ for each generators s_i of S_n and $v, w \in B$ so that

$$C'_v(q)\widetilde{T}_{s_i} = \sum_{w \in B} a^{s_i}_{v,w} C'_w(q). \tag{50}$$

Then from the proof of Lemma 4.4 and Corollary 5.3 we see that for all $v \in B$

$$\mathrm{Imm}_{vw_0}(x_{L,[n]}; q) \circ \widetilde{T}_{s_i} = \sum_{w \in B} a^{s_i}_{v,w} \mathrm{Imm}_{ww_0}(x_{L,[n]}; q). \tag{51}$$

Thus $X_1(\widetilde{T}_{s_i}) = X_2(\widetilde{T}_{s_i})$. Since any element of $v \in S_n$ is a product of generators we have that $X_1(\widetilde{T}_v) = X_2(\widetilde{T}_v)$ and thus for any element $h \in H_n(q)$ we have that $X_1(h) = X_2(h)$. $\square$

Corollary 5.6 *The modules $\mathbb{C}(q^{\frac{1}{2}}) \otimes W^\lambda$ indexed by partitions $\lambda \vdash n$ are the irreducible $\mathbb{C}(q^{\frac{1}{2}}) \otimes H_n(q)$-modules.*

This result follows immediately from the fact that the modules K^λ are the irreducible $H_n(q)$-modules.

References

[1] D. BARBASCH AND D. VOGAN. Primitive ideals and orbital integrals in complex exceptional groups. *J. Algebra*, **80**, 2 (1983) pp. 350–382.

[2] A. BJÖRNER AND F. BRENTI. *Combinatorics of Coxeter groups*, vol. 231 of *Graduate Texts in Mathematics*. Springer, New York (2005).

[3] C. BUEHRLE AND M. SKANDERA. Relations between the Clausen and Kazhdan-Lusztig representations of the symmetric group. *J. Pure Appl. Algebra*, **214**, 5 (2010) pp. 689–700.

[4] M. CLAUSEN. Multivariate polynomials, standard tableaux, and representations of symmetric groups. *J. Symbolic Comput.*, **11**, 5-6 (1991) pp. 483–522. Invariant-theoretic algorithms in geometry (Minneapolis, MN, 1987).

[5] J. DU. Canonical bases for irreducible representations of quantum GL_n. *Bull. London Math. Soc.*, **24**, 4 (1992) pp. 325–334.

[6] J. DU. Canonical bases for irreducible representations of quantum GL_n. II. *J. London Math. Soc. (2)*, **51**, 3 (1995) pp. 461–470.

[7] M. HAIMAN. Hecke algebra characters and immanant conjectures. *J. Amer. Math. Soc.*, **6**, 3 (1993) pp. 569–595.

[8] D. KAZHDAN AND G. LUSZTIG. Representations of Coxeter groups and Hecke algebras. *Invent. Math.*, **53** (1979) pp. 165–184.

[9] D. E. LITTLEWOOD AND A. R. RICHARDSON. Group characters and algebra. *Phil. Trans. R. Soc. Lond. A*, **233** (1934) pp. 99–141.

[10] B. RHOADES AND M. SKANDERA. Kazhdan-Lusztig immanants and products of matrix minors. *J. Algebra*, **304**, 2 (2006) pp. 793–811.

[11] B. RHOADES AND M. SKANDERA. Bitableaux and the dual canonical basis of the polynomial ring (2008). Submitted to *Adv. in Math.*

[12] B. SAGAN. *The Symmetric Group*. Springer, New York (2001).

[13] M. TAŞKIN. Properties of four partial orders on standard Young tableaux. *J. Combin. Theory Ser. A*, **113**, 6 (2006) pp. 1092–1119.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 467–474

On k-crossings and k-nestings of permutations

Sophie Burrill[1] and Marni Mishna[1] and Jacob Post[2]

[1]*Department of Mathematics, Simon Fraser University, Burnaby, Canada, V5A 1S6*
[2]*Department of Computer Science, Simon Fraser University, Canada, V5A 1S6*

Abstract. We introduce k-crossings and k-nestings of permutations. We show that the crossing number and the nesting number of permutations have a symmetric joint distribution. As a corollary, the number of k-noncrossing permutations is equal to the number of k-nonnesting permutations. We also provide some enumerative results for k-noncrossing permutations for some values of k.

Résumé. Nous introduisons les k-chevauchement d'arcs et les k-empilements d'arcs de permutations. Nous montrons que l'index de chevauchement et l'index de empilement ont une distribution conjointe symétrique pour les permutations de taille n. Comme corollaire, nous obtenons que le nombre de permutations n'ayant pas un k-chevauchement est égal au nombre de permutations n'ayant un k-empilement. Nous fournissons également quelques résultats énumératifs.

Keywords: crossing, nesting, permutation, enumeration

1 Introduction

Nestings and crossings are equidistributed in many combinatorial objects, such as matchings, set partitions, permutations, and large classes of embedded labelled graphs [2, 3, 5]. More surprising is the symmetric joint distribution of the crossing and nesting numbers: A set of k arcs forms a k-crossing (respectively nesting) if each of the $\binom{k}{2}$ pairs of arcs cross (resp. nest). The *crossing number* of an object is the largest k for which there is a k-crossing, and the *nesting number* is defined similarly. Chen *et al.* [2] proved the symmetric joint distribution of the nesting and crossing numbers for set partitions and matchings. Although they describe explicit involutions, they do not use simple local operations on the partitions. Recently, de Mier [5] interpreted the work of Krattenthaler [6] to show that k-crossings and k-nestings satisfy a similar distribution in embedded labelled graphs.

A hole in this family of results is the extension of the notions of k-crossings and k-nestings to permutations. This note fills this gap. We also give exact enumerative formulas for permutations of size n with crossing numbers 1 (non-crossing) and $\lceil n/2 \rceil$.

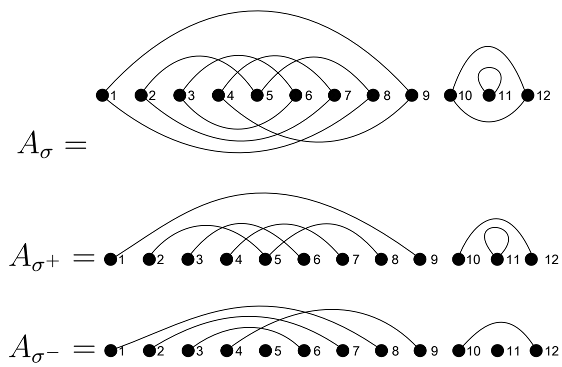


Fig. 1: An arc diagram representation for the permutation $\sigma = [9\,5\,6\,7\,8\,3\,2\,1\,4\,12\,11\,10]$, and its decomposition into upper and lower arc diagrams (A_σ^+, A_σ^-). In this example, $\mathrm{cr}(\sigma) = 4$, $\mathrm{ne}(\sigma) = 3$, and the degree sequence is given by $D_\sigma = (1,0)(1,0)(1,0)(1,0)(1,1)(0,1)(0,1)(0,1)(0,1)(1,0)(1,1)(0,1)$.

2 Introducing k-crossings and k-nestings of permutations

2.1 Crossings and nestings

The *arc annotated sequence* associated to the permutation $\sigma \in \mathfrak{S}_n$ is the directed graph on the vertex set $V(\sigma) = \{1, \ldots, n\}$ with arc set $A(\sigma) = \{(a, \sigma(a)) : 1 \leq a \leq n\}$, drawn in a particular way. It is also known as the standard representation, or simply, the arc diagram. It is embedded in the plane by drawing an increasing linear sequence of the vertices, with edges $(a, \sigma(a))$ satisfying $a \leq \sigma(a)$ drawn above the vertices (the upper arcs), and the remaining lower arcs satisfying $a > \sigma(a)$ drawn below. We refer to this graph as A_σ; the subgraph induced by the upper arcs and $V(\sigma)$ is A_σ^+; and the subgraph induced by the lower arcs and $V(\sigma)$ is A_σ^-. Additionally, we reverse the orientation of the arcs in A_σ^-, and view it as a classic arc diagram above the horizon. Because of these rules, the direction of the arcs is determined, and hence we simplify our drawings by not showing arrows on the arcs.

These two subgraphs are arc diagrams in their own right: for example A_σ^- represents a set partition, and A_σ^+ is a set partition with some additional loops.

Crossings and nestings are defined for permutations by considering the upper and lower arcs separately. A crossing is a pair of arcs $\{(a, \sigma(a)), (b, \sigma(b))\}$ satisfying either $a < b \leq \sigma(a) < \sigma(b)$ (an upper crossing) or $\sigma(a) < \sigma(b) < a < b$ (a lower crossing). A nesting is a pair of arcs $(a, \sigma(a))$ $(b, \sigma(b))$ satisfying $a < b \leq \sigma(b) < \sigma(a)$ (an upper nesting) or $\sigma(a) < \sigma(b) < b < a$ (a lower nesting).

There is a slight asymmetry to the treatment of upper and lower arcs in this definition which we shall see is inconsequential. However, the reader should recall that what is considered a crossing (resp. nesting) in the upper diagram is elsewhere called an *enhanced* crossing (resp. enhanced nesting).

Crossings and nestings were defined in this way by Corteel [3] because they represent better known permutation statistics. Corteel's Theorem 1 states that the number of top arcs in this representation of a permutation is equal to the number of weak excedances, the number of arcs on the bottom is the number of descents, each crossing is equivalent to an occurrence of the pattern $2-31$, and each nesting is an occurrence of the pattern $31-2$. Corteel's Proposition 4 states nestings and crossings occur in equal number across all permutations of length n.

2.2 k-nestings and k-crossings

To generalize Corteel's work we define k-crossings and k-nestings in the same spirit as set partitions and matchings. A k-crossing in a permutation ard diagram A_σ is a set of k arcs $\{(a_i, \sigma(a_i)) : 1 \leq i \leq k\}$ that satisfy either the relation $a_1 < a_2 < \cdots < a_k \leq \sigma(a_1) < \sigma(a_2) < \cdots < \sigma(a_k)$ (upper k-crossing) or $\sigma(a_1) < \sigma(a_2) < \cdots < \sigma(a_k) < a_1 < a_2 < \cdots < a_k$ (lower k-crossing). Similarly, a k-nesting is a set of k arcs $\{(a_i, \sigma(a_i)) : 1 \leq i \leq n\}$ that satisfy either the relation $a_1 < a_2 < \cdots < a_k \leq \sigma(a_k) < \cdots < \sigma(a_2) < \sigma(a_k)$ (upper k-nesting) or $\sigma(a_1) < \sigma(a_2) < \cdots < \sigma(a_k) < a_k < \cdots < a_2 < a_1$ (lower k-nesting).

The *crossing number* of a permutation σ, denoted by $\mathrm{cr}(\sigma)$, is the size of the largest k such that A_σ contains a k-crossing. In this case we also say σ is $k+1$-noncrossing. Likewise, the nesting number of a permutation $\mathrm{ne}(\sigma)$ is the size of the largest nesting in A_σ, and define $k+1$-noncrossing similarly. Occasionally we consider the top and lower diagrams in their own right as *graphs*, and then we use the definition of deMier [5], and hence distinguish separately the enhanced crossing number of the *graph* A_σ^+ denoted $\mathrm{cr}^*(A_\sigma^+)$ from the permutation crossing number, and likewise for the enhanced nesting number ne^*. The number of permutations of $\mathfrak{S}_n$ with crossing number equal to k is $\mathrm{C_n}(k)$, and we likewise define $\mathrm{N_n}(k)$ for nestings.

The *degree sequence* D_g of a graph g is the sequence of indegree and outdegrees of the vertices, when considered as a directed graph:

$$D_g \equiv (D_g(i))_i = \left(\mathrm{indegree}_g(i), \mathrm{outdegree}_g(i)\right)_{i=1}^n.$$

Some sources call these left-right degree sequences since in other arc diagrams the incoming arcs always come in on the left, and the outgoing arcs go out to the right. As a graph, the degree sequence of a permutation is trivial: $(1,1)^n$, since a permutation is a map in which every point has a unique image, and a unique pre-image. To define a more useful entity, we define the degree sequence of a permutation to be the degree sequence of only the upper arc diagram: $D_\sigma \equiv D_{A_\sigma^+}$. The degree sequence defined by the lower arc diagram can be computed coordinate-wise directly from the upper by simple transformations given in Table 2.2, and we denote this sequence $\overline{D_\sigma}$. (The sums of the vertex degrees is not (1,1) because the lower arcs have their orientation reversed, and hence the indegree, and the outdegree have switched) An example is in Figure 1. The vertices with degree $(0,1)$ are called "openers" and those with degree $(1,0)$ are "closers".

The main theorem can now be stated.

Theorem 1 *Let $NC_n(i,j,D)$ be the number of permutations of n with crossing number i, nesting number j, and left-right degree sequence specified by D. Then*

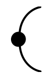
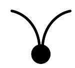
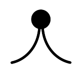
$$NC_n(i,j,D) = NC_n(j,i,D). \tag{1}$$

There is an explicit involution behind this enumerative result, and the proof is in Section 4.

2.3 Preliminary enumerative results

The number of permutations of $\mathfrak{S}_n$ with crossing number equal to k is directly computable for small values of n and k.

We immediately notice the first column of Table 2, the non-crossing permutations, are counted by Catalan numbers: $\mathrm{C_n}(1) = \frac{1}{n+1}\binom{2n}{n}$. This has a simple explanation: non-crossing *partitions* have long

Type	**vertex** i	$D_\sigma(i)$	$\overline{D_\sigma}(i)$
opener		(1,0)	(1,0)
closer		(0,1)	(0,1)
loop		(1,1)	(0,0)
upper transient		(1,1)	(0,0)
lower transient		(0,0)	(1,1)

Tab. 1: *The five vertex types that appear in permutations, and their associated upper degree value, and lower degree value.*

$n\backslash k$	1	2	3	4	5
1	1				
2	2				
3	5	1			
4	14	10			
5	42	76	2		
6	132	543	45		
7	429	3904	701	6	
8	1430	29034	9623	233	
9	4862	225753	126327	5914	24

Tab. 2: $\mathrm{C_n}(k)$*: The number of permutations of* $\mathfrak{S}_n$ *with crossing number* k*. A crossing number of 1 is equivalent to non-crossing.*

been known to be counted by Catalan numbers and there is a simple bijection between non-crossing permutations and non-crossing partitions. Essentially, to go from a non-crossing permutation to a non-crossing partition, flip the arc diagram upside down, convert the loops to fixed points, and then remove the lower arcs. This defines a unique set partition, and is easy to reverse. This bijection is easy to formalize, but it is not the main topic of this note.

3 Enumeration of maximum nestings and crossings

To get a sense of how Theorem 1 is proved, and to obtain some new enumerative results, we consider the set of maximum nestings and crossings. A *maximum nesting* is the largest possible: a $\lceil n/2 \rceil$-nesting is maximum in a permutation on n elements. We can compute $\mathrm{N_n}(\lceil n/2 \rceil)$ explicitly.

Theorem 2 *The number of permutations with a maximum nesting satisfies the following formula:*

$$\mathrm{N_n}(\lceil n/2 \rceil) = \begin{cases} m! & n = 2m+1 \\ 2(m+1)! - (m-1)! - 1 & n = 2m \end{cases}. \tag{2}$$

Proof: We divide the result into a few cases, but each one is resolved the same way: For each permutation $\sigma \in \mathfrak{S}_n$ with a maximum nesting, the $\lceil n/2 \rceil$-nesting comes from either A^+_σ or A^-_σ, and in most cases defines that subgraph. Once one side is fixed, and there is a given degree sequence, it is straightforward to compute the number of ways to place the remaining arcs. Some cases are over counted, and tallying these gives the final result.

Odd n: $n = 2m + 1$ To achieve an $m + 1$-nesting, it must be an enhanced nesting in the upper arc diagram, and it uses all vertices, including a loop: $\sigma(i) = n - i + 1 : 1 \le i \le m$. It remains to define $\sigma(i)$ for $m < i \le n$. The lower degree sequence is fixed, and so $1 \le \sigma(i) < m$ for each i, but other than that there is no restriction. Thus, there are $m!$ possibilities.

Even n: $n = 2m$ The even case is slightly more complicated, owing to the fact that three different ways to achieve an m-nesting:

1. **An m-nesting in A^+_σ** These permutations satisfy $\sigma(i) = n - i, 1 \le i \le m$. As before, there are $m!$ ways to define $\sigma(i), m < i \le n$.

2. **An m-nesting in A^-_σ** These permutations satisfy $\sigma(n - i) = i, 1 \le i \le m$. Again, there are $m!$ possibilities to define $\sigma(i), 1 < i \le m$. Only the involution $[n\, n-1 \ldots 2\, 1]$ is in the intersection of these sets.

3. **An enhanced m-nesting in A^+_σ** If the m-nesting uses only $2m - 1$ vertices, there is one left over. It must either be a lower transient vertex, or a loop since there is nothing left to connect to it. We count these by considering the different ways to construct it from a smaller permutation diagram. Suppose we have a permutation with an m-nesting on $2m - 1$ vertices. By the first part, we know there are $(m - 1)!$ of these. We place it on $2m$ points, by first selecting our special vertex i, and placing the permutation on the rest. There are $2m$ ways to pick this special vertex. Finally, we create the new permutation σ by connecting the new vertex to the rest of the structure. We choose a point j to be the value $\sigma(i)$. We can choose i and thus i is a loop. Otherwise, j must be before the loop in σ'. We then set $\sigma^{-1}(i)$ to be $\sigma'^{-1}(j)$. There are m choices for $\sigma(i)$.

Over counting we have counted twice the family of diagrams with two loops in the center. There are $(m - 1)!$ of these.

Putting all of the pieces together, and simplifying the expression we get the formula:

$$\mathrm{N}_{2m}(m) = 2(m + 1)! - (m - 1)! - 1.$$

□

This proof suggests a direct involution on the permutation which switches a maximum nesting for a maximum crossing, since the degree sequence of a nesting and a crossing have the same shape. Thus the formula for maximum crossings is the same. However, in certain diagrams, this involution sends a k-crossing to a $k+1$-nesting, and so it can not be used to prove equidistribution for general k.

3.1 Other enumerative questions

From the formula, we see that a very small proportion of permutations have maximum crossings, ($\leq 2\frac{n+1}{2}!/n!$) or are non-crossing $\approx 4^n n^{-3/2}/n!$. What can be said of the nature of the distribution, or the even simply the average crossing number? What is the nature of the generating function $P(z;u)$ where u marks the crossing number, or even simply the generating function for k-noncrossing permutations? Bousquet-Mélou and Xin [1] consider this question for partitions: 2-noncrossing partitions are counted by Catalan numbers, (as we mentioned before), and thus the generating function is algebraic; the counting sequence for 3-noncrossing partitions is P-recursive, and so the generating function is D-finite, and they conjecture that the generating function for k-noncrossing partitions, $k > 3$ are likely not D-finite. How can these results be adapted to permutations, given the similar structure?

4 Proof of Theorem 1

We restate and prove our main theorem. The proof first decomposes a permutation into its upper and lower arc diagrams and then applies the results for graphs separately to each part.

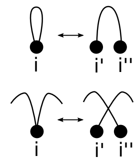
Theorem 1 *Let $NC_n(i,j,D)$ be the number of permutations of n with crossing number i, nesting number j, and left-right degree sequence specified by D. Then*

$$NC_n(i,j,D) = NC_n(j,i,D).$$

Proof: We consider the top and the bottom in turn, and to each apply the consequence of Chen *et al.*, that the pair $(\mathrm{cr}(g), \mathrm{ne}(g)))$ is symmetrically distributed across all arc diagrams g on n vertices with degree sequence a fixed element of $\{(0,0),(0,1),(1,0),(1,1)\}^n$, which is the case for our graphs here. Furthermore, we apply their degree preserving involution $\Psi : \mathfrak{S}_n \to \mathfrak{S}_n$ which swaps nesting and crossing number. That is, $D_\sigma = D_{\Psi(\sigma)}$, and $\mathrm{ne}(\sigma) = \mathrm{cr}(\Psi(\sigma))$, $\mathrm{cr}(\sigma) = \mathrm{ne}(\Psi(\sigma))$.

This consequence can also be seen as an example of de Miers' Theorem 3.3 [5]. Vertices with maximum left or right degree at most one avoid multiple edges, as is the case with our graphs, and hence the result applies. Furthermore, her interpretation of graphs as fillings of growth diagrams apply.

In order to apply the above results, the first step is to re-write A_σ^+ so that we only consider proper crossings and nestings instead of enhanced crossing and nestings. This is a common trick, known as inflation. Essentially, we create the graph g from A_σ^+ by adding some supplementary vertices to eliminate loops and transitory vertices:



Now each nesting and crossing is proper, and by [5, Lemma 3.4] $\mathrm{ne}^*(A_\sigma^+) = \mathrm{ne}(g)$ and $\mathrm{cr}^*(A_\sigma^+) = \mathrm{cr}(g)$.

Let Ψ be the map on embedded labelled graphs described implicitly in de Mier's proof. Because Ψ is a left-right degree preserving map, we can identify the supplementary vertices in $\Psi(g)$ to get a graph with the correct kind of vertices. Call this new graph g'. We now extend the definition of Ψ to A_σ^+ by $\Psi(A_\sigma^+) \equiv g'$.

Consider the pair of graphs $(\Psi(A_\sigma^+), \Psi(A_\sigma^-))$.

Proving our main theorem now reduces to showing that there is a unique $\tau \in \mathfrak{S}_n$ such that $A_\tau = (\Psi(A_\sigma^+), \Psi(A_\sigma^-))$, which we do next. For every vertex in A_τ the indegree and the outdegree are equal to one. This is because the left-right degree sequence of both the top and the bottom are preserved in the map, and hence the vector sum of their degree sequence is unchanged, i.e. $(1,1)^n$, and has all the correct partial sum properties. The map is a bijection and so τ is unique.

This map swaps the upper nesting and the upper crossing number, and also the lower nesting and the lower crossing number. Thus $\mathrm{cr}(\tau) = \max\{\mathrm{cr}^*(A_\tau^+), \mathrm{cr}(A_\tau^-)\} = \max\{\mathrm{ne}^*(A_\sigma^+), \mathrm{ne}(A_\sigma^-)\} = \mathrm{ne}(\sigma)$. Thus, the crossing and the nesting number are switched under the map Ψ. $\square$

Figure 2 illustrates our involution on an example. Remark that the degree sequence is fixed.

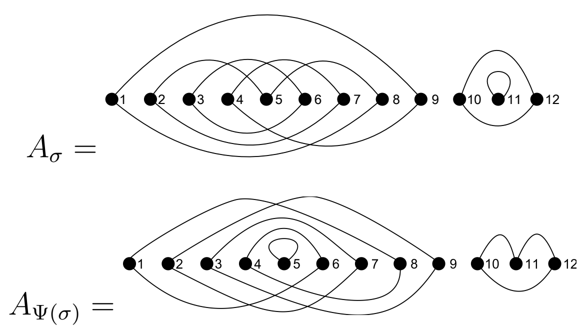


Fig. 2: *The permutation σ and its image in the involution $\Psi(\sigma)$. Note that* $\mathrm{ne}(\Psi(\sigma)) = 4$, $\mathrm{cr}(\Psi(\sigma)) = 3$.

4.1 Equidistribution in permutation subclasses

Involutions are in bijection with partial matchings, and have thus been previously considered. What of other subclasses of permutations? The map presented here does not fix involutions, because loops are mapped to upper transient vertices, but it does fix any class that is closed under degree sequence, for example, permutations with no lower transitory vertices, or permutations with no upper transitory vertices nor loops. These conditions have interpretations in terms of other permutation statistics, if we consider the initial motivations of Corteel.

5 Conclusions and open questions

The main open question, aside from the enumerative, and probabilistic questions we have already raised, is to find a direct permutation description of our involution, i.e. a description avoiding the passage through tableaux or fillings of Ferrers diagrams. Is this involution already part of the vast canon of permutation automorphisms? de Mier's original involution for graphs [4] applies in our situation, and is apparently a different map. How does it compare?

Which subclasses of permutations preserve the symmetric distribution? From our example, we remark that cycle type is not neccesarily conserved (since loops are always mapped to upper transitory vertices), but non-intersecting intervals are preserved. Involution permutations are in bijection with partial matchings, and so this subclass has this property.

Is there an interpretation of crossing and nesting numbers in terms of other permutations statistics? Which other statistics does this involution preserve?

Ultimately we have considered a type of graph with two edge colours and strict degree restrictions. Can this be generalized to a larger class of graphs with fewer degree restrictions? What of a generalization of graphs with multiple edge colours?

Acknowledgements

The authors thank Cedric Chauve, Lily Yen, Sylvie Corteel, Anna de Mier and Eric Fusy for useful discussions. SB and MM are both funded by NSERC (Canada), and MM did part of this work as part of a CNRS Poste Rouge (France).

References

[1] M. Bousquet-Mélou and G. Xin. On partitions avoiding 3-crossings. *Sém. Lothar. Combin.*, 54:Art. B54e, 21 pp. (electronic), 2005/07.

[2] William Y. C. Chen, Eva Y. P. Deng, Rosena R. X. Du, Richard P. Stanley, and Catherine H. Yan. Crossings and nestings of matchings and partitions. *Trans. Amer. Math. Soc.*, 359(4):1555–1575 (electronic), 2007.

[3] Sylvie Corteel. Crossings and alignments of permutations. *Adv. in Appl. Math.*, 38(2):149–163, 2007.

[4] Anna de Mier. On the symmetry of the distribution of k-crossings and k-nestings in graphs. *Electron. J. Combin.*, 13(1):Note 21, 6 pp. (electronic), 2006.

[5] Anna de Mier. k-noncrossing and k-nonnesting graphs and fillings of Ferrers diagrams. *Combinatorica*, 27(6):699–720, 2007.

[6] C. Krattenthaler. Growth diagrams, and increasing and decreasing chains in fillings of Ferrers shapes. *Adv. in Appl. Math.*, 37(3):404–431, 2006.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 475–485

On joint distribution of adjacencies, descents and some Mahonian statistics

Alexander Burstein[1]

[1]*Department of Mathematics, Howard University, Washington, DC 20059 USA*

Abstract. We prove several conjectures of Eriksen regarding the joint distribution on permutations of the number of adjacencies (descents with consecutive values in consecutive positions), descents and some Mahonian statistics. We also prove Eriksen's conjecture that a certain bistatistic on Viennot's alternative tableaux is Euler-Mahonian.

Résumé. Nous demontrons plusieurs conjectures d'Eriksen concernant la distribution conjointe sur les permutations du nombre de contiguïtés (descentes avec des valeurs consécutives en positions consécutives), les descentes et quelques statistiques mahoniennes. Nous demontrons également une conjecture d'Eriksen qui affirme qu'une certaine bistatistique sur les tableaux alternatifs de Viennot est euler-mahonienne.

Keywords: permutation statistic, Eulerian, Mahonian, descent, adjacency, pattern, permutation tableau

1 Introduction

Eriksen [3] defined a new statistic `adj` on permutations that has the same distribution as the number of fixed points. He also conjectured that certain Euler-Mahonian pairs of statistics together with `adj` have the same joint distribution on permutations. Here we prove this conjecture. This refines a result of Foata and Zeilberger [5] that proves a conjecture of Babson and Steingrímsson [1].

1.1 Permutation statistics

We will start with some definitions. A combinatorial *statistic* on a set S is a map $\mathbf{f} : S \to \mathbb{N}^m$ for some integer $m \geq 0$. The *distribution* of $\mathbf{f}$ is the map $\mathrm{d}_{\mathbf{f}} : \mathbb{N}^m \to \mathbb{N}$ with $\mathrm{d}_{\mathbf{f}}(\mathbf{i}) = |\mathbf{f}^{-1}(\mathbf{i})|$ for $\mathbf{i} \in \mathbb{N}^m$, where $|\mathbf{f}^{-1}(\mathbf{i})|$ is the number of objects $s \in S$ such that $\mathbf{f}(s) = \mathbf{i}$.

Let $\mathfrak{S}_n$ be the set of permutations of $[n] = \{1, \dots, n\}$. A *descent* of a permutation $\pi \in \mathfrak{S}_n$ is a position $i < n$ such that $\pi(i) > \pi(i+1)$. Then $\pi(i)$ and $\pi(i+1)$ are called *descent top* and *descent bottom*, respectively. A non-descent position is called an *ascent*. Ascent tops and ascent bottoms are defined similarly. An *adjacency* is a descent such that $\pi(i) - \pi(i+1) = 1$. Where the context is unambiguous, we will also refer to the sequence of descent top and descent bottom as the descent, and do likewise for adjacencies.

Let $\mathtt{des}\,\pi$ be the number of descents of π, and let $\mathtt{adj}\,\pi$ be the number of adjacencies of $\pi 0$, i.e. a permutation of $[0, n] = \{0, 1, \dots, n\}$ obtained by appending 0 to the end of π. Eriksen proved that `adj` has the same distribution as `fix`, the number of *fixed points*, i.e. positions i such that $\pi(i) = i$.

A statistic is *Eulerian* if its distribution is the same as that of the descent statistic, $\mathtt{des}$. One well known Eulerian statistic is $\mathtt{exc}$, the number of *excedances*, i.e. $\mathtt{exc}\,\pi$ is the number of positions i of π such that $\pi(i) > i$. Another "almost" Eulerian statistic $\mathtt{wex}$, the number of *weak excedances* (i.e. $\mathtt{wex}\,\pi$ is the number of positions i of π such that $\pi(i) \geq i$) has the same distribution as $\mathtt{des} + 1$. Eriksen [3] has also proved that the bistatistic $(\mathtt{adj}, \mathtt{des} + 1)$ has the same distribution on permutations as $(\mathtt{fix}, \mathtt{wex})$.

An *inversion* of π is a pair of positions (i, j), such that $i < j$ and $\pi(i) > \pi(j)$. Let $\mathtt{inv}\,\pi$ be the number of inversions of π. A statistic is *Mahonian* if it has the same distribution as $\mathtt{inv}$. The first Mahonian statistic other than $\mathtt{inv}$ was discovered by MacMahon himself and is the *major index* $\mathtt{maj}$, the sum of positions of descents of a permutation. Other Mahonian statistics include, e.g., Denert's statistic $\mathtt{den}$ (see [1, 2]).

1.2 Permutation patterns

A *pattern* is an order-isomorphism type of a string over a totally ordered alphabet. An *occurrence* (or *instance*) of a pattern τ in a permutation π is a subsequence of π that is order-isomorphic to τ. A *generalized pattern*, first defined in [1], is a pattern where some consecutive entries must be adjacent in all occurrences of the pattern as well. Consecutive entries of the pattern that need not be adjacent in the containing permutation are separated by a hyphen.

Example 1.1 An occurrence of the generalized pattern 2-31 in a permutation π is a subsequence $(\pi(i), \pi(j), \pi(j + 1))$ of π such that $i < j$ and $\pi(j + 1) < \pi(i) < \pi(j)$.

Given a pattern τ and a permutation π, we denote by $(\tau)\pi$ the number of occurrences of the pattern τ in π. Thus, (τ) is a pattern occurrence statistic. Babson and Steingrímsson [1] showed that many Mahonian statistics can be expressed as sums of pattern occurrence statistics. For example,

$$\begin{aligned} \mathtt{inv} &= (2\text{-}1) = (21) + (3\text{-}12) + (3\text{-}21) + (2\text{-}31) = (21) + (31\text{-}2) + (32\text{-}1) + (23\text{-}1), \\ \mathtt{maj} &= (21) + (1\text{-}32) + (2\text{-}31) + (3\text{-}21), \\ \mathtt{mak} &= (21) + (1\text{-}32) + (2\text{-}31) + (32\text{-}1), \\ \mathtt{stat} &= (21) + (13\text{-}2) + (21\text{-}3) + (32\text{-}1) \end{aligned} \tag{1.1}$$

are Mahonian statistics. In fact, the last line is the definition [1] of the $\mathtt{stat}$ statistic.

We also need to define some permutation symmetries. *Reversal*, r, and *complement*, c, are the operations of reading a permutation back-to-front and upside-down, respectively. In other words, for $\pi \in \mathfrak{S}_n$, $\pi^{\mathrm{r}}(i) = \pi(n + 1 - i)$ and $\pi^{\mathrm{c}}(i) = n + 1 - \pi(i)$. Note that the composition $\mathrm{rc} = \mathrm{cr}$ is equivalent to rotating the permutation diagram by $180°$. It is well known (and easily seen) that for any permutation π and pattern τ, we have $(\tau^{\mathrm{r}})\pi^{\mathrm{r}} = (\tau^{\mathrm{c}})\pi^{\mathrm{c}} = (\tau^{\mathrm{rc}})\pi^{\mathrm{rc}} = (\tau)\pi$.

2 Main result

Eriksen [3] conjectured that $(\mathtt{adj}, \mathtt{des}, \mathtt{stat})$ and $(\mathtt{adj}, \mathtt{des}, \mathtt{maj})$ are equidistributed on permutations. This (up to permutation reversal) is a refinement of a result of Foata and Zeilberger [5, Theorem 3] who use q-enumeration and generating functions and an almost completely automated proof via Maple packages ROTA and PERCY. In this paper, we produce a "handmade" bijective proof of Eriksen's conjecture. A possible redeeming feature of this approach is that the bijection is a nice and simple involution yielding a slightly better result than may have been expected.

Let $\mathtt{F}\,\pi = \pi(1)$ be the first (leftmost) letter of π. Then the following result holds.

Theorem 2.1 *Statistics* $(\mathtt{adj}, \mathtt{des}, \mathtt{F}, \mathtt{maj}, \mathtt{stat})$ *and* $(\mathtt{adj}, \mathtt{des}, \mathtt{F}, \mathtt{stat}, \mathtt{maj})$ *have the same joint distribution on* $\mathfrak{S}_n$ *for all* n.

Eriksen's conjecture is an immediate corollary of Theorem 2.1.

Corollary 2.2 *Statistics* $(\mathtt{adj}, \mathtt{des}, \mathtt{stat})$ *and* $(\mathtt{adj}, \mathtt{des}, \mathtt{maj})$ *have the same joint distribution on* $\mathfrak{S}_n$ *for all* n.

To prove Theorem 2.1, we will define a map on permutations and prove that it preserves the values of $\mathtt{adj}$, $\mathtt{des}$ and $\mathtt{F}$ and switches the values of $\mathtt{maj}$ and $\mathtt{stat}$.

Given a permutation $\pi \in \mathfrak{S}_n$ with $\pi(1) = k$, define the permutation $\pi' \in \mathfrak{S}_n$ as follows: $\pi'(1) = \pi(1) = k$, and for $i \in [2, n]$,

$$\pi'(i) = \begin{cases} k - \pi(n+2-i), & \text{if } \pi(n+2-i) < k, \\ n+k+1-\pi(n+2-i), & \text{if } \pi(n+2-i) > k. \end{cases}$$

This is better visualized as follows. Let π_{b} (resp. π_{t}) be the subsequence of π consisting of values that are lesser (resp. greater) than k. Then we can represent π and π' graphically as

$$\pi = k\frac{\pi_{\mathrm{t}}}{\pi_{\mathrm{b}}} \quad \text{and} \quad \pi' = k\left(\frac{\pi_{\mathrm{t}}^{\mathrm{c}}}{\pi_{\mathrm{b}}^{\mathrm{c}}}\right)^{\mathrm{r}}.$$

In other words, we take the complements of π_{t} and π_{b} separately, but then take the reversal of the whole π except the first letter. Note that this implies that

$$\pi'_{\mathrm{b}} = \pi_{\mathrm{b}}^{\mathrm{rc}}, \quad \pi'_{\mathrm{t}} = \pi_{\mathrm{t}}^{\mathrm{rc}},$$

and the operation rc takes descents to descents and adjacencies to adjacencies. Note also that the map $p : \pi \mapsto \pi'$ (where p stands for "prime") is an involution on the set of permutations in $\mathfrak{S}_n$ that start with k, i.e. $(\pi')' = \pi$. Hence, p is a bijection on $\mathfrak{S}_n$.

We can also describe the bijection p is as follows. Insert bars between adjacent elements of π_{t} and π_{b} that are not adjacent in π. Also insert a bar at the start (resp. end) of the one of two sequences π_{t} and π_{b} not containing $\pi(2)$ (resp. $\pi(n)$). Let $\bar{\pi}_{\mathrm{t}}$ and $\bar{\pi}_{\mathrm{b}}$ be the resulting top and bottom sequences with bars, and write

$$\pi = k\frac{\bar{\pi}_{\mathrm{t}}}{\bar{\pi}_{\mathrm{b}}}.$$

Then

$$\pi' = k\frac{(\bar{\pi}_{\mathrm{t}})^{\mathrm{rc}}}{(\bar{\pi}_{\mathrm{b}})^{\mathrm{rc}}}.$$

In other words, we rotate (separately) the permutation diagrams of $\bar{\pi}_{\mathrm{t}}$ and $\bar{\pi}_{\mathrm{b}}$ by $180°$ *with bars*.

Example 2.3 Let $\pi = 543617982 = 5\dfrac{|6|798|}{43|1|2}$. Then $\pi' = 5\dfrac{|768|9|}{3|4|21} = 537684921$.

We claim that the following is true:

$$\begin{aligned} \texttt{adj}\,\pi' &= \texttt{adj}\,\pi \\ \texttt{des}\,\pi' &= \texttt{des}\,\pi \\ \texttt{F}\,\pi' &= \texttt{F}\,\pi \\ \texttt{maj}\,\pi' &= \texttt{stat}\,\pi \\ \texttt{stat}\,\pi' &= \texttt{maj}\,\pi \end{aligned} \tag{2.1}$$

Obviously, the first letter is preserved under this map, so the third equation is certainly true. To prove the fourth equation, we will show that

$$\texttt{maj}\,\pi + \texttt{stat}\,\pi = (n+1)\,\texttt{des}\,\pi - (k-1) = \texttt{maj}\,\pi + \texttt{maj}\,\pi' \tag{2.2}$$

for all $\pi \in \mathfrak{S}_n$. Finally, the fifth equation follows from the fourth equation and the fact that the map $\mathrm{p} : \pi \mapsto \pi'$ is an involution.

Example 2.4 Take $\pi = 543617982$ and $\pi' = 537684921$ from Example 2.3. Then $n = 9$, $k = \texttt{F}\,\pi = \texttt{F}\,\pi' = 5$, $\texttt{adj}\,\pi = \texttt{adj}\,\pi' = 3$ (recall that we actually count adjacencies in $\pi 0 = 5436179820$ and $\pi' 0 = 5376849210$), $\texttt{des}\,\pi = \texttt{des}\,\pi' = 5$, $\texttt{maj}\,\pi = \texttt{stat}\,\pi' = 22$, $\texttt{stat}\,\pi = \texttt{maj}\,\pi' = 24$, and $(n+1)\texttt{des}\,\pi - (k-1) = 46 = 22 + 24$.

Lemma 2.5 $\texttt{adj}\,\pi' = \texttt{adj}\,\pi$.

Proof: No descent of $\pi 0$ that starts in π_{t} and ends in $\pi_{\mathrm{b}}0$ can be an adjacency since the values of the descent top and descent bottom in this case differ by at least 2. Therefore, adjacencies of $\pi 0$ can be of four types:

- both top and bottom are in π_{t},
- both top and bottom are in π_{b},
- $k(k-1)$, if $\pi(2) = \pi_{\mathrm{b}}(1) = k-1$,
- 10, if $\pi(n) = \pi_{\mathrm{b}}(|\pi_{\mathrm{b}}|) = 1$.

If both entries are in π_{t} or in π_{b}, then these adjacencies are mapped to adjacencies in $\pi_{\mathrm{t}}^{\mathrm{rc}} = \pi'_{\mathrm{t}}$ or $\pi_{\mathrm{b}}^{\mathrm{rc}} = \pi'_{\mathrm{b}}$, respectively. If $\pi(2) = k-1$, then $\pi'(n) = 1$, so $\pi 0$ contains the adjacency $k(k-1)$ if and only if $\pi' 0$ contains the adjacency 10. Likewise, if $\pi(n) = 1$, then $\pi'(2) = k-1$, so $\pi 0$ contains the adjacency 10 if and only if $\pi' 0$ contains the adjacency $k(k-1)$. Thus, all adjacencies of $\pi 0$ map to adjacencies in $\pi' 0$, and vice versa, so $\texttt{adj}\,\pi = \texttt{adj}\,\pi'$. □

Lemma 2.6 $\texttt{des}\,\pi' = \texttt{des}\,\pi$.

Proof: As in the previous lemma, note that descents where both top and bottom are in π_{t} or in π_{b} map to descents in π' where both entries are in $\pi_{\mathrm{t}}^{\mathrm{rc}} = \pi'_{\mathrm{t}}$ and $\pi_{\mathrm{b}}^{\mathrm{rc}} = \pi'_{\mathrm{b}}$, respectively. Thus, we only need to look at descents of π and π' that start at or above k (i.e. in $k\pi_{\mathrm{t}}$ or $k\pi'_{\mathrm{t}}$, respectively) and end below k (i.e. in π_{b} or π'_{b}, respectively). The number of such descents equals the number of blocks in the bottom part π_{b} of π, which is the same as the number of blocks in the bottom part π'_{b} of π'.

This implies that $\mathtt{des}\,\pi' = \mathtt{des}\,\pi$ as desired. □

For the next lemma, we will need a bit of notation. Given a pattern τ, let $\tau_{i_1,i_2,\ldots,i_h}$ denote the pattern τ with distinguished entries $i_1 < i_2 < \cdots < i_h$. Usually, we will write τ with distinguished entries italicized. Now given a list of letters $j_1 < j_2 < \cdots < j_h$ and a permutation π, let $(\tau_{i_1,i_2,\ldots,i_h})(j_1, j_2, \ldots, j_h)$ be the permutation statistic of the number of occurrences of τ where each entry i_s in τ corresponds to the entry j_s in the containing permutation, for $s = 1, 2, \ldots, h$. For example, (2-*31*)$(a, b)\pi$ counts all occurrences of 2-31 in π where a and b in π correspond to 1 and 3 in 2-31, respectively.

Lemma 2.7 $\mathtt{maj}\,\pi + \mathtt{stat}\,\pi = (n+1)\,\mathtt{des}\,\pi - (k-1)$.

Proof: From Equation (1.1), it follows that

$$\mathtt{maj} + \mathtt{stat} = (21) + (21) + (3\text{-}21) + (21\text{-}3) + (1\text{-}32) + (32\text{-}1) + (2\text{-}31) + (13\text{-}2).$$

Let ba, $b > a$, be a descent of π (in other words, b is the descent top and a is the descent bottom here). Then

$$(\mathit{21})(a, b) + (\mathit{21})(a, b)$$

counts the entries a and b themselves,

$$(3\text{-}\mathit{21})(a, b) + (\mathit{21}\text{-}3)(a, b)$$

counts all entries greater than b (split into those to the left or to the right of the descent ba),

$$(1\text{-}\mathit{32})(a, b) + (\mathit{32}\text{-}1)(a, b)$$

counts all entries less than a (split into the same two groups), and

$$(2\text{-}\mathit{31})(a, b) + (\mathit{31}\text{-}2)(a, b)$$

counts all entries between a and b (split likewise). This implies that

$$\Big((\mathit{21}) + (\mathit{21}) + (3\text{-}\mathit{21}) + (\mathit{21}\text{-}3) + (1\text{-}\mathit{32}) + (\mathit{32}\text{-}1) + (2\text{-}\mathit{31}) + (\mathit{31}\text{-}2)\Big)(a, b)\pi = |\pi| = n$$

for any permutation π of size n and any descent ba of π. Summing over all descents of π, we get

$$\Big((21) + (21) + (3\text{-}21) + (21\text{-}3) + (1\text{-}32) + (32\text{-}1) + (2\text{-}31)\Big)\pi = n\,\mathtt{des}\,\pi - (31\text{-}2)\pi,$$

so that

$$\mathtt{maj}\,\pi + \mathtt{stat}\,\pi = n\,\mathtt{des}\,\pi + (13\text{-}2)\pi - (31\text{-}2)\pi.$$

Now for each $c \in [n]$, lets us count $(13\text{-}2)(c)\pi - (31\text{-}2)(c)\pi$. Recall that $\pi(1) = k$. We claim that the following is true:

$$(13\text{-}2)(c)\pi - (31\text{-}2)(c)\pi = \begin{cases} 0, & \text{if } c = k, \\ 0, & \text{if } c > k \text{ is an ascent top}, \\ 1, & \text{if } c > k \text{ is an descent bottom}, \\ -1, & \text{if } c < k \text{ is an ascent top}, \\ 0, & \text{if } c < k \text{ is an descent bottom.} \end{cases}$$

Let us call instances of $(13\text{-}2)(c)$ and $(31\text{-}2)(c)$ left ascents *across* c and left descents *across* c, respectively. Since the first line is obvious, assume $c \neq k$ and let d be the entry immediately preceding c in π.

If $c > k$ is an ascent top, then $c > d$ as well, so k and d are on the same side of c (on the value axis of the permutation diagram), and hence, from k to d, the number of ascents across c is equal to the number of descents across c, so their difference is 0. The same result is obtained when $c < k$ is a descent bottom, i.e. when $c < d$ as well, and hence k and d are again on the same side of c. If $c > k$ is a descent bottom, then $k < c < d$, so k is below c and d is above c. Hence, from k to d, there are 1 more ascents across c than descents across c. Finally, if $c < k$ is an ascent top, then $k > c > d$, so k is above c and d is below c. Hence, from k to d, there are 1 more descents across c than ascents across c.

Therefore, summing over all entries $c \in \pi$, we get

$$\begin{aligned} (13\text{-}2)\pi - (31\text{-}2)\pi &= |\{\text{descent bottoms} > k\}| - |\{\text{ascent tops} < k\}| = \\ &= (|\{\text{descent bottoms} > k\}| + |\{\text{descent bottoms} < k\}|) \\ &\quad - (\{|\text{ascent tops} < k\}| + |\{\text{descent bottoms} < k\}|) = \\ &= |\{\text{all descent bottoms}\}| - |\{\text{all entries} < k\}| \\ &= \texttt{des}\,\pi - |\pi_{\mathrm{b}}| = \texttt{des}\,\pi - (k-1). \end{aligned} \tag{2.3}$$

We will explain the passage from the second to the third equality in some detail. For the first parenthesis, note that k cannot be a descent bottom and that every descent bottom corresponds to a unique descent. For the second parenthesis, note that every entry less than k must be at the end of an ascent or descent.

Thus,

$$\texttt{maj}\,\pi + \texttt{stat}\,\pi = n\,\texttt{des}\,\pi + \texttt{des}\,\pi - (k-1) = (n+1)\,\texttt{des}\,\pi - (k-1).$$

This ends the proof. □

Remark 2.8 The Equation (2.3) can also be proved as follows. Note that $\texttt{des} = (21)$ and $k - 1 = \texttt{F}\,\pi - 1 = [2\text{-}1)$, where the initial bracket means that the first letter of the pattern must also be the first letter in the permutation. In our case, $[2\text{-}1)$ counts all inversions starting from the leftmost letter of π, i.e. all letters less than $\pi(1) = k$. We can write

$$(2\text{-}1) = [2\text{-}1) + (32\text{-}1) + (23\text{-}1) + (13\text{-}2), \tag{2.4}$$

since the first letter in an inversion is either the initial letter in the permutation (the first summand on the right) or preceded by another letter (the other three summands). On the other hand,

$$(2\text{-}1) = (21) + (32\text{-}1) + (23\text{-}1) + (31\text{-}2), \tag{2.5}$$

since the first letter in an inversion is immediately followed either by the second letter in that inversion (the first summand on the right) or by some other letter (the other three summands). Comparing Equations (2.4) and (2.5), we obtain

$$(13\text{-}2) - (31\text{-}2) = (21) - [2\text{-}1) = \mathtt{des} - (\mathtt{F} - 1). \tag{2.6}$$

Lemma 2.9 $\mathtt{maj}\,\pi + \mathtt{maj}\,\pi' = (n+1)\mathtt{des}\,\pi - (k-1)$.

Proof: Suppose that i is a descent of π and that $\pi(i)$ and $\pi(i+1)$ are both in π_{t} or both in π_{b}. Then the map $p : \pi \mapsto \pi'$ takes this descent to the descent at position $n+1-i$ since $\pi'(n+1-i) > \pi'(n+2-i)$ and these values are also both in π_{t} or both in π_{b}. Therefore, each descent of π within π_{t} or within π_{b} and its corresponding descent in π' within π'_{t} or π'_{b} contributes

$$i + (n+1-i) = n+1$$

to the sum $\mathtt{maj}\,\pi + \mathtt{maj}\,\pi'$.

Now consider descents in π and π' from not below k to below k. Suppose that $\pi_{\mathrm{b}} = \pi_{\mathrm{b}}^{(1)}\pi_{\mathrm{b}}^{(2)}\ldots\pi_{\mathrm{b}}^{(m)}$, where each $\pi_{\mathrm{b}}^{(s)}$ is a maximal block of consecutive entries of π that are in π_{b}. Suppose also that $\pi_{\mathrm{b}}^{(s)}$ starts at position j_s+1 for some $j_s \le n-1$. Then $\pi_{\mathrm{b}}^{(s)}$ ends at position $j_s + |\pi_{\mathrm{b}}^{(s)}|$. Therefore, we can partition the descents of π and π' from not below k to below k into pairs, where the descent from $k\pi_{\mathrm{t}}$ to $\pi_{\mathrm{b}}^{(s)}$ at position j_s corresponds to the descent from $k\pi'_{\mathrm{t}}$ to $(\pi_{\mathrm{b}}^{(s)})^{\mathrm{rc}}$ at position

$$(n+2) - (j_s + |\pi_{\mathrm{b}}^{(s)}| + 1) = (n+1) - (j_s + |\pi_{\mathrm{b}}^{(s)}|).$$

Therefore, each such pair (for $s = 1, \ldots, m$) together contributes

$$j_s + \Big((n+1) - (j_s + |\pi_{\mathrm{b}}^{(s)}|)\Big) = (n+1) - |\pi_{\mathrm{b}}^{(s)}|$$

to the sum $\mathtt{maj}\,\pi + \mathtt{maj}\,\pi'$. Summing over all $\pi_{\mathrm{b}}^{(s)}$ for $s = 1, \ldots, m$, we get that the descents in π and π' from not below k to below k together contribute

$$(n+1)|\{\text{descents from } k\pi_{\mathrm{t}} \text{ to } \pi_{\mathrm{b}}\}| - \sum_{s=1}^{m} |\pi_{\mathrm{b}}^{(s)}| = (n+1)|\{\text{descents from } k\pi_{\mathrm{t}} \text{ to } \pi_{\mathrm{b}}\}| - |\pi_{\mathrm{b}}|$$

to $\mathtt{maj}\,\pi + \mathtt{maj}\,\pi'$. Thus, all descents in π and π' together sum to

$$\begin{aligned}&(n+1)|\{\text{descents of } \pi \text{ in } \pi_{\mathrm{t}}\}| + (n+1)|\{\text{descents of } \pi \text{ in } \pi_{\mathrm{b}}\}| + (n+1)|\{\text{descents from } k\pi_{\mathrm{t}} \text{ to } \pi_{\mathrm{b}}\}| - |\pi_{\mathrm{b}}|\\ &= (n+1)|\{\text{all descents of } \pi\}| - |\pi_{\mathrm{b}}| = (n+1)\,\mathtt{des}\,\pi - (k-1),\end{aligned}$$

in other words,

$$\mathtt{maj}\,\pi + \mathtt{maj}\,\pi' = (n+1)\,\mathtt{des}\,\pi - (k-1).$$

This ends the proof. □

Thus, we proved all the equalities in Equation (2.1). This ends the proof of Theorem 2.1.

3 Distribution of $(\mathtt{adj}, \mathtt{des}+1)$

In [3], Eriksen gives a proof of equidistribution of $(\mathtt{adj}, \mathtt{des}+1)$ and $(\mathtt{fix}, \mathtt{wex})$ on $\mathfrak{S}_n$ using two bijections from permutations to permutation tableaux. Here we give a direct bijection on permutations that maps the former bistatistic to the latter. This bijection is different [4] from the composition of Eriksen's two bijections. In fact, letting $\mathfrak{S}_n^0 = \{\pi 0 \,|\, \pi \in \mathfrak{S}_n\}$, we get $\mathtt{des}\,\pi 0 = \mathtt{des}\,\pi + 1$, $\mathtt{fix}\,\pi 0 = \mathtt{fix}\,\pi$, $\mathtt{wex}\,\pi 0 = \mathtt{wex}\,\pi$, so $(\mathtt{adj}, \mathtt{des})$ and $(\mathtt{fix}, \mathtt{wex})$ are equidistributed on $\mathfrak{S}_n^0$.

Given a permutation $\pi \in \mathfrak{S}_n$, for each entry $m \in [n]$ of π, define $\ell(m)$ to be the leftmost entry of $\pi 0$ to the right of m that is less than m. Note that $\ell(m)+1 \le m$ for all $m \in [n]$. Then π is mapped to

$$\tilde{\pi} = (1\ \ \ell(1)+1)(2\ \ \ell(2)+1)\ldots(n-1\ \ \ell(n-1)+1)(n\ \ \ell(n)+1). \tag{3.1}$$

The map $t : \pi \mapsto \tilde{\pi}$ is obviously a bijection since $\tilde{\pi}$ in (3.1) is the product of transpositions required to make entries $n, n-1, \ldots, 2, 1$ of $\tilde{\pi}$ fixed points in order of decreasing magnitude (i.e., first, n is moved from position $\ell(n)+1$ to position n, then $n-1$ is moved from position $\ell(n-1)+1$ to position $n-1$, and so on). Moreover, the inverse map $t^{-1} : \tilde{\pi} \mapsto \pi$ amounts to starting with the string 0, then inserting the entries $1, 2, \ldots, n$ in increasing order so that each i is inserted immediately to the left of $\ell(i)$, then deleting the 0.

Example 3.1 Let $\pi = 543617982$, then $\pi 0 = 5436179820$, so

$$\tilde{\pi} = (11)(21)(32)(44)(55)(62)(73)(83)(99) = 268453179.$$

Note that $\mathtt{adj}\,\pi = 3 = \mathtt{fix}\,\tilde{\pi}$ and $\mathtt{des}\,\pi + 1 = \mathtt{des}(\pi 0) = 6 = \mathtt{wex}\,\tilde{\pi}$.

Proposition 3.2 *The map $t : \pi \mapsto \tilde{\pi}$ is a bijection on S_n such that* $(\mathtt{adj}, \mathtt{des}+1)\pi = (\mathtt{fix}, \mathtt{wex})\tilde{\pi}$.

Proof: Suppose that, scanning the cycles of $\tilde{\pi}$ in formula (3.1) from right to left, we see that the element j occurs first as $\ell(i)+1$ for some i. Then $j \le i$ and i does not occur to the left of the cycle $(i\, j) = (i\ \ell(i)+1)$ in (3.1). Therefore, $\tilde{\pi}(j) = i \ge j$, i.e. j is a weak excedance of $\tilde{\pi}$. But such a situation arises exactly when i is to the left of $j-1 = \ell(i)$, and between i and $j-1$ there is no element less than $j-1$ (by definition of $\ell(i)$) and no element greater than $j-1$ (or j would occur in a transposition earlier to the right in (3.1)), i.e. exactly when $\pi 0$ contains a descent from i to $j-1$.

On the other hand, suppose the above situation does not occur. Then j first occurs as the greater entry of the transposition $(j\ \ \ell(j)+1)$ in (3.1), and $\ell(j)+1 < j$. Then j is first mapped to $\ell(j)+1 < j$ by $(j\ \ \ell(j)+1)$, and since there are no elements greater than j to the left of $(j\ \ \ell(j)+1)$ in (3.1), it follows that $\tilde{\pi}(j) < j$.

Thus, j is a weak excedance of $\tilde{\pi}$ exactly when $j-1$ is a descent bottom of $\pi 0$. Moreover, j is fixed point of $\tilde{\pi}$ exactly when $\pi 0$ contains a descent from j to $j-1$, i.e. when j is an adjacency of $\pi 0$. $\square$

4 Euler-Mahonian statistics on permutation tableaux

Here we will give a simple proof of another conjecture of Eriksen [3]: that a certain bistatistic on permutation tableaux is Euler-Mahonian, i.e. has the same distribution as $(\mathtt{des}, \mathtt{maj})$.

A *Ferrers diagram* of a partition is a left-justified column of nonincreasing rows of identical squares (cells), where some rows may be of length 0. A *permutation tableau* $\mathcal{T}$ (see [7]) is a $(0,1)$-filling of a Ferrers diagram that satisfies the following properties:

(1-hinge) every cell that has a 1 to its left in the same row and a 1 above it in the same column must also contain a 1 (such a 1 is called an *induced* 1),

(column) every column contains at least one 1.

Permutation tableaux of semiperimeter n are in bijection with permutations of length n (see [7]). Note that there is no corresponding restrictions on rows, i.e. a row may contain all 0s.

A 0 in a permutation tableau is called a *restricted zero* if there is a 1 above it in the same column. Note that all cells to the left of a restricted 0 in a permutation tableau must also be filled with 0s.

Viennot [8] considered so-called *alternative tableaux*, that are closely related to the regular permutation tableaux. An alternative tableau $\bar{\mathcal{T}}$ is obtained from an regular permutation tableau $\mathcal{T}$ by replacing the top 1 in each column with a blue dot, the rightmost restricted 0 (if any) in each row with a red dot, then removing all nontop 1s from their cells and deleting the top row except for its bottom boundary (so that we know the length of the deleted top row).

This operation is a bijection. Indeed, given an alternative tableau $\mathcal{A}$, we can recover the permutation tableau $\hat{\mathcal{A}} = \mathcal{T}$ such that $\bar{\mathcal{T}} = \mathcal{A}$ by adjoining back the top row, inserting blue dots in those columns of the top row which had no blue dots in $\bar{\mathcal{T}}$, filling all red dot cells, cells to the left of red dots and cells above the blue dots with 0s, and filling the remaining cells with 1s.

The alternative tableaux have an advantage over the original permutation tableaux in that they are closed under the involution that consists of transposition and switching the colors of red and blue dots.

Eriksen [3] conjectured that the following statistic on alternative tableaux,

$$\texttt{Astat} = \binom{\text{rows}+1}{2} + \text{ red dots } + \text{ blue dots } + \text{ cells to the left of red dots } + \text{ cells above blue dots},$$

is Mahonian and, in fact, that (rows, $\texttt{Astat}$) is Euler-Mahonian. Here we prove this conjecture.

Theorem 4.1 *The bistatistic* (rows, $\texttt{Astat}$) *on alternative tableaux is Euler-Mahonian.*

Proof: Let $\mathcal{A}$ be an alternative tableau, and let $\mathcal{T} = \hat{\mathcal{A}}$ be its corresponding permutation tableau. Then the 0s of $\mathcal{T}$ that are not in the top row are exactly in the cells that either contain the red dots or are to the left of the red dots or are above the blue dots in $\mathcal{A}$. On the other hand, the 0s in the top row of $\mathcal{T}$ are exactly in the columns that contain the blue dots of $\mathcal{T}$. Note also that $\text{rows}(\mathcal{T}) = \text{rows}(\mathcal{A}) + 1$. Thus,

$$\texttt{Astat}(\mathcal{A}) = \binom{\text{rows}(\mathcal{T})}{2} + \text{ zeros}(\mathcal{T}).$$

Steingrímsson and Williams [7] give a bijection from permutation tableaux to permutations that converts tableau statistics to pattern statistics. If a permutation tableau $\mathcal{T}$ corresponds to a permutation π via that bijection ($\Psi^{-1} \circ \Phi$ in the notation of [7]), then the following holds

$$\begin{aligned} \texttt{des}\,\pi &= \text{rows}(\mathcal{T}) - 1 = \text{ rows}(\mathcal{A}), \\ [(31\text{-}2) + (21\text{-}3) + (3\text{-}21)]\pi - \binom{\texttt{des}\,\pi}{2} &= \text{ zeros}(\mathcal{T}). \end{aligned}$$

Thus,

$$\begin{aligned}\texttt{Astat}(\mathcal{A}) &= \binom{\texttt{des}\,\pi+1}{2} + [(31\text{-}2)+(21\text{-}3)+(3\text{-}21)]\pi - \binom{\texttt{des}\,\pi}{2}\\ &= \texttt{des}\,\pi + [(31\text{-}2)+(21\text{-}3)+(3\text{-}21)]\pi\\ &= [(21)+(31\text{-}2)+(21\text{-}3)+(3\text{-}21)]\pi\\ &= [(21)+(2\text{-}31)+(1\text{-}32)+(32\text{-}1)]\pi^{\mathrm{rc}}\\ &= \texttt{mak}\,\pi^{\mathrm{rc}} = \texttt{mak}^{\mathrm{rc}}(\pi),\end{aligned}$$

and mak is a Mahonian statistic (e.g., see [2]). Moreover, $\texttt{des}\,\pi^{\mathrm{rc}} = \texttt{des}\,\pi$, and $(\texttt{des}, \texttt{mak})$ is Euler-Mahonian [6], and thus, $(\mathrm{rows}, \texttt{Astat})$ is Euler-Mahonian as well. □

5 Further questions

In the earlier sections we have proved that the bistatistics $(\texttt{des}, \texttt{mak})$ (or, more specifically, $(\texttt{des}, \texttt{mak})^{\mathrm{rc}}$) on permutations and $(\mathrm{rows}, \texttt{Astat})$ on permutation tableaux are equidistributed. Steingrímsson and Williams [7] show that the bistatistics $(\texttt{fix}, \texttt{wex})$ on permutations and $(\text{rows-with-no-1s}, \mathrm{rows})$ on permutation tableaux are equidistributed. Furthermore, Eriksen [3] gives a direct bijection that shows equidistribution of $(\texttt{adj}, \texttt{des}+1)$ and $(\text{rows-with-no-1s}, \mathrm{rows})$.

However, the triples of statistics $(\texttt{adj}, \texttt{des}, \texttt{mak})$ (or $(\texttt{adj}, \texttt{des}+1, \texttt{mak})$, or $(\texttt{adj}, \texttt{des}, \texttt{mak})^{\mathrm{rc}}$) and $(\text{rows-with-no-1s}, \mathrm{rows}, \texttt{Astat})$ are not equidistributed. This leads us to ask if there are more or less natural statistics on permutations or permutation tableaux that fill the position of the question mark, for example, in the following equidistributions, among others:

$$\begin{aligned}(\texttt{adj}, \texttt{des}, \texttt{mak}) &\sim (\text{rows-with-no-1s}, \mathrm{rows}, ?)\\ (\texttt{adj}, \texttt{des}, \texttt{mak}) &\sim (?, \mathrm{rows}, \texttt{Astat})\\ (\texttt{adj}, \texttt{des}, ?) &\sim (\text{rows-with-no-1s}, \mathrm{rows}, \texttt{Astat})\\ (?, \texttt{des}, \texttt{mak}) &\sim (\text{rows-with-no-1s}, \mathrm{rows}, \texttt{Astat})\end{aligned}$$

References

[1] E. Babson, E. Steingrímsson, Generalized permutation patterns and a classification of the Mahonian statistics, *Sém. Lothar. Combin.* **B44b** (2000), 18 pp.

[2] R.J. Clarke, E. Steingrímsson, J. Zeng, New Euler-Mahonian statistics on permutations and words, *Adv. Appl. Math.* **18** (1997), 237–270.

[3] N. Eriksen, Pattern and position based permutation statistics, in preparation, available online at `http://www.math.chalmers.se/~ner/artiklar/bpMahonianPP.pdf`.

[4] N. Eriksen, personal communication.

[5] D. Foata, D. Zeilberger, Babson-Steingrímsson statistics are indeed Mahonian (and sometimes even Euler-Mahonian), *Adv. Appl. Math.* **27** (2001), 390–404.

[6] D. Foata, D. Zeilberger, Denert's permutation statistic is indeed Euler-Mahonian, *Studies in Appl. Math.* **83** (1990), 31–59.

[7] E. Steingrímsson, L.K. Williams, Permutation tableaux and permutation patterns, *J. Combin. Th. Ser. A* **114** (2007), no. 2, 211–234.

[8] X. Viennot, Alternative tableaux, permutations and partially asymmetric exclusion process. Talk given at the Isaac Newton Institute, Cambridge, available online at `http://www.newton.ac.uk/programmes/CSM/seminars/042314001.html`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 486–497

Mixed Statistics on 01-Fillings of Moon Polyominoes

William Y. C. Chen[1] and Andrew Y. Z. Wang[1] and Catherine H. Yan[2] and Alina F. Y. Zhao[1†]

[1]*Center for Combinatorics, LPMC-TJKLC*
Nankai University, Tianjin 300071, P. R. China
[2] *Department of Mathematics*
Texas A&M University, College Station, TX 77843, USA

Abstract. We establish a stronger symmetry between the numbers of northeast and southeast chains in the context of 01-fillings of moon polyominoes. Let $\mathcal{M}$ be a moon polyomino. Consider all the 01-fillings of $\mathcal{M}$ in which every row has at most one 1. We introduce four mixed statistics with respect to a bipartition of rows or columns of $\mathcal{M}$. More precisely, let S be a subset of rows of $\mathcal{M}$. For any filling M, the top-mixed (resp. bottom-mixed) statistic $\alpha(S;M)$ (resp. $\beta(S;M)$) is the sum of the number of northeast chains whose top (resp. bottom) cell is in S, together with the number of southeast chains whose top (resp. bottom) cell is in the complement of S. Similarly, we define the left-mixed and right-mixed statistics $\gamma(T;M)$ and $\delta(T;M)$, where T is a subset of the columns. Let $\lambda(A;M)$ be any of these four statistics $\alpha(S;M)$, $\beta(S;M)$, $\gamma(T;M)$ and $\delta(T;M)$. We show that the joint distribution of the pair $(\lambda(A;M), \lambda(M/A;M))$ is symmetric and independent of the subsets S, T. In particular, the pair of statistics $(\lambda(A;M), \lambda(M/A;M))$ is equidistributed with $(\mathrm{se}(M), \mathrm{ne}(M))$, where $\mathrm{se}(M)$ and $\mathrm{ne}(M)$ are the numbers of southeast chains and northeast chains of M, respectively.

Résumé. Nous établissons une symétrie plus forte entre les nombres de chaînes nord-est et sud-est dans le cadre des remplissages 01 des polyominos lune. Soit $\mathcal{M}$ un polyomino lune. Considérez tous les remplissages 01 de $\mathcal{M}$ dans lesquels chaque rangée contient au plus un 1. Nous présentons quatre statistiques mixtes sur les bipartitions des rangées et des colonnes de $\mathcal{M}$. Plus précisément, soit S un sous-ensemble de rangées de $\mathcal{M}$. Pour tout remplissage M, la statistique mixte du dessus (resp. du dessous) $\alpha(S;M)$ (resp. $\beta(S;M)$) est la somme du nombre de chaînes nord-est dont le dessus (resp. le dessous) est dans S, et du nombre de chaînes sud-est dont la cellule supérieure (resp. inférieure) est dans le complément de S. De même, nous définissons les statistiques mixtes à gauche et à droite $\gamma(T;M)$ et $\delta(T;M)$, où T est un sous-ensemble des colonnes. Soit $\lambda(A;M)$ une des quatre statistiques $\alpha(S;M)$, $\beta(S;M)$, $\gamma(T;M)$ et $\delta(T;M)$. Nous montrons que la distribution commune des paires $(\lambda(A;M), \lambda(M/A;M))$ est symétrique et indépendante des sous-ensembles S, T. En particulier, la paire de statistiques $(\lambda(A;M), \lambda(M/A;M))$ est équidistribuée avec $(\mathrm{se}(M), \mathrm{ne}(M))$, où $\mathrm{se}(M)$ et $\mathrm{ne}(M)$ sont les nombres de chaînes sud-est et nord-est de M respectivement.

Keywords: mixed statistic, polyomino, symmetric distribution.

[†]The first, the second, and the fourth authors were supported by the 973 Project, the PCSIRT Project of the Ministry of Education, and the National Science Foundation of China. The third author was supported in part by a US-NSF grant #DMS-0653846. The authors thank Peter McNamara and Alain Goupil for the french translation of the abstract.

1 Introduction

Recently it is observed that the numbers of crossings and nestings have a symmetric distribution over many families of combinatorial objects, such as matchings and set partitions. Recall that a matching of $[2n] = \{1, 2, \ldots, 2n\}$ is a partition of the set $[2n]$ with the property that each block has exactly two elements. It can be represented as a graph with vertices $1, 2, \ldots, 2n$ drawn on a horizontal line in increasing order, where two vertices i and j are connected by an edge if and only if $\{i, j\}$ is a block. We say that two edges (i_1, j_1) and (i_2, j_2) form a *crossing* if $i_1 < i_2 < j_1 < j_2$; they form a *nesting* if $i_1 < i_2 < j_2 < j_1$. The symmetry of the joint distribution of crossings and nestings follows from the bijections of de Sainte-Catherine, who also found the generating functions for the number of crossings and the number of nestings. Klazar [12] further studied the distribution of crossings and nestings over the set of matchings obtained from a given matching by successfully adding edges.

The symmetry between crossings and nestings was extended by Kasraoui and Zeng [11] to set partitions, and by Chen, Wu and Yan [3] to linked set partitions. Poznanović and Yan [15] determined the distribution of crossings and nestings over the set of partitions which are identical to a given partition π when restricted to the last n elements.

Many classical results on enumerative combinatorics can be put in the larger context of counting submatrices in fillings of certain polyominoes. For example, words and permutations can be represented as 01-fillings of rectangular boards, and general graphs can be represented as $\mathbb{N}$-fillings of arbitrary Ferrers shapes, as studied by [13, 6, 7]. Other polyominoes studied include stack polyominoes [9], and moon polyominoes [16, 10]. It is well-known that crossings and nestings in matchings and set partitions correspond to northeast chains and southeast chains of length 2 in a filling of polyominoes. The symmetry between crossings and nestings has been extended by Kasraoui [10] to 01-fillings of moon polyominoes where either every row has at most one 1, or every column has at most one 1. In both cases, the joint distribution of the numbers of northeast and southeast chains can be expressed as a product of p, q-Gaussian coefficients. Other known statistics on fillings of moon polyominoes are the length of the longest northeast/southeast chains [2, 13, 16], and the major index [4].

The main objective of this paper is to present a stronger symmetry between the numbers of northeast and southeast chains in the context of 01-fillings of moon polyominoes. Given a bipartition of the rows (or columns) of a moon polyomino, we define four statistics by considering mixed sets of northeast and southeast chains according to the bipartition. Let M be a 01-filling of a moon polyomino $\mathcal{M}$ with n rows and m columns. These statistics are the top-mixed and the bottom-mixed statistics $\alpha(S; M), \beta(S; M)$ with respect to a row-bipartition $(S, \bar{S})$, and the left-mixed and the right-mixed statistics $\gamma(T; M), \delta(T; M)$ with respect to a column-bipartition $(T, \bar{T})$. We show that for any of these four statistics $\lambda(A; M)$, namely, $\alpha(S; M), \beta(S; M)$ for $S \subseteq [n]$ and $\gamma(T; M), \delta(T; M)$ for $T \subseteq [m]$, the joint distribution of the pair $(\lambda(A; M), \lambda(\bar{A}; M))$ is symmetric and independent of the subsets S, T. Consequently, we have the equidistribution

$$\sum_{M} p^{\lambda(A;M)} q^{\lambda(\bar{A};M)} = \sum_{M} p^{\mathrm{se}(M)} q^{\mathrm{ne}(M)},$$

where M ranges over all 01-fillings of $\mathcal{M}$ with the property that either every row has at most one 1, or every column has at most one 1, and $\mathrm{se}(M)$ and $\mathrm{ne}(M)$ are the numbers of southeast and northeast chains of M, respectively.

The paper is organized as follows. Section 2 contains necessary notation and the statements of the main results. We present the proofs in Section 3, and show by bijections in Section 4 that these new statistics are invariant under a permutation of columns or rows on moon polyominoes.

2 Notation and the Main Results

A *polyomino* is a finite subset of $\mathbb{Z}^2$, where every element of $\mathbb{Z}^2$ is represented by a square cell. The polyomino is *convex* if its intersection with any column or row is connected. It is *intersection-free* if every two columns are comparable, i.e., the row-coordinates of one column form a subset of those of the other column. Equivalently, it is intersection-free if every two rows are comparable. A *moon polyomino* is a convex and intersection-free polyomino.

Given a moon polyomino $\mathcal{M}$, we assign 0 or 1 to each cell of $\mathcal{M}$ so that there is at most one 1 in each row. Throughout this paper we will simply use the term *filling* to denote such 01-fillings. We say that a cell is empty if it is assigned 0, and it is a 1-cell otherwise. Assume $\mathcal{M}$ has n rows and m columns. We label the rows $R_1, \ldots, R_n$ from top to bottom, and the columns $C_1, \ldots, C_m$ from left to right. Let $\mathbf{e} = (\varepsilon_1, \ldots, \varepsilon_n) \in \{0,1\}^n$ and $\mathbf{s} = (s_1, \ldots, s_m) \in \mathbb{N}^m$ with $\sum_{i=1}^n \varepsilon_i = \sum_{j=1}^m s_j$. We denote by $\mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ the set of fillings M of $\mathcal{M}$ such that the row R_i has exactly ε_i many 1's, and the column C_j has exactly s_j many 1's, for $1 \le i \le n$ and $1 \le j \le m$. See Figure 1 for an illustration.

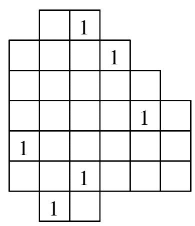


Fig. 1: A filling M with $\mathbf{e} = (1,1,0,1,1,1,1)$ and $\mathbf{s} = (1,1,2,1,1,0)$.

A *northeast* (resp. *southeast*) *chain* in a filling M of $\mathcal{M}$ is a set of two 1-cells such that one of them is strictly above (resp. below) and to the right of the other and the smallest rectangle containing them is contained in $\mathcal{M}$. Northeast (resp. southeast) chains will be called NE (resp. SE) chains. The number of NE (resp. SE) chains of M is denoted by $\mathrm{ne}(M)$ (resp. $\mathrm{se}(M)$). It is proved by Kasraoui [10] that $\mathrm{ne}(M)$ and $\mathrm{se}(M)$ have a symmetric joint distribution over $\mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$.

Theorem 2.1

$$\sum_{M \in \mathbf{F}(\mathcal{M},\mathbf{e},\mathbf{s})} p^{\mathrm{ne}(M)} q^{\mathrm{se}(M)} = \sum_{M \in \mathbf{F}(\mathcal{M},\mathbf{e},\mathbf{s})} p^{\mathrm{se}(M)} q^{\mathrm{ne}(M)} = \prod_{i=1}^m \begin{bmatrix} h_i \\ s_i \end{bmatrix}_{p,q}.$$

Let $\mathcal{R}$ be the set of rows of the moon polyomino $\mathcal{M}$. For $S \subseteq [n]$, let $\mathcal{R}(S) = \bigcup_{i \in S} R_i$. We say a 1-cell is an *S-cell* if it lies in $\mathcal{R}(S)$. An NE chain is called a *top S-NE chain* if its northeast 1-cell is an S-cell. Similarly, an SE chain is called a *top S-SE chain* if its northwest 1-cell is an S-cell. In other words, an NE/SE chain is a top S-NE/SE chain if the upper 1-cell of the chain is in $\mathcal{R}(S)$. Similarly, an NE/SE chain is a *bottom S-NE/SE chain* if the lower 1-cell of the chain is in $\mathcal{R}(S)$.

Let $\bar{S} = [n] \setminus S$ be the complement of S. Given a filling $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$, we define the *top-mixed statistic* $\alpha(S; M)$ and the *bottom-mixed statistic* $\beta(S; M)$ *with respect to* S as

$$\begin{aligned}
\alpha(S; M) &= \#\{\text{top } S\text{-NE chain of } M\} + \#\{\text{top } \bar{S}\text{-SE chain of } M\}, \\
\beta(S; M) &= \#\{\text{bottom } S\text{-NE chain of } M\} + \#\{\text{bottom } \bar{S}\text{-SE chain of } M\}.
\end{aligned}$$

See Example 2.3 for some of these statistics on the filling M in Figure 1.

Let $F_S^t(p, q)$ and $F_S^b(p, q)$ be the bi-variate generating functions for the pairs $(\alpha(S; M), \alpha(\bar{S}; M))$ and $(\beta(S; M), \beta(\bar{S}; M))$ respectively, namely,

$$F_S^t(p, q) = \sum_{M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})} p^{\alpha(S;M)} q^{\alpha(\bar{S};M)} \quad \text{and} \quad F_S^b(p, q) = \sum_{M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})} p^{\beta(S;M)} q^{\beta(\bar{S};M)}.$$

Note that

$$(\alpha(\emptyset; M), \alpha([n]; M)) = (\beta(\emptyset; M), \beta([n]; M)) = (\mathrm{se}(M), \mathrm{ne}(M)).$$

Our first result is the following property.

Theorem 2.2 $F_S^t(p, q) = F_{S'}^t(p, q)$ *for any two subsets* S, S' *of* $[n]$. *In other words, the bi-variate generating function* $F_S^t(p, q)$ *does not depend on* S. *Consequently,*

$$F_S^t(p, q) = F_\emptyset^t(p, q) = \sum_{M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})} p^{\mathrm{se}(M)} q^{\mathrm{ne}(M)}$$

is a symmetric function. The same statement holds for $F_S^b(p, q)$.

We can also define the mixed statistics with respect to a subset of columns. Let $\mathcal{C}$ be the set of columns of $\mathcal{M}$. For $T \subseteq [m]$, let $\mathcal{C}(T) = \bigcup_{j \in T} C_j$. An NE chain is called a *left* T*-NE chain* if the southwest 1-cell of the chain lies in $\mathcal{C}(T)$. Similarly, an SE chain is called a *left* T*-SE chain* if the northwest 1-cell of the chain lies in $\mathcal{C}(T)$. In other words, an NE/SE chain is a left T-NE/SE chain if its left 1-cell is in $\mathcal{C}(T)$. Similarly, an NE/SE chain is a *right* T*-NE/SE chain* if its right 1-cell is in $\mathcal{C}(T)$.

Let $\bar{T} = [m] \setminus T$ be the complement of T. For any filling M of $\mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$, we define the *left-mixed statistic* $\gamma(T; M)$ and the *right-mixed statistic* $\delta(T; M)$ *with respect to* T as

$$\begin{aligned}
\gamma(T; M) &= \#\{\text{left } T\text{-NE chain of } M\} + \#\{\text{left } \bar{T}\text{-SE chain of } M\}, \\
\delta(T; M) &= \#\{\text{right } T\text{-NE chain of } M\} + \#\{\text{right } \bar{T}\text{-SE chain of } M\}.
\end{aligned}$$

Example 2.3 Let M be the filling in Figure 1, where $\mathrm{ne}(M) = 6$ and $\mathrm{se}(M) = 1$. Let $S = \{2, 4\}$, i.e., $\mathcal{R}(S)$ contains the second and the fourth rows. Then

$$\alpha(S; M) = 5, \quad \alpha(\bar{S}; M) = 2, \quad \beta(S; M) = 1, \quad \beta(\bar{S}; M) = 6.$$

Let $T = \{1, 3\}$, i.e., $\mathcal{C}(T)$ contains the first and the third columns. Then

$$\gamma(T; M) = 4, \quad \gamma(\bar{T}; M) = 3, \quad \delta(T; M) = 2, \quad \delta(\bar{T}; M) = 5. \qquad \blacksquare$$

Let $G^l_T(p,q)$ and $G^r_T(p,q)$ be the bi-variate generating functions of the pairs $(\gamma(T;M),\gamma(\bar{T};M))$ and $(\delta(T;M),\delta(\bar{T};M))$ respectively, namely,

$$G^l_T(p,q) = \sum_{M\in\mathbf{F}(\mathcal{M},\mathbf{e},\mathbf{s})} p^{\gamma(T;M)}q^{\gamma(\bar{T};M)} \quad \text{and} \quad G^r_T(p,q) = \sum_{M\in\mathbf{F}(\mathcal{M},\mathbf{e},\mathbf{s})} p^{\delta(T;M)}q^{\delta(\bar{T};M)}.$$

Again note that

$$(\gamma(\emptyset;M),\gamma([m];M)) = (\delta(\emptyset;M),\delta([m];M)) = (\mathrm{se}(M),\mathrm{ne}(M)).$$

Our second result shows that the generating function $G^l_T(p,q)$ possesses a similar property as $F^t_S(p,q)$.

Theorem 2.4 *$G^l_T(p,q) = G^l_{T'}(p,q)$ for any two subsets T,T' of $[m]$. In other words, the bi-variate generating function $G^l_T(p,q)$ does not depend on T. Consequently,*

$$G^l_T(p,q) = G^l_\emptyset(p,q) = \sum_{M\in\mathbf{F}(\mathcal{M},\mathbf{e},\mathbf{s})} p^{\mathrm{se}(M)}q^{\mathrm{ne}(M)}$$

is a symmetric function. The same statement holds for $G^r_T(p,q)$.

We notice that the set $\mathbf{F}(\mathcal{M},\mathbf{e},\mathbf{s})$ appeared as $\mathcal{N}^r(T,\mathbf{m},A)$ in Kasraoui [10], where $\mathbf{m}$ is the column sum vector, and A is the set of empty rows, i.e., $A=\{i:\varepsilon_i=0\}$. Kasraoui also considered the set $\mathcal{N}^c(T,\mathbf{n},B)$ of fillings whose row sum is an arbitrary $\mathbb{N}$-vector $\mathbf{n}$ under the condition that there is at most one 1 in each column and where B is the set of empty columns. By a rotation of moon polyominoes, it is easily seen that Theorem 2.2 and Theorem 2.4 also hold for the set $\mathcal{N}^c(T,\mathbf{n},B)$, as well as for the set of fillings such that there is at most one 1 in each row and in each column.

As an interesting example, we explain how Theorems 2.2 and 2.4 specialize to permutations and words, which are in bijections with fillings of squares or rectangles. More precisely, a word $w=w_1w_2\cdots w_n$ on $[m]$ can be represented as a filling M on an $n\times m$ rectangle $\mathcal{M}$ in which the cell in row $n+1-i$ and column j is assigned the integer 1 if and only if $w_i=j$. In the word $w_1w_2\cdots w_n$, a pair (w_i,w_j) is an *inversion* if $i<j$ and $w_i>w_j$; we say that it is a *co-inversion* if $i<j$ and $w_i<w_j$, see also [14]. Denote by $\mathrm{inv}(w)$ the number of inversions of w, and by $\mathrm{coinv}(w)$ the number of co-inversions of w.

For $S\subseteq[n]$ and $T\subseteq[m]$, we have

$$\begin{aligned}
\alpha(S;w) &= \#\{(w_i,w_j): n+1-j\in S \text{ and } (w_i,w_j) \text{ is a co-inversion}\}\\
&\qquad +\#\{(w_i,w_j): n+1-j\notin S \text{ and } (w_i,w_j) \text{ is an inversion}\},\\
\beta(S;w) &= \#\{(w_i,w_j): n+1-i\in S \text{ and } (w_i,w_j) \text{ is a co-inversion}\}\\
&\qquad +\#\{(w_i,w_j): n+1-i\notin S \text{ and } (w_i,w_j) \text{ is an inversion}\}.\\
\gamma(T,w) &= \#\{(w_i,w_j): w_i\in T \text{ and } (w_i,w_j) \text{ is a co-inversion}\}\\
&\qquad +\#\{(w_i,w_j): w_j\notin T \text{ and } (w_i,w_j) \text{ is an inversion}\},\\
\delta(T,w) &= \#\{(w_i,w_j): w_j\in T \text{ and } (w_i,w_j) \text{ is a co-inversion}\}\\
&\qquad +\#\{(w_i,w_j): w_i\notin T \text{ and } (w_i,w_j) \text{ is an inversion}\}.
\end{aligned}$$

Let $W=\{1^{s_1},2^{s_2},\ldots,m^{s_m}\}$ be a multiset with $s_1+\cdots+s_m=n$, and $R(W)$ be the set of permutations, also called rearrangements, of the elements in W. Let $\lambda(A;w)$ denote any of the four statistics

$\alpha(S;w), \beta(S;w), \gamma(T;w), \delta(T;w)$. Theorems 2.2 and 2.4 imply that the bi-variate generating function for $(\lambda(A;w), \lambda(\bar{A};w))$ is symmetric and

$$\sum_{w\in R(W)} p^{\lambda(A;w)} q^{\lambda(\bar{A};w)} = \sum_{w\in R(W)} p^{\mathrm{inv}(w)} q^{\mathrm{coinv}(w)} = \begin{bmatrix} n \\ s_1, \ldots, s_m \end{bmatrix}_{p,q}, \tag{1}$$

where $\left[\begin{smallmatrix} n \\ s_1,\ldots,s_m \end{smallmatrix}\right]_{p,q}$ is the p,q-Gaussian coefficient with the p,q-integer $[r]_{p,q}$ given by $[r]_{p,q} = p^{r-1} + p^{r-2}q + \cdots + pq^{r-2} + q^{r-1}$.

3 Proof of the Main Results

It is sufficient to prove our results for $\alpha(S;M)$ and $\gamma(T;M)$ only, since conclusions for $\beta(S;M)$ and $\delta(T;M)$ can be obtained by reflecting the moon polyomino with respect to a horizontal line or a vertical line.

In Subsection 3.1, we recall Kasraoui's bijection Ψ from $\mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ to sequences of compositions [10]. Kasraoui's construction is stated for the set $\mathcal{N}^c(T, \mathbf{n}, B)$. We shall modify the description to fit our notation. This bijection will be used in the proof of Lemma 3.2 which states that the pair of the top-mixed statistics $(\alpha(\{1\};M), \alpha(\overline{\{1\}};M))$ is equidistributed with $(\mathrm{se}(M), \mathrm{ne}(M))$. Theorem 2.2 follows from an iteration of Lemma 3.2. In Subsection 3.3 we prove Theorem 2.4. Again the crucial step is the observation that $(\gamma(\{1\};M), \gamma(\overline{\{1\}};M))$ has the same distribution as $(\mathrm{se}(M), \mathrm{ne}(M))$.

Due to the space limit, in this extended abstract we would just describe the main ideas and the construction of the bijections, and leave out the detailed proofs. A complete version of the present paper is available in [5].

3.1 Kasraoui's bijection Ψ

Assume the columns of $\mathcal{M}$ are $C_1, \ldots, C_m$ from left to right. Let $|C_i|$ be the length of the column C_i. Assume that k is the smallest index such that $|C_k| \geq |C_i|$ for all i. Define the left part of $\mathcal{M}$, denoted $L(\mathcal{M})$, to be the union $\cup_{1\leq i\leq k-1} C_i$, and the right part of $\mathcal{M}$, denoted $R(\mathcal{M})$, to be the union $\cup_{k\leq i\leq m} C_i$. Note that the columns of maximal length in $\mathcal{M}$ belong to $R(\mathcal{M})$.

We order the columns $C_1, \ldots, C_m$ by a total order $\prec$ as follows: $C_i \prec C_j$ if and only if

- $|C_i| < |C_j|$ or
- $|C_i| = |C_j|$, $C_i \in L(\mathcal{M})$ and $C_j \in R(\mathcal{M})$, or
- $|C_i| = |C_j|$, $C_i, C_j \in L(\mathcal{M})$ and C_i is on the left of C_j, or
- $|C_i| = |C_j|$, $C_i, C_j \in R(\mathcal{M})$ and C_i is on the right of C_j.

For every column $C_i \in L(\mathcal{M})$, we define the rectangle $\mathcal{M}(C_i)$ to be the largest rectangle that contains C_i as the leftmost column. For $C_i \in R(\mathcal{M})$, the rectangle $\mathcal{M}(C_i)$ is taken to be the largest rectangle that contains C_i as the rightmost column and does not contain any column $C_j \in L(\mathcal{M})$ such that $C_j \prec C_i$.

Given $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$, we define a coloring of M by the following steps.

The coloring of the filling M

1. Color the cells of empty rows;

2. For each $C_i \in L(\mathcal{M})$, color the cells which are contained in the rectangle $\mathcal{M}(C_i)$ and on the right of any 1-cell in C_i.
3. For each $C_i \in R(\mathcal{M})$, color the cells which are contained in the rectangle $\mathcal{M}(C_i)$ and on the left of any 1-cell in C_i.

For $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$, let a_i be the number of empty rows (i.e., $\{R_i : \varepsilon_i = 0\}$) that intersect the column C_i. Suppose that $C_{i_1} \prec C_{i_2} \prec \cdots \prec C_{i_m}$. For $j = 1, \ldots, m$, we define

$$h_{i_j} = |C_{i_j}| - a_{i_j} - (s_{i_1} + s_{i_2} + \cdots + s_{i_{j-1}}). \tag{2}$$

For positive integers n and k, denote by $\mathcal{C}_k(n)$ the set of compositions of n into k nonnegative parts, that is, $\mathcal{C}_k(n) = \{(\lambda_1, \lambda_2, \ldots, \lambda_k) \in \mathbb{N}^k : \sum_{i=1}^k \lambda_i = n\}$. The bijection Ψ is constructed as follows.

The bijection $\Psi : \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) \longrightarrow \mathcal{C}_{s_1+1}(h_1 - s_1) \times \mathcal{C}_{s_2+1}(h_2 - s_2) \times \cdots \times \mathcal{C}_{s_m+1}(h_m - s_m)$.

For each $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ with the coloring, $\Psi(M)$ is a sequence of compositions $(c^{(1)}, c^{(2)}, \ldots, c^{(m)})$, where

- $c^{(i)} = (0)$ if $s_i = 0$. Otherwise
- $c^{(i)} = (c_1^{(i)}, c_2^{(i)}, \ldots, c_{s_i+1}^{(i)})$ where
  - $c_1^{(i)}$ is the number of uncolored cells above the first 1-cell in the column C_i;
  - $c_k^{(i)}$ is the number of uncolored cells between the $(k-1)$-st and the k-th 1-cells in the column C_i, for $2 \le k \le s_i$;
  - $c_{s_i+1}^{(i)}$ is the number of uncolored cells below the last 1-cell in the column C_i.

The statistics $\mathrm{ne}(M)$ and $\mathrm{se}(M)$ can be written in terms of the compositions.

Theorem 3.1 *Let* $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ *and* $\mathbf{c} = \Psi(M) = (c^{(1)}, c^{(2)}, \ldots, c^{(m)})$. *Then*

$$\mathrm{ne}(M) = \sum_{C_i \in L(\mathcal{M})} \sum_{k=1}^{s_i} (c_1^{(i)} + c_2^{(i)} + \cdots + c_k^{(i)}) + \sum_{C_j \in R(\mathcal{M})} \sum_{k=1}^{s_j} (h_j - s_j - c_1^{(j)} - c_2^{(j)} - \cdots - c_k^{(j)}),$$

$$\mathrm{se}(M) = \sum_{C_i \in L(\mathcal{M})} \sum_{k=1}^{s_i} (h_i - s_i - c_1^{(i)} - c_2^{(i)} - \cdots - c_k^{(i)}) + \sum_{C_j \in R(\mathcal{M})} \sum_{k=1}^{s_j} (c_1^{(j)} + c_2^{(j)} + \cdots + c_k^{(j)}).$$

Summing over the sequences of compositions yields the symmetric generating function of $\mathrm{ne}(M)$ and $\mathrm{se}(M)$, c.f. Theorem 2.1.

3.2 Proof of Theorem 2.2

To prove Theorem 2.2 for the top-mixed statistic $\alpha(S; M)$, we first consider the special case when $\mathcal{R}(S)$ contains the first row only.

Lemma 3.2 *For* $S = \{1\}$, *we have*

$$F_{\{1\}}^t(p, q) = F_{\emptyset}^t(p, q) = \sum_{M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})} p^{\mathrm{se}(M)} q^{\mathrm{ne}(M)}.$$

Proof: We assume that the first row is nonempty. Otherwise the identity is obvious. Given a filling $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$, assume that the unique 1-cell of the first row lies in the column C_t. Let the upper polyomino $\mathcal{M}_u$ be the union of the rows that intersect C_t, and the lower polyomino $\mathcal{M}_d$ be the complement of $\mathcal{M}_u$, i.e., $\mathcal{M}_d = \mathcal{M} \setminus \mathcal{M}_u$. We aim to construct a bijection $\phi_\alpha : \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) \to \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ such that for any filling M,

$$(\alpha(\{1\}; M), \alpha(\overline{\{1\}}; M)) = (\mathrm{se}(\phi_\alpha(M)), \mathrm{ne}(\phi_\alpha(M))),$$

and $\phi_\alpha(M)$ is identical to M on $\mathcal{M}_d$ (which depends on M).

Let $M_u = M \cap \mathcal{M}_u$ and $M_d = M \cap \mathcal{M}_d$. Let s'_i be the number of 1-cells of M in the column $C_i \cap \mathcal{M}_u$, and $\mathbf{s}' = (s'_1, \ldots, s'_m)$. Let $\mathbf{e}' = (\varepsilon_1, \ldots, \varepsilon_r)$, where r is the number of rows in $\mathcal{M}_u$. We shall define ϕ_α on $\mathbf{F}(\mathcal{M}_u, \mathbf{e}', \mathbf{s}')$ first.

Let $C'_i = C_i \cap \mathcal{M}_u$. Suppose that in $\mathcal{M}$ the columns intersecting with the first row are $C_a, \ldots, C_t, \ldots, C_b$ from left to right. Then $C_t = C'_t$, and in $\mathcal{M}_u$ the columns $C'_a, \ldots, C'_t, \ldots, C'_b$ intersect the first row. Assume that among them the ones with the same length as C'_t are $C'_u, \ldots, C'_t, \ldots, C'_v$ from left to right. Clearly, the columns $C'_u, \ldots, C'_t, \ldots, C'_v$ are those with maximal length and belong to $R(\mathcal{M}_u)$. Note that in M_u, the number of top $\{1\}$-NE chains is $\sum_{a \le i < t} s'_i$, while the number of top $\{1\}$-SE chains is $\sum_{t < i \le b} s'_i$. Let h'_i be given as in Eq. (2) for $\mathbf{F}(\mathcal{M}_u, \mathbf{e}', \mathbf{s}')$. Let $\mathbf{c} = \Psi(M_u) = (c^{(1)}, c^{(2)}, \ldots, c^{(m)})$. Then we can compute that

$$\begin{aligned} \alpha(\{1\}; M_u) \;\; = \;\; & \sum_{a \le i < u} s'_i + (h'_t - s'_t) + \sum_{C'_i \in L(\mathcal{M}_u)} \sum_{k=1}^{s'_i} (h'_i - s'_i - c_1^{(i)} - c_2^{(i)} - \cdots - c_k^{(i)}) \\ & + \sum_{C'_j \in R(\mathcal{M}_u)} \sum_{k=1}^{s'_j} (c_1^{(j)} + c_2^{(j)} + \cdots + c_k^{(j)}) - \sum_{t < i \le b} s'_i. \end{aligned} \tag{3}$$

and

$$\begin{aligned} \alpha(\overline{\{1\}}; M_u) \;\; = \;\; & \sum_{t < i \le b} s'_i + \sum_{C'_i \in L(\mathcal{M}_u)} \sum_{k=1}^{s'_i} (c_1^{(i)} + c_2^{(i)} + \cdots + c_k^{(i)}) \\ & + \sum_{C'_j \in R(\mathcal{M}_u)} \sum_{k=1}^{s'_j} (h'_j - s'_j - c_1^{(j)} - c_2^{(j)} - \cdots - c_k^{(j)}) - \sum_{a \le i < u} s'_i - (h'_t - s'_t). \end{aligned} \tag{4}$$

The fact that the 1-cell of the first row lies in the column C'_t implies that $c_1^{(t)} = 0$, and $c_1^{(i)} > 0$ for $a \le i < u$ or $t < i \le b$. We define the filling $\phi_\alpha(M_u)$ by setting $\phi_\alpha(M_u) = \Psi^{-1}(\tilde{\mathbf{c}})$, where $\tilde{\mathbf{c}}$ is obtained from $\mathbf{c}$ as follows:

$$\begin{cases} \tilde{c}^{(i)} = (c_1^{(i)} - 1, c_2^{(i)}, \ldots, c_{s_i}^{(i)}, c_{s_i+1}^{(i)} + 1), & \text{if } a \le i < u \text{ or } t < i \le b, \text{ and } s'_i \ne 0, \\ \tilde{c}^{(t)} = (c_2^{(t)}, c_3^{(t)}, \ldots, c_{s_t+1}^{(t)}, c_1^{(t)}), & \text{if } i = t, \\ \tilde{c}^{(i)} = c^{(i)}, & \text{for any other } i. \end{cases}$$

Comparing the formulas (3) and (4) with Theorem 3.1 for $\tilde{\mathbf{c}}$, one easily verifies that

$$(\alpha(\{1\}; M_u), \alpha(\overline{\{1\}}; M_u)) = (\mathrm{se}(\phi_\alpha(M_u)), \mathrm{ne}(\phi_\alpha(M_u))).$$

Now $\phi_\alpha(M)$ is obtained from M by replacing M_u with $\phi_\alpha(M_u)$. □

Proposition 3.3 *Assume* $S = \{r_1, r_2, \ldots, r_s\} \subseteq [n]$ *with* $r_1 < r_2 < \cdots < r_s$. *Let* $S' = \{r_1, r_2, \ldots, r_{s-1}\}$. *Then* $F^t_S(p, q) = F^t_{S'}(p, q)$.

Proof: Let $X = \{R_i : 1 \leq i < r_s\}$ be the set of rows above the row R_{r_s}, and Y be the set of remaining rows. Given a filling $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$, let $\mathcal{T}(M)$ be the set of fillings $M' \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ that are identical to M in the rows of X. Construct a map $\theta_{r_s} : \mathcal{T}(M) \to \mathcal{T}(M)$ by setting $\theta_{r_s}(M)$ to be the filling obtained from M by replacing $M \cap Y$ with $\phi_\alpha(M \cap Y)$. Then it is a bijection with the property that

$$(\alpha(S; M), \alpha(\bar{S}; M)) = (\alpha(S'; \theta_{r_s}(M)), \alpha(\overline{S'}; \theta_{r_s}(M))). \tag{5}$$

□

Proof of Theorem 2.2. Assume $S = \{r_1, r_2, \ldots, r_s\} \subseteq \mathcal{R}$ with $r_1 < r_2 < \cdots < r_s$. Let $\Theta_\alpha = \theta_{r_1} \circ \theta_{r_2} \circ \cdots \circ \theta_{r_s}$, where θ_r is defined in the proof of Prop. 3.3. Then Θ_α is a bijection on $\mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ with the property that $(\alpha(S; M), \alpha(\bar{S}; M)) = (\mathrm{se}(\Theta_\alpha(M)), \mathrm{ne}(\Theta_\alpha(M)))$. The symmetry of $F^t_S(p, q)$ follows from Theorem 2.1. □

3.3 Proof of Theorem 2.4

Theorem 2.4 is concerned with the left-mixed statistic $\gamma(T; M)$. The idea of the proof is similar to that of Theorem 2.2: we show that the statement is true when T contains the left-most column only. However, Kasraoui's bijection ϕ does not help here, since the columns and rows play different roles in the fillings. Instead, we give an algorithm which gradually maps the left-mixed statistics with respect to the first column to the pair $(\mathrm{ne}, \mathrm{se})$.

Lemma 3.4 *For* $T = \{1\}$, *we have*

$$G^l_{\{1\}}(p, q) = G^l_{\emptyset}(p, q) = \prod_{i=1}^{m} \begin{bmatrix} h_i \\ s_i \end{bmatrix}_{p,q}.$$

The proof is built on an involution ρ on the fillings of a rectangular shape $\mathcal{M}$.

An involution ρ on rectangular shapes.

Let $\mathcal{M}$ be an $n \times m$ rectangle, and M a filling of $\mathcal{M}$. Let C_1 be the left-most column of M, in which the 1-cells are in the $l_1, \ldots, l_k$ rows from top to bottom. Replace C_1 by the column C^r_1 so that the 1's in C^r_1 appear in the $l_1, \ldots, l_k$ rows from bottom to top. This is $\rho(M)$. Note that this map does not change the relative positions of those 1-cells that are not in C_1. It is easy to verify that $\rho(\rho(M)) = M$ and $(\gamma(\{1\}; M), \gamma(\overline{\{1\}}; M)) = (\mathrm{se}(\rho(M)), \mathrm{ne}(\rho(M)))$.

Proof: Given a general moon polyomino $\mathcal{M}$, assume that the rows intersecting the first column are $\{R_a, \ldots, R_b\}$. Let $\mathcal{M}_c$ be the union $R_a \cup \cdots \cup R_b$. Clearly, for any $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$, a left $\{1\}$-NE (SE) chain consists of two 1-cells in $\mathcal{M}_c$. Let $C'_i = C_i \cap \mathcal{M}_c$ be the restriction of the column C_i on $\mathcal{M}_c$. Then $C'_1 = C_1$ and $|C'_1| \geq |C'_2| \geq \cdots \geq |C'_m|$.

Suppose that

$$|C'_1| = |C'_2| = \cdots = |C'_{j_1}| > |C'_{j_1+1}| = |C'_{j_1+2}| = \cdots = |C'_{j_2}| > |C'_{j_2+1}| \cdots$$
$$\cdots = |C'_{j_{k-1}}| > |C'_{j_{k-1}+1}| = |C'_{j_{k-1}+2}| = \cdots = |C'_{j_k}| = |C'_m|.$$

Let B_i be the greatest rectangle contained in $\mathcal{M}_c$ whose right most column is C'_{j_i} ($1 \leq i \leq k$), and $B'_i = B_i \cap B_{i+1}$ ($1 \leq i \leq k-1$).

We define $\phi_\gamma\colon \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) \to \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ by constructing a sequence of fillings $(M, M_k, \ldots, M_1)$ starting from M.

The map $\phi_\gamma\colon \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) \to \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$

Let $M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$.

1. The filling M_k is obtained from M by replacing $M \cap B_k$ with $\rho(M \cap B_k)$.
2. For i from $k-1$ to 1:
   (a) Define a filling N_i on B'_i by setting $N_i = \rho(M_{i+1} \cap B'_i)$. Let the filling M'_i be obtained from M_{i+1} by replacing $M_{i+1} \cap B'_i$ with N_i.
   (b) The filling M_i is obtained from M'_i by replacing $M'_i \cap B_i$ with $\rho(M'_i \cap B_i)$.
3. Set $\phi_\gamma(M) = M_1$.

Then ϕ_γ is a bijection satisfying the equation $(\gamma(\{1\}; M), \gamma(\overline{\{1\}}; M)) = (\mathrm{se}(\phi_\gamma(M)), \mathrm{ne}(\phi_\gamma(M)))$. $\square$

Proposition 3.5 *Assume* $T = \{c_1, c_2, \ldots, c_t\} \subseteq [m]$ *with* $c_1 < c_2 < \cdots < c_t$. *Let* $T' = \{c_1, c_2, \ldots, c_{t-1}\}$. *Then* $G^l_T(p,q) = G^l_{T'}(p,q)$.

The proof is similar to that of Prop. 3.3. Iterating Prop. 3.5 leads to Theorem 2.4.

4 Invariance Properties

The bi-variate generating function of $(\mathrm{ne}, \mathrm{se})$ (cf. Theorem 2.1) implies that the mixed statistics are invariant under any permutation of rows and/or columns. To be more specific, let $\mathcal{M}$ be a moon polyomino. For any moon polyomino $\mathcal{M}'$ obtained from $\mathcal{M}$ by permuting the rows and/or the columns of $\mathcal{M}$, we have

$$\begin{aligned}&\#\{M \in \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) : \lambda(A; M) = i, \lambda(\bar{A}; M) = j\}\\ &\quad = \#\{M' \in \mathbf{F}(\mathcal{M}', \mathbf{e}', \mathbf{s}') : \lambda(A; M') = i, \lambda(\bar{A}; M') = j\}\end{aligned}$$

for any nonnegative integers i and j, where $\mathbf{e}'$ (resp. $\mathbf{s}'$) is the sequence obtained from $\mathbf{e}$ (resp. $\mathbf{s}$) in the same ways as the rows (resp. columns) of $\mathcal{M}'$ are obtained from the rows (resp. columns) of $\mathcal{M}$, and $\lambda(A; M)$ is any of the four statistics $\alpha(S; M), \beta(S; M), \gamma(T; M)$, and $\delta(T; M)$. In this section we present bijective proofs of such phenomena.

Let $\mathcal{M}$ be a general moon polyomino. Let $\mathcal{N}_l$ be the unique left-aligned moon polyomino whose sequence of row lengths is equal to $|R_1|, \ldots, |R_n|$ from top to bottom. In other words, $\mathcal{N}_l$ is the left-aligned polyomino obtained by rearranging the columns of $\mathcal{M}$ by length in weakly decreasing order from left to right. We shall use an algorithm developed in [4] that rearranges the columns of $\mathcal{M}$ to generate $\mathcal{N}_l$.

The algorithm α **for rearranging** $\mathcal{M}$**:**

Step 1 Set $\mathcal{M}' = \mathcal{M}$.

Step 2 If $\mathcal{M}'$ is left aligned, go to Step 4.

Step 3 If $\mathcal{M}'$ is not left-aligned, consider the largest rectangle $\mathcal{B}$ completely contained in $\mathcal{M}'$ that contains C_1, the leftmost column of $\mathcal{M}'$. Update $\mathcal{M}'$ by setting $\mathcal{M}'$ to be the polyomino obtained by moving the leftmost column of $\mathcal{B}$ to the right end. Go to Step 2.

Step 4 Set $\mathcal{N}_l = \mathcal{M}'$.

Based on the algorithm α, Chen et al. constructed a bijection $g = g_{\mathcal{M}} : \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) \to \mathbf{F}(\mathcal{N}_l, \mathbf{e}, \mathbf{s}')$ such that $(\mathrm{se}(M), \mathrm{ne}(M)) = (\mathrm{se}(g(M)), \mathrm{ne}(g(M)))$, see [4, Section 5.3.2].

Combining $g_{\mathcal{M}}$ with the bijection Θ_α constructed in the proof of Theorem 2.2, we are led to the following invariance property.

Theorem 4.1 *Let $\mathcal{M}$ be a moon polyomino. For any moon polyomino $\mathcal{M}'$ obtained from $\mathcal{M}$ by permuting the columns of $\mathcal{M}$, the map*

$$\Phi_\alpha = \Theta_\alpha^{-1} \circ g_{\mathcal{M}'}^{-1} \circ g_{\mathcal{M}} \circ \Theta_\alpha : \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) \to \mathbf{F}(\mathcal{M}', \mathbf{e}, \mathbf{s}') \tag{6}$$

is a bijection with the property that

$$(\alpha(S; M), \alpha(\bar{S}; M)) = (\alpha(S; M'), \alpha(\bar{S}; M')).$$

Similarly, let $\mathcal{N}_t$ be the top aligned polyomino obtained from $\mathcal{M}$ by rotating 90 degrees counterclockwise first, followed by applying the algorithm α, and finally rotating 90 degrees clockwise. Such operations enable us to establish a bijection $h = h_{\mathcal{M}}$ from $\mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s})$ to $\mathbf{F}(\mathcal{N}_t, \mathbf{e}', \mathbf{s})$ that keeps the statistics $(\mathrm{se}, \mathrm{ne})$. The exact description of $h_{\mathcal{M}}$ is given in [5]. Combining the bijection Θ_α with $h_{\mathcal{M}}$, we arrive at the second invariance property.

Theorem 4.2 *Let $\mathcal{M}$ be a moon polyomino. For any moon polyomino $\mathcal{M}'$ obtained from $\mathcal{M}$ by permuting the rows of $\mathcal{M}$, the map*

$$\Lambda_\alpha = \Theta_\alpha^{-1} \circ h_{\mathcal{M}'}^{-1} \circ h_{\mathcal{M}} \circ \Theta_\alpha : \mathbf{F}(\mathcal{M}, \mathbf{e}, \mathbf{s}) \to \mathbf{F}(\mathcal{M}', \mathbf{e}', \mathbf{s}) \tag{7}$$

is a bijection with the property that

$$(\alpha(S; M), \alpha(\bar{S}; M)) = (\alpha(S; M'), \alpha(\bar{S}; M')).$$

Similar statements hold for the statistics $\beta(S; M)$, $\gamma(T; M)$ and $\delta(T; M)$.

References

[1] D. Chebikin, Variations on descents and inversions in permutations, Electron. J. Combin. 15 (2008) R132.

[2] W. Y. C. Chen, E. Y. P. Deng, R. R. X. Du, R. P. Stanley and C. H. Yan, Crossings and nestings of matchings and partitions, Trans. Amer. Math. Soc. 359 (2007) 1555–1575.

[3] W. Y. C. Chen, S. Y. Wu and C. H. Yan, Linked partitions and linked cycles, European J. Combin. 29 (2008) 1408–1426.

[4] W. Y. C. Chen, S. Poznanović, C. H. Yan and A. L. B. Yang, Major index for 01-fillings of moon polyominoes, arXiv:math.CO/0902.3637.

[5] W. Y. C. Chen, A.Y.Z. Wang, C. H. Yan and A. F.Y. Zhao, Mixed statistics on 01-fillings of moon polyominoes, (full version), preprint. arXiv:math.CO/0908.0073.

[6] A. de Mier, On the symmetry of the distribution of crossings and nestings in graphs, Electron. J. Combin. 13 (2006), N21.

[7] A. de Mier, k-noncrossing and k-nonnesting graphs and fillings of Ferrers diagrams, Combinatorica 27 (2007) 699–720.

[8] M. de Sainte-Catherine, Couplages et Pfaffiens en Combinatoire, Physique et Informatique, Ph.D. Thesis, University of Bordeaux I, Talence, France, 1983.

[9] J. Jonsson, Generalized triangulations and diagonal-free subsets of stack polyominoes, J. Combin. Theory Ser. A 112 (2005), 117–142.

[10] A. Kasraoui, Ascents and descents in 01-fillings of moon polyominoes, European J. Combin., to appear.

[11] A. Kasraoui and J. Zeng, Distributions of crossings, nestings and alignments of two edges in matchings and partitions, Electron. J. Combin. 13 (2006) R33.

[12] M. Klazar, On identities concerning the numbers of crossings and nestings of two edges in matchings, SIAM J. Discrete Math. 20 (2006) 960–976.

[13] C. Krattenthaler, Growth diagrams, and increasing and decreasing chains in fillings of Ferrers shapes, Adv. in Appl. Math. 37 (2006), 404–431.

[14] A. Mendes, J. Remmel, Descents, inversions, and major indices in permutation groups, Discrete Math. 308 (2008) 2509–2524.

[15] S. Poznanović, C. H. Yan, Crossings and nestings of two edges in set partitions, SIAM J. Discrete Math., to appear.

[16] M. Rubey, Increasing and decreasing sequences in fillings of moon polyominoes, arXiv:math.CO/0604140.

Pattern avoidance in partial permutations (extended abstract)

Anders Claesson[1†] and Vít Jelínek[1†] and Eva Jelínková[2‡] and Sergey Kitaev[1†]

[1]*The Mathematics Institute, School of Computer Science, Reykjavik University, Menntavegur 1, IS-101 Reykjavik, Iceland*
[2]*Department of Applied Mathematics, Charles University in Prague, Malostranské nám. 25, 118 00 Praha 1, Czech Republic*

Abstract. Motivated by the concept of partial words, we introduce an analogous concept of partial permutations. A *partial permutation of length n with k holes* is a sequence of symbols $\pi = \pi_1\pi_2\cdots\pi_n$ in which each of the symbols from the set $\{1, 2, \ldots, n-k\}$ appears exactly once, while the remaining k symbols of π are "holes".

We introduce pattern-avoidance in partial permutations and prove that most of the previous results on Wilf equivalence of permutation patterns can be extended to partial permutations with an arbitrary number of holes. We also show that Baxter permutations of a given length k correspond to a Wilf-type equivalence class with respect to partial permutations with $(k-2)$ holes. Lastly, we enumerate the partial permutations of length n with k holes avoiding a given pattern of length at most four, for each $n \geq k \geq 1$.

Résumé. Nous introduisons un concept de permutations partielles. *Une permutation partielle de longueur n avec k trous* est une suite finie de symboles $\pi = \pi_1\pi_2\cdots\pi_n$ dans laquelle chaque nombre de l'ensemble $\{1, 2, \ldots, n-k\}$ apparaît précisement une fois, tandis que les k autres symboles de π sont des "trous".

Nous introduisons l'étude des permutations partielles à motifs exclus et nous montrons que la plupart des résultats sur l'équivalence de Wilf peuvent être généralisés aux permutations partielles avec un nombre arbitraire de trous. De plus, nous montrons que les permutations de Baxter d'une longueur donnée k forment une classe d'équivalence du type Wilf par rapport aux permutations partielles avec $(k-2)$ trous. Enfin, nous présentons l'énumeration des permutations partielles de longueur n avec k trous qui évitent un motif de longueur $\ell \leq 4$, pour chaque $n \geq k \geq 1$.

Keywords: partial permutation, pattern avoidance, Wilf-equivalence, bijection, generating function, Baxter permutation

†Supported by the Icelandic Research Fund, grant no. 090038011.
‡Supported by project 1M0021620838 of the Czech Ministry of Education. The work was done while the third author was visiting ICE-TCS, Reykjavik University, Iceland.

1 Introduction

Let A be a nonempty set, which we call an *alphabet*. A *word* over A is a finite sequence of elements of A, and the *length* of the word is the number of elements in the sequence. Assume that $\diamond$ is a special symbol not belonging to A. The symbol $\diamond$ will be called *a hole*. A *partial word* over A is a word over the alphabet $A \cup \{\diamond\}$. In the study of partial words, the holes are usually treated as gaps that may be filled by an arbitrary letter of A. The *length* of a partial word is the number of its symbols, including the holes.

The study of partial words was initiated by Berstel and Boasson [BB99]. Partial words appear in comparing genes [Leu05]; alignment of two sequences can be viewed as a construction of two partial words that are compatible in the sense defined in [BB99]. Combinatorial aspects of partial words that have been studied include periods in partial words [BB99, SK01], avoidability/unavoidability of sets of partial words [BBSGR09, BSBK$^+$09], squares in partial words [HHK08], and overlap-freeness [HHKS09]. For more see the book by Blanchet-Sadri [BS08].

Let V be a set of symbols not containing $\diamond$. A *partial permutation of* V is a partial word π such that each symbol of V appears in π exactly once, and all the remaining symbols of π are holes. Let $\mathcal{S}_n^k$ denote the set of all partial permutations of the set $[n-k] = \{1, 2, \dots, n-k\}$ that have exactly k holes. For example, $\mathcal{S}_3^1$ contains the six partial permutations $12\diamond$, $1\diamond 2$, $21\diamond$, $2\diamond 1$, $\diamond 12$, and $\diamond 21$. Obviously, all elements of $\mathcal{S}_n^k$ have length n, and $|\mathcal{S}_n^k| = \binom{n}{k}(n-k)! = n!/k!$. Note that $\mathcal{S}_n^0$ is the familiar symmetric group $\mathcal{S}_n$. For a set $H \subset [n]$ of size k, we let $\mathcal{S}_n^H$ denote the set of partial permutations $\pi_1 \cdots \pi_n \in \mathcal{S}_n^k$ such that $\pi_i = \diamond$ if and only if $i \in H$. We remark that our notion of partial permutations is somewhat reminiscent of the notion of insertion encoding of permutations, introduced by Albert et al. [ALR05]. However, the interpretation of holes in the two settings is different.

In this paper, we extend the classical notion of pattern-avoiding permutations to the more general setting of partial permutations. Let us first recall some definitions related to pattern avoidance on permutations. Let $V = \{v_1, \dots, v_n\}$ with $v_1 < \cdots < v_n$ be any finite subset of $\mathbb{N}$. The *standardization* of a permutation π on V is the permutation $\mathrm{st}(\pi)$ on $[n]$ obtained from π by replacing the letter v_i with the letter i. As an example, $\mathrm{st}(19452) = 15342$. Given $p \in \mathcal{S}_k$ and $\pi \in \mathcal{S}_n$, an *occurrence* of p in π is a subword $\omega = \pi_{i(1)} \cdots \pi_{i(k)}$ of π such that $\mathrm{st}(\omega) = p$; in this context p is called a *pattern*. If there are no occurrences of p in π we also say that π *avoids* p. Two patterns p and q are called *Wilf-equivalent* if for each n, the number of p-avoiding permutations in $\mathcal{S}_n$ is equal to the number of q-avoiding permutations in $\mathcal{S}_n$.

Let $\pi \in \mathcal{S}_n^k$ be a partial permutation and let $i(1) < \cdots < i(n-k)$ be the indices of the non-hole elements of π. A permutation $\sigma \in \mathcal{S}_n$ is an *extension* of π if

$$\mathrm{st}(\sigma_{i(1)} \cdots \sigma_{i(n-k)}) = \pi_{i(1)} \cdots \pi_{i(n-k)}.$$

For example, the partial permutation $2\diamond 1$ has three extensions, namely 312, 321 and 231. In general, the number of extensions of $\pi \in \mathcal{S}_n^k$ is $\binom{n}{k}k! = n!/(n-k)!$.

We are now ready to define pattern avoidance on partial permutations. We say that $\pi \in \mathcal{S}_n^k$ *avoids* the *pattern* $p \in \mathcal{S}_\ell$ if each extension of π avoids p. For example, $\pi = 32\diamond 154$ avoids 1234, but π does not avoid 123: the permutation 325164 is an extension of π and it contains two occurrences of 123. Let $\mathcal{S}_n^k(p)$ be the set of all the partial permutations in $\mathcal{S}_n^k$ that avoid p, and let $s_n^k(p) = |\mathcal{S}_n^k(p)|$. Similarly, if $H \subseteq [n]$ is a set of indices, then $\mathcal{S}_n^H(p)$ is the set of p-avoiding permutations in $\mathcal{S}_n^H$, and $s_n^H(p)$ is its cardinality.

We say that two patterns p and q are k-*Wilf-equivalent* if $s_n^k(p) = s_n^k(q)$ for all n. Notice that 0-Wilf equivalence coincides with the standard notion of Wilf equivalence. We also say that two patterns p and

q are *$\star$-Wilf-equivalent* if p and q are k-Wilf-equivalent for all $k \geq 0$. Two patterns p and q are *strongly k-Wilf-equivalent* if $s_n^H(p) = s_n^H(q)$ for each n and for each k-element subset $H \subseteq [n]$. Finally, p and q are *strongly $\star$-Wilf-equivalent* if they are strongly k-Wilf-equivalent for all $k \geq 0$.

We note that although strong k-Wilf equivalence implies k-Wilf-equivalence, and strong $\star$-Wilf equivalence implies $\star$-Wilf equivalence, the converse implications are not true. Consider for example the patterns $p = 1342$ and $q = 2431$. A partial permutation avoids p if and only if its reverse avoids q, and thus p and q are $\star$-Wilf-equivalent. However, p and q are not strongly 1-Wilf-equivalent, and hence not strongly $\star$-Wilf equivalent either. To see this, we fix $H = \{2\}$ and easily check that $s_5^H(p) = 13$ while $s_5^H(q) = 14$.

Our Results

The main goal of this paper is to establish criteria for k-Wilf equivalence and $\star$-Wilf equivalence of permutation patterns. We are able to show that most pairs of Wilf-equivalent patterns that were discovered so far are in fact $\star$-Wilf-equivalent. The only exception is the pair of patterns $p = 2413$ and $q = 1342$. Although these patterns are known to be Wilf-equivalent [Sta94], they are neither 1-Wilf-equivalent nor 2-Wilf equivalent (see Section 6).

Many of our arguments rely on a close relationship between partial permutations and partial 01-fillings of Ferrers diagrams. These fillings are introduced in Section 2, where we also establish the link between partial fillings and partial permutations.

Our first main result is Theorem 5 in Section 3, which states that a permutation pattern of the form $123 \cdots \ell X$ is strongly $\star$-Wilf-equivalent to the pattern $\ell(\ell-1)\cdots 321X$, where $X = x_{\ell+1}x_{\ell+2}\cdots x_m$ is any permutation of $\{\ell+1, \dots, m\}$. This theorem is a strengthening of a result of Backelin, West and Xin [BWX07], who show that patterns of this form are Wilf-equivalent. Our proof is based on a different argument than the original proof of Backelin, West and Xin. The main ingredient of our proof is an involution on a set of fillings of Ferrers diagrams, discovered by Krattenthaler [Kra06]. We adapt this involution to partial fillings and use it to obtain a bijective proof of our result.

Our next main result is Theorem 6 in Section 4, which states that for any permutation X of the set $\{4, 5, \dots, k\}$, the two patterns $312X$ and $231X$ are strongly $\star$-Wilf-equivalent. This is also a refinement of an earlier result involving Wilf equivalence, due to Stankova and West [SW02]. As in the previous case, the refined version requires a different proof than the weaker version.

In Section 5, we study the k-Wilf equivalence of patterns whose length is small in terms of k. It is not hard to see that all patterns of length ℓ are k-Wilf equivalent whenever $\ell \leq k+1$, because $s_n^k(p) = 0$ for every such pattern p. Thus, the shortest patterns that exhibit nontrivial behaviour are the patterns of length $k+2$. For these patterns, we show that k-Wilf equivalence yields a new characterization of Baxter permutations: a pattern p of length $k+2$ is a Baxter permutation if and only if $s_n^k(p) = \binom{n}{k}$. For any non-Baxter permutation q of length $k+2$, $s_n^k(q)$ is strictly smaller than $\binom{n}{k}$ and is in fact a polynomial in n of degree at most $k-1$.

In Section 6, we focus on explicit enumeration of $s_n^k(p)$ for small patterns p. We obtain explicit closed-form formulas for $s_n^k(p)$ for every p of length at most four and every $k \geq 1$.

In view of the space constraints, most of the proofs have been omitted from this extended abstract.

An example: monotone patterns

Before we present our main results, let us illustrate the concept of pattern-avoiding partial permutations on the example of partial permutations avoiding the monotone pattern $12\cdots\ell$. Let $\pi \in \mathcal{S}_n^k$, and let

$\pi' \in \mathcal{S}_{n-k}$ be the permutation obtained from π by deleting all $\diamond$'s. Note that π avoids the pattern $12\cdots\ell$ if and only if π' avoids $12\cdots(\ell-k)$. Thus,

$$s_n^k(12\cdots\ell) = \binom{n}{k} s_n^0(12\cdots(\ell-k)), \tag{1}$$

where $\binom{n}{k}$ counts the possibilities of placing k $\diamond$'s. For instance, if $\ell = k+3$ then $s_n^k(12\cdots\ell) = \binom{n}{k}s_n^0(123)$, and it is well known that $s_n^0(123) = C_n$, the n-th Catalan number. We remark that for general ℓ, Regev [Reg81] found an asymptotic formula for $s_n^0(12\cdots\ell)$, which can be used to obtain a (rather complicated) asymptotic formula for $s_n^k(12\cdots\ell)$ as n tends to infinity.

2 Partial fillings

In this section, we introduce the necessary definitions related to partial fillings of Ferrers diagrams. These notions will later be useful in our proofs of $\star$-Wilf equivalence of patterns.

Let $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k)$ be a non-increasing sequence of k nonnegative integers. A *Ferrers diagram with shape* λ is a bottom-justified array D of cells arranged into k columns, such that the j-th column from the left has exactly λ_j cells. Note that our definition of Ferrers diagram is slightly more general than usual, in that we allow columns with no cells. If each column of D has at least one cell, then we call D a *proper Ferrers diagram*. Every row of a Ferrers diagram D has nonzero length (while we allow columns of zero height). If all the columns of D have zero height—in other words, D has no rows—then D is called *degenerate*.

For the sake of consistency, we assume throughout this paper that the rows of each diagram and each matrix are numbered from bottom to top, with the bottom row having number 1. Similarly, the columns are numbered from left to right, with column 1 being the leftmost column.

By *cell* (i,j) of a Ferrers diagram D we mean the cell of D that is the intersection of i-th row and j-th column of the diagram. We assume that the cell (i,j) is a unit square whose corners are lattice points with coordinates $(i-1, j-1)$, $(i, j-1)$, $(i-1, j)$ and (i,j). The point $(0,0)$ is the bottom-left corner of the whole diagram. We say a diagram D *contains* a lattice point (i,j) if either $j=0$ and the first column of D has height at least i, or $j>0$ and the j-th column of D has height at least i. A point (i,j) of the diagram D is a *boundary point* if the cell $(i+1, j+1)$ does not belong to D (see Figure 1). Note that a Ferrers diagram with r rows and c columns has $r+c+1$ boundary points.

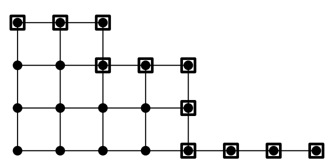


Fig. 1: A Ferrers diagram with shape $(3,3,2,2,0,0,0)$. The black dots represent the points. The black dots in squares are the boundary points.

A *01-filling* of a Ferrers diagram assigns to each cell the value 0 or 1. A 01-filling is *transversal* if each row and each column has exactly one 1-cell. A 01-filling is *sparse* if each column and each row has at most one 1-cell. A permutation $p = p_1 p_2 \cdots p_\ell \in \mathcal{S}_\ell$ can be represented by a *permutation matrix* which is a 01-matrix of size $\ell \times \ell$, whose cell (i, j) is equal to 1 if and only if $p_j = i$. If there is no risk of confusion, we abuse terminology by identifying a permutation pattern p with the corresponding permutation matrix. Note that a permutation matrix is a transversal filling of a diagram with square shape.

Let P be a permutation matrix of size $n \times n$, and let F be a sparse filling of a Ferrers diagram. We say that F *contains* P if F has a (not necessarily contiguous) square subdiagram of size $n \times n$ which induces in F a subfilling equal to P. This notion of containment generalizes usual permutation containment.

We now extend the notion of partial permutations to partial fillings of diagrams. Let D be a Ferrers diagram with k columns. Let H be a subset of columns of D. Let ϕ be a function that assigns to every cell of D one of the three symbols 0, 1 and $\diamond$, such that every cell in a column belonging to H is filled with $\diamond$, while every cell in a column not belonging to H is filled with 0 or 1. The pair $F = (\phi, H)$, will be referred to as *a partial 01-filling (or a partial filling) of the diagram* D. See Figure 2. The columns from the set H will be called *the $\diamond$-columns* of F, while the remaining columns will be called *the standard columns*. Observe that if the diagram D has columns of height zero, then ϕ itself is not sufficient to determine the filling F, because it does not allow us to determine whether the zero-height columns are $\diamond$-columns or standard columns. For our purposes, it is necessary to distinguish between partial fillings that differ only by the status of their zero-height columns.

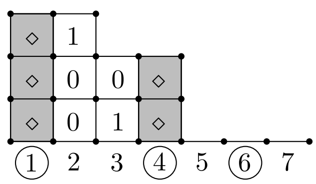


Fig. 2: A partial filling with $\diamond$-columns 1, 4 and 6.

We say that a partial 01-filling is *transversal* if every row and every standard column has exactly one 1-cell, and we say that a partial 01-filling is *sparse* if every row and every standard column has at most one 1-cell. A *partial 01-matrix* is a partial filling of a (possibly degenerate) rectangular diagram .

There is a natural correspondence between partial permutations and transversal partial 01-matrices. Let $\pi \in \mathcal{S}_n^k$ be a partial permutation. A *partial permutation matrix representing* π is a partial 01-matrix P with $n - k$ rows and n columns, with the following properties:

- If the j-th symbol of π is $\diamond$, then the j-th column of P is a $\diamond$-column.
- If the j-th symbol of π is a number i, then the j-th column is a standard column. Also, the cell in column j and row i is filled with 1, and the remaining cells in column j are filled with 0's.

To define pattern-avoidance for partial fillings, we first introduce the concept of substitution into a $\diamond$-column, which is analogous to substituting a number for a $\diamond$ in a partial permutation. The idea is to insert a new row with a 1-cell in the $\diamond$-column; this increases the height of the diagram by one. Let us now describe the substitution formally.

Let F be a partial filling of a Ferrers diagram with m columns. Assume that the j-th column of F is a $\diamond$-column. Let h be the height of the j-th column. A *substitution* into the j-th column is an operation consisting of the following steps:

1. Choose a number i, with $1 \leq i \leq h+1$.

2. Insert a new row into the diagram, between rows $i-1$ and i. The newly inserted row must not be longer than the $(i-1)$-th row, and it must not be shorter than the i-th row, so that the new diagram is still a Ferrers diagram. If $i=1$, we also assume that the length of the new row is at most m, so that the number of columns does not increase during the substitution.

3. Fill all the cells in column j with the symbol 0, except for the cell in the newly inserted row, which is filled with 1. Remove column j from the set of $\diamond$-columns.

4. Fill all the remaining cells of the new row with 0 if they belong to a standard column, and with $\diamond$ if they belong to a $\diamond$-column.

Figure 3 illustrates an example of substitution.

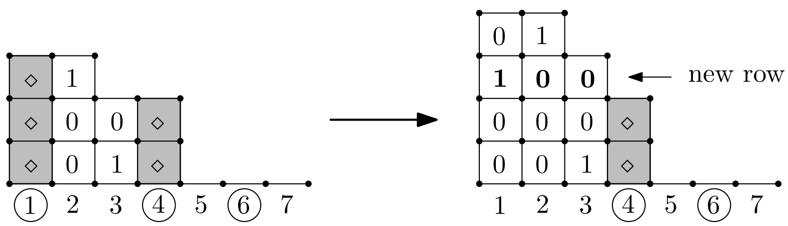


Fig. 3: A substitution into the first column of a partial filling, involving an insertion of a new row between the second and third rows of the original partial filling.

Note that a substitution into a partial filling increases the number of rows by 1. A substitution into a transversal (resp. sparse) partial filling produces a new transversal (resp. sparse) partial filling. A partial filling F with k $\diamond$-columns can be transformed into a (non-partial) filling F' by a sequence of k substitutions; we then say that F' is an *extension* of F.

Let P be a permutation matrix. We say that a partial filling F of a Ferrers diagram *avoids* P if every extension of F avoids P. Note that a partial permutation $\pi \in S_k^n$ avoids a permutation p, if and only if the partial permutation matrix representing π avoids the permutation matrix representing p.

We say that two permutation matrices P and Q are *shape-$\star$-Wilf-equivalent*, if for every Ferrers diagram D there is a bijection between P-avoiding and Q-avoiding partial transversal fillings of D that preserves the set of $\diamond$-columns. Observe that if two permutations are shape-$\star$-Wilf-equivalent, then they are also strongly $\star$-Wilf-equivalent, because a partial permutation is a special case of a partial transversal filling of a Ferrers diagram.

The notion of shape-$\star$-Wilf-equivalence is motivated by the following proposition, which extends an analogous result of Babson and West [BW00] for shape-Wilf-equivalence of non-partial permutations.

Proposition 1 *Let P and Q be shape-$\star$-Wilf-equivalent permutations, let X be an arbitrary permutation. Then the two permutations $\left(\begin{smallmatrix} 0 & X \\ P & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & X \\ Q & 0 \end{smallmatrix}\right)$ are strongly $\star$-Wilf-equivalent.*

Due to space constraints, the proof of Proposition 1 is omitted in this extended abstract.

3 Strong $\star$-Wilf-equivalence of $12\cdots\ell X$ and $\ell(\ell-1)\cdots 1X$

We will use Proposition 1 as the main tool to prove strong $\star$-Wilf equivalence. To apply the proposition, we need to find pairs of shape-$\star$-Wilf-equivalent patterns. A family of such pairs is provided by the next proposition, which extends previous results of Backelin, West and Xin [BWX07].

Proposition 2 *Let $I_\ell = 12\cdots\ell$ be the identity permutation of order ℓ, and let $J_\ell = \ell(\ell-1)\cdots 21$ be the anti-identity permutation of order ℓ. The permutations I_ℓ and J_ℓ are shape-$\star$-Wilf-equivalent.*

Before sketching the proof of this proposition, we introduce some notation and terminology. Let F be a partial filling with r rows and c columns. Let i and j be numbers such that the point (i,j) is in F. Let $F(\leq i, \leq j)$ denote the submatrix of F whose bottom-left corner is the point $(0,0)$ and whose top-right corner is the point (i,j); in other words, $F(\leq i, \leq j)$ is the intersection of the bottom i rows with the leftmost j columns of F. We assume that $F(\leq i, \leq 0)$ is the empty matrix, while $F(\leq 0, \leq j)$ is the degenerate matrix with no rows but with j columns of zero height.

Let F be a sparse partial filling of a Ferrers diagram, and let (i,j) be a boundary point of F. Let $h(F,j)$ denote the number of $\diamond$-columns among the first j columns of F. Let $I(F,i,j)$ denote the largest integer ℓ such that the partial matrix $F(\leq i, \leq j)$ contains I_ℓ. Similarly, let $J(F,i,j)$ denote the largest ℓ such that $F(\leq i, \leq j)$ contains J_ℓ.

We let F_0 denote the (non-partial) sparse filling obtained by replacing all the symbols $\diamond$ in F by zeros.

Let us state without proof the following simple observation.

Observation 3 *Let F be a sparse partial filling.*

1. *F contains a permutation matrix P if and only if F has a boundary point (i,j) such that $F(\leq i, \leq j)$ contains P.*
2. *For any boundary point (i,j), we have $I(F,i,j) = h(F,j) + I(F_0,i,j)$ and $J(F,i,j) = h(F,j) + J(F_0,i,j)$.*

The key to the proof of Proposition 2 is the following theorem, which follows directly from the powerful results of Krattenthaler [Kra06] obtained using the theory of growth diagrams.

Theorem 4 (Krattenthaler [Kra06]) *Let D be a Ferrers diagram. There is a bijective mapping κ from the set of all (non-partial) sparse fillings of D onto itself, with the following properties.*

1. *For any boundary point (i,j) of D, and for any sparse filling F, we have $I(F,i,j) = J(\kappa(F),i,j)$ and $J(F,i,j) = I(\kappa(F),i,j)$.*
2. *The mapping κ preserves the number of 1-cells in each row and column. In other words, if a row (or column) of a sparse filling F has no 1-cell, then the same row (or column) of $\kappa(F)$ has no 1-cell either.*

In Krattenthaler's paper, the results are stated in terms of proper Ferrers diagrams. However, the bijection obviously extends to Ferrers diagrams with zero-height columns as well. This is because adding zero-height columns to a (non-partial) filling does not affect pattern containment.

From the previous theorem, we easily obtain the proof of the main proposition in this section.

Proof of Proposition 2: Let D be a Ferrers diagram. Let F be an I_ℓ-avoiding transversal partial filling of D. Let F_0 be the sparse filling obtained by replacing all the $\diamond$ characters of F by zeros. Define

$G_0 = \kappa(F_0)$, where κ is the bijection from Theorem 4. Note that all the $\diamond$-columns of F are filled with zeros both in F_0 and G_0. Let G be the sparse partial filling obtained from G_0 by replacing zeros with $\diamond$ in all such columns. Then G is a sparse partial filling with the same set of $\diamond$-columns as F.

We see that for any boundary point (i, j) of the diagram D, $h(F, j) = h(G, j)$. By the properties of κ, we further obtain $I(F_0, i, j) = J(G_0, i, j)$. In view of Observation 3, this implies that G is a J_ℓ-avoiding filling. It is clear that this construction can be inverted, thus giving the required bijection between I_ℓ-avoiding and J_ℓ-avoiding transversal partial fillings of D. □

Combining Proposition 1 with Proposition 2, we get directly the main result of this section.

Theorem 5 *For any $\ell \leq m$, and for any permutation X of $\{\ell+1, \ldots, m\}$, the permutation pattern $123 \cdots (\ell-1)\ell X$ is strongly $\star$-Wilf-equivalent to the pattern $\ell(\ell-1) \cdots 21X$.*

Notice that Theorem 5 implies, among other things, that all the patterns of size three are strongly $\star$-Wilf-equivalent.

4 Strong $\star$-Wilf-equivalence of $312X$ and $231X$

We will now focus on the two patterns 312 and 231. The main result of this section is the following theorem.

Theorem 6 *The patterns 312 and 231 are shape-$\star$-Wilf-equivalent. By Proposition 1, this implies that for any permutation X of the set $\{4, 5, \ldots, m\}$, the two permutations $312X$ and $231X$ are strongly $\star$-Wilf-equivalent.*

Theorem 6 generalizes a result of Stankova and West [SW02], who have shown that 312 and 231 are shape-Wilf equivalent. The original proof of Stankova and West [SW02] is rather complicated, and does not seem to admit a straightforward generalization to the setting of shape-$\star$-Wilf-equivalence. Our proof of Theorem 6 is different from the argument of Stankova and West, and it is based on a bijection of Jelínek [Jel07], obtained in the context of pattern-avoiding ordered matchings.

Due to space limitations, we omit the whole very long proof from this extended abstract.

5 The k-Wilf-equivalence of patterns of length $k+2$

We will now consider the structure of pattern-avoiding partial permutations in which the number of $\diamond$'s is close to the length of the forbidden pattern.

Let us begin by an easy observation.

Observation 7 *Assume that p is a pattern of length ℓ. Any partial permutation with at least ℓ occurrences of $\diamond$ contains p. Almost as obviously, a partial permutation with $\ell - 1$ occurrences of $\diamond$ and of length at least ℓ, contains p as well. In particular, for every $k \geq \ell - 1$, we have $s_n^k(p) = 0$, and hence all the patterns of length ℓ are k-Wilf-equivalent.*

In the rest of this section, we will deal with k-Wilf-equivalence of patterns of length $\ell = k + 2$.

As we will see, an important part in k-Wilf-equivalence is played by Baxter permutations, which are defined as follows.

Definition 8 *A permutation $p \in \mathcal{S}_\ell$ is called a* Baxter permutation*, if there is no four-tuple of indices $a < b < c < d \in [\ell]$ such that*

- *$c = b + 1$, and*
- *the subpermutation p_a, p_b, p_c, p_d is order-isomorphic to 2413 or to 3142.*

In the terminology of Babson and Steingrímsson [BS00], Baxter permutations are exactly the permutations avoiding simultaneously the two patterns 2-41-3 and 3-14-2.

Baxter permutations have been introduced by G. Baxter [Bax64] in 1964. They were originally encountered in the context of common fixed points of commuting continuous functions [Bax64, Boy81]. Later, it has been discovered that Baxter permutations are also closely related to other combinatorial structures, such as plane bipolar orientations [BBMF08], noncrossing triples of lattice paths [FFNO08], and standard Young tableaux [DG96]. An explicit formula for the number of Baxter permutations has been found by Chung et al. [CGJK78], with several later refinements [Mal79, Vie81, DG98].

It is not hard to see that for any pattern of length $\ell = k+2$, and for any n from the set $\{k, k+1, k+2\}$, we always have $s_n^k(p) = \binom{n}{k}$. Thus, for these small values of n, all patterns have the same behavior. However, for all larger values of n, the Baxter patterns are separated from the rest, as the next proposition and theorem show. We omit the proofs of these results.

Proposition 9 *Let p be a permutation pattern of size ℓ, and let $k = \ell - 2$. The following statements are equivalent.*

1. *The pattern p is a Baxter permutation.*
2. *For each $n \geq k$ and each k-element subset $H \subseteq [n]$, $s_n^H(p) = 1$.*
3. *For $n = k+3$ and each k-element subset $H \subseteq [n]$, $s_n^H(p) = 1$.*
4. *There exists $n \geq k+3$ such that for each k-element subset $H \subseteq [n]$, $s_n^H(p) = 1$.*

Theorem 10 *Let $p \in \mathcal{S}_\ell$ be a permutation pattern. Let $k = \ell - 2$. If p is a Baxter permutation then $s_n^k(p) = \binom{n}{k}$ for each $n \geq k$. If p is not a Baxter permutation, then $s_n^k(p) < \binom{n}{k}$ whenever $n \geq k+3$. Moreover, all the Baxter permutations are strongly k-Wilf equivalent.*

We remark that by a slightly more careful analysis of the arguments leading to Proposition 9 and Theorem 10, we could give a stronger upper bound for $s_n^k(p)$ when p is not a Baxter permutation. In particular, it is not hard to show that in that case, $s_n^k(p)$ is eventually equal to a polynomial in n of degree at most $k-1$, with coefficients depending on k.

6 Short patterns

In the rest of this paper, we focus on explicit formulas for $s_n^k(p)$, where p is a pattern of length ℓ. We may assume that $k < \ell - 1$, and $\ell > 2$, since for any other values of (k, ℓ) the enumeration is trivial (see Observation 7). We also restrict ourselves to $k \geq 1$, since the case $k = 0$, which corresponds to classical pattern-avoidance in permutations, has already been extensively studied [BÓ4].

For a pattern p of length three, the situation is very simple. Theorem 10 implies that $s_n^1(p) = n$, since all permutations of length three are Baxter permutations.

Let us now deal with patterns of length four. In Figure 4, we depict the k-Wilf equivalence classes, where the four rows, top to bottom, correspond to the four values $k = 0, 1, 2, 3$. Since all the k-Wilf equivalences are closed under complements and reversals (but not inversions), we represent the 24 patterns of length four by eight representatives, one from each symmetry class. For instance, $\{1342, 1423\}$ in the second row represents the union of $\{1342, 2431, 3124, 4213\}$ and $\{1423, 2314, 3241, 4132\}$.

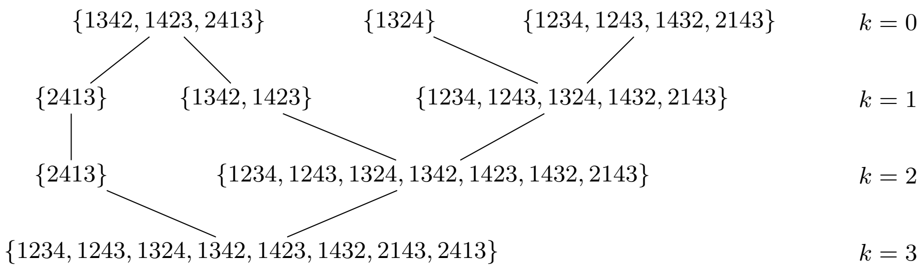


Fig. 4: The k-Wilf-equivalence classes of permutations of size 4.

All patterns p of length four except 2413 and 3142 are Baxter permutations, and hence they satisfy $s_n^2(p) = \binom{n}{2}$ by Theorem 10. It is possible to show that $s_n^2(2413) = s_n^2(3142) = 3n - 6$. We omit the details of this routine argument in this extended abstract.

In the rest of the paper, we deal with 1-Wilf equivalence of patterns of length four, and with the enumeration of the corresponding avoidance classes. Theorem 5 and symmetry arguments imply that all the patterns 1234, 1243, 1432 and 2143 are strongly $\star$-Wilf-equivalent, and Theorem 6 with appropriate symmetry arguments shows that 1342 and 1423 are strongly $\star$-Wilf-equivalent as well. The only case not covered by these general theorems is the 1-Wilf equivalence of 1324 and 1234, which is handled separately by the next proposition.

Proposition 11 *The patterns 1234 and 1324 are strongly 1-Wilf-equivalent.*

The proof of Proposition 11 is omitted.

Let us now state the formulas for $s_n^1(p)$, where p has length four. The proofs are omitted.

Theorem 12 *The number of partial permutations of length $n \geq 1$ with a single hole, avoiding a pattern of length four, is given by these formulas:*

- $s_n^1(1234) = s_n^1(1243) = s_n^1(1324) = s_n^1(1432) = s_n^1(2143) = \binom{2n-2}{n-1}$,
- $s_n^1(1342) = s_n^1(1423) = \binom{2n-2}{n-1} - \binom{2n-2}{n-5}$, *and*
- $s_n^1(2413) = \frac{2}{n+1}\binom{2n}{n} - 2^{n-1}$.

7 Directions of further research

We have shown that classical Wilf equivalence may be regarded as a special case in a hierarchy of k-Wilf equivalence relations, and that many properties previously established in the context of Wilf equivalence can be generalized to all the k-Wilf equivalences. In many situations, understanding the k-Wilf equivalence class of a given pattern p becomes easier as k increases. In particular, the k-Wilf equivalence class of the permutation $p = 12 \cdots (k+1)(k+2)$ contains exactly the Baxter permutations of length $k + 2$.

What can we say about the k-Wilf equivalence class of the permutation $12 \cdots (k+3)$? For $k = 0$ and $k = 1$ this class contains exactly the layered permutations of length $k + 3$. Computer enumeration suggests that the same is true for larger values of k as well, but we have no proof.

References

[ALR05] Michael H. Albert, Steve Linton, and Nik Ruškuc. The insertion encoding of permutations. *Electron. J. Combin*, 12(1), 2005. Research paper 47, 31 pp.

[B0́4] M. Bóna. *Combinatorics of Permutations*. Discrete Mathematics and its Applications. Chapman and Hall/CRC Press, 2004.

[Bax64] G. Baxter. On fixed points of the composite of commuting functions. *Proceedings of the American Mathematical Society*, 15(6):851 – 855, 1964.

[BB99] J. Berstel and L. Boasson. Partial words and a theorem of Fine and Wilf. *Theoretical Computer Science*, 218(1):135 – 141, 1999.

[BBMF08] N. Bonichon, M. Bousquet-Mélou, and É. Fusy. Baxter permutations and plane bipolar orientations. *Electronic Notes in Discrete Mathematics*, 31:69 – 74, 2008. The International Conference on Topological and Geometric Graph Theory.

[BBSGR09] B. Blakeley, F. Blanchet-Sadri, J. Gunter, and N. Rampersad. *Developments in Language Theory*, volume 5583 of *LNCS*, chapter On the Complexity of Deciding Avoidability of Sets of Partial Words, pages 113 – 124. Springer, 2009.

[Boy81] W. M. Boyce. Baxter permutations and functional composition. *Houston Journal of Mathematics*, 7(2):175 – 189, 1981.

[BS00] E. Babson and E. Steingrímsson. Generalized permutation patterns and a classification of the Mahonian statistics. *Séminaire Lotharingien de Combinatoire*, 44:18, 2000.

[BS08] F. Blanchet-Sadri. *Algorithmic combinatorics on partial words*. Discrete Mathematics and its Applications. Chapman & Hall/CRC, Boca Raton, FL, 2008.

[BSBK$^+$09] F. Blanchet-Sadri, N. C. Brownstein, A. Kalcic, J. Palumbo, and T. Weyand. Unavoidable sets of partial words. *Theory of Computing Systems*, 45(2):381–406, 2009.

[BW00] E. Babson and J. West. The permutations $123p_4 \cdots p_m$ and $321p_4 \cdots p_m$ are Wilf-equivalent. *Graphs and Combinatorics*, 16(4):373 – 380, 2000.

[BWX07] J. Backelin, J. West, and G. Xin. Wilf-equivalence for singleton classes. *Advances in Applied Mathematics*, 38(2):133 – 148, 2007.

[CGJK78] F. R. K. Chung, R. L. Graham, V. E. Hoggatt Jr., and M. Kleiman. The number of Baxter permutations. *J. Combin. Theory Ser. A*, 24(3):382 – 394, 1978.

[DG96] S. Dulucq and O. Guibert. Stack words, standard tableaux and Baxter permutations. *Discrete Mathematics*, 157(1-3):91 – 106, 1996.

[DG98] S. Dulucq and O. Guibert. Baxter permutations. *Discrete Mathematics*, 180(1-3):143 – 156, 1998. Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics.

[FFNO08] S. Felsner, E. Fusy, M. Noy, and D. Orden. Bijections for Baxter families and related objects. `arXiv:0803.1546`, 2008.

[HHK08] V. Halava, T. Harju, and T. Kärki. Square-free partial words. *Information Processing Letters*, 108(5):290 – 292, 2008.

[HHKS09] V. Halava, T. Harju, T. Kärki, and P. Séébold. Overlap-freeness in infinite partial words. *Theoretical Computer Science*, 410(8-10):943 – 948, 2009.

[Jel07] V. Jelínek. Dyck paths and pattern-avoiding matchings. *European Journal of Combinatorics*, 28(1):202 – 213, 2007.

[Kra06] C. Krattenthaler. Growth diagrams, and increasing and decreasing chains in fillings of Ferrers shapes. *Advances in Applied Mathematics*, 37(3):404 – 431, 2006.

[Leu05] P. Leupold. Partial words for dna coding. In *DNA 10, Tenth International Meeting on DNA Computing*, volume 3384 of *LNCS*, pages 224 – 234. Springer-Verlag, Berlin, 2005.

[Mal79] C. L. Mallows. Baxter permutations rise again. *Journal of Combinatorial Theory, Series A*, 27(3):394 – 396, 1979.

[Reg81] Amitai Regev. Asymptotic values for degrees associated with strips of Young diagrams. *Advances in Mathematics*, 41(2):115 – 136, 1981.

[SK01] A. M. Shur and Y. V. Konovalova. On the periods of partial words. In *MFCS '01: Proceedings of the 26th International Symposium on Mathematical Foundations of Computer Science*, volume 2136 of *LNCS*, pages 657 – 665. Springer-Verlag, 2001.

[Sta94] Z. Stankova. Forbidden subsequences. *Discrete Mathematics*, 132(1-3):291–316, 1994.

[SW02] Z. Stankova and J. West. A new class of Wilf-equivalent permutations. *J. Algebraic Comb.*, 15(3):271–290, 2002.

[Vie81] G. Viennot. A bijective proof for the number of Baxter permutations. *Séminaire Lotharingien de Combinatoire*, 1981.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 510–521

$n!$ *matchings,* $n!$ *posets (extended abstract)*

Anders Claesson[1†] and Svante Linusson[2‡]

[1] *The Mathematics Institute, School of Computer Science, Reykjavik University, Menntavegi 1, IS-101 Reykjavik, Iceland*
[2] *Department of Mathematics, KTH-Royal Institute of Technology, SE-100 44 Stockholm, Sweden*

Abstract. We show that there are $n!$ matchings on $2n$ points without, so called, left (neighbor) nestings. We also define a set of naturally labeled $(2+2)$-free posets, and show that there are $n!$ such posets on n elements. Our work was inspired by Bousquet-Mélou, Claesson, Dukes and Kitaev [J. Combin. Theory Ser. A. 117 (2010) 884–909]. They gave bijections between four classes of combinatorial objects: matchings with no neighbor nestings (due to Stoimenow), unlabeled $(2+2)$-free posets, permutations avoiding a specific pattern, and so called ascent sequences. We believe that certain statistics on our matchings and posets could generalize the work of Bousquet-Mélou et al. and we make a conjecture to that effect. We also identify natural subsets of matchings and posets that are equinumerous to the class of unlabeled $(2+2)$-free posets.

We give bijections that show the equivalence of (neighbor) restrictions on nesting arcs with (neighbor) restrictions on crossing arcs. These bijections are thought to be of independent interest. One of the bijections maps via certain upper-triangular integer matrices that have recently been studied by Dukes and Parviainen [Electron. J. Combin. 17 (2010) #R53]

Résumé. Nous montrons qu'il y a $n!$ couplages sur $2n$ points sans emboîtement (de voisins) à gauche. Nous définissons aussi un ensemble d'EPO (ensembles partiellement ordonnés) sans motif $(2+2)$ naturellement étiquetés, et montrons qu'il y a $n!$ tels EPO sur n éléments. Notre travail a été inspiré par Bousquet-Mélou, Claesson, Dukes et Kitaev [J. Combin. Theory Ser. A. 117 (2010) 884–909]. Ces auteurs donnent des bijections entre quatre classes d'objets combinatoires: couplages sans emboîtement de voisins (dû à Stoimenow), EPO sans motif $(2+2)$ non étiquetés, permutations évitant un certain motif, et des objets appelés suites à montées. Nous pensons que certaines statistiques sur nos couplages et nos EPO pourraient généraliser le travail de Bousquet-Mélou et al. et nous proposons une conjecture à ce sujet. Nous identifions aussi des sous-ensembles naturels de couplages et d'EPO qui sont énumérés par la même séquence que la classe des EPO sans motif $(2+2)$ non étiquetés.

Nous donnons des bijections qui démontrent l'équivalence entre les restrictions sur les emboîtements (d'arcs voisins) et les restrictions sur les croisements (d'arcs voisins). Nous pensons que ces bijections présentent un intérêt propre. L'une de ces bijections passe par certaines matrices triangulaires supérieures à coefficients entiers qui ont été récemment étudiées par Dukes et Parviainen [Electron. J. Combin. 17 (2010) #R53]

Keywords: matching, poset, inversion table, permutation, ascent sequence, nesting, crossing

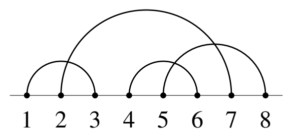
† Claesson was supported by grant no. 090038011 from the Icelandic Research Fund.
‡ Linusson is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

1 Introduction

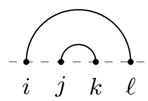
A *matching* of the integers $\{1, 2, \ldots, 2n\}$ is a partition of that set into blocks of size 2. An example of a matching is

$$M = \{(1, 3), (2, 7), (4, 6), (5, 8)\}.$$

In the diagram below there is an *arc* connecting i with j precisely when $(i, j) \in M$.



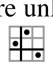
A *nesting* of M is a pair of arcs (i, ℓ) and (j, k) with $i < j < k < \ell$:

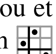


We call such a nesting a *left-nesting* if $j = i + 1$. Similarly, we call it a *right-nesting* if $\ell = k + 1$. The example matching has one nesting, formed by the two arcs $(2, 7)$ and $(4, 6)$. It is a right-nesting.

To give upper bounds on the dimension of the space of Vassiliev's knot invariants of a given degree, Stoimenow [14] was led to introduce what he calls regular linearized chord diagrams. In the terminology of this paper, Stoimenow's diagrams are matchings with no *neighbor nestings*, that is, matchings with neither left-nestings, nor right-nestings. Following Stoimenow's paper, Zagier [16] derived the following beautiful generating function enumerating such matchings with respect to size:

$$\sum_{n \geq 0} \prod_{i=1}^{n} \left(1 - (1 - t)^i\right).$$

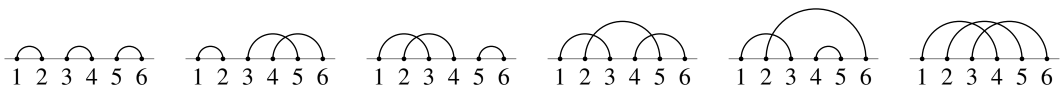
Recently, Bousquet-Mélou et al. [2] gave bijections between matchings on $[2n]$ with no neighbor nestings and three other classes of combinatorial objects, thus proving that they are equinumerous. The other classes were unlabeled $(\mathbf{2} + \mathbf{2})$-free posets (or interval orders) on n nodes; permutations on $[n]$ avoiding the pattern ; and ascent sequences of length n. Let $\mathfrak{f}_n$ be the cardinality of any, and thus all, of the above classes—it is the coefficient in front of t^n in Zagier's generating function. We call $\mathfrak{f}_n$ the nth *Fishburn number*; the first few numbers are $1, 1, 2, 5, 15, 53, 217, 1014, 5335, 31240$. Fishburn [7, 8, 9] did pioneering work on interval orders; for instance, he showed the basic theorem that a poset is an interval order if and only if it is $(\mathbf{2} + \mathbf{2})$-free.

The pattern avoiding permutations and the ascent sequences were both defined by Bousquet-Mélou et al. We shall recall those definitions there. In a permutation $\pi = a_1 \ldots a_n$, an occurrence of the pattern  is a 3 letter subsequence $a_i a_{i+1} a_j$ of π such that $a_j + 1 = a_i < a_{i+1}$. As an example, the permutation $\pi = 351426$ contains one such occurrence, namely 352. If π contains no such occurrence we say that π *avoids* the pattern. An integer sequence $(x_1, \ldots, x_n)$ is an *ascent sequences* if

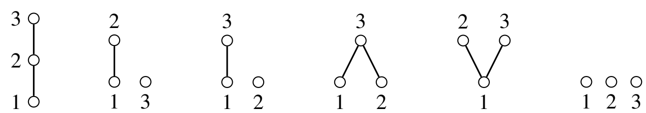
$$x_1 = 0 \quad \text{and} \quad 0 \leq x_i \leq 1 + \operatorname{asc}(x_1, \ldots, x_{i-1}),$$

for $2 \leq i \leq n$. Here, $\operatorname{asc}(x_1, \ldots, x_k)$ denotes the number of ascents in $(x_1, \ldots, x_k)$, and an ascent is a $j \in [k-1]$ such that $x_j < x_{j+1}$. Bousquet-Mélou et al. [2] derived a closed expression for the generating function enumerating ascent sequences with respect to length and number of ascents; hence they gave a new proof of Zagier's result, or rather a refinement of it.

Recall that Stoimenow's diagrams are matchings with no neighbor nestings. The discovery that led to the present paper is that there are exactly $n!$ matchings on $[2n]$ with no left-nestings (Theorem 1). As an example, these are the 6 such matchings on $\{1, \ldots, 6\}$:

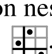


Can we also "lift" ascent sequences and unlabeled $(\mathbf{2}+\mathbf{2})$-free posets to the level of all permutations? That is, can we define "certain sequences" and "certain posets", both of cardinality $n!$, that are supersets of ascent sequences and unlabeled $(\mathbf{2}+\mathbf{2})$-free posets, respectively? For ascents sequences this is easy, and inversion tables is a natural choice. The poset case is more challenging. However, we show (Definition 2 and Theorem 5) that there are exactly $n!$ naturally labeled posets P on $[n]$ such that $i <_P k$ whenever $i < j <_P k$ for some $j \in [n]$; we call them *factorial posets*. Here is a list of the 6 factorial posets on $\{1, 2, 3\}$:



It is not hard to see (Proposition 4) that factorial posets are $(\mathbf{2}+\mathbf{2})$-free. Moreover, we give an additional restriction on the labeling of factorial posets under which the labeling is unique (Proposition 6), and thus the subset of factorial posets meeting that restriction is trivially in bijection with unlabeled $(\mathbf{2}+\mathbf{2})$-free posets.

The bijections we give to prove that inversion tables, factorial posets and matchings with no left-nesting are equinumerous do however not specialize to give back the results from [2]. This remains an interesting challenge. In Section 5 we prove that we could have studied matchings with restrictions on crossings instead of on nestings and present bijections to verify this.

Let $p =$ . As mentioned before, Bousquet-Mélou et al. [2] gave a bijection between matchings with no neighbor nestings and p-avoiding permutations. We conjecture (Conjecture 19) a generalization of that result. Namely, we conjecture that the distribution of right-nestings over matchings on $[2n]$ with no left-nestings coincides with the distribution of p over permutations on $[n]$.

In a recent paper, Dukes and Parviainen [6] study upper triangular matrices with non-negative integer entries such that each row and column has at least one nonzero entry and the total sum of the entries is n. They provide a recursive encoding of those matrices as ascent sequences. We have found a direct bijection (Theorem 8) from the same matrices to matchings with no neighbor nestings. In addition, we show (Proposition 11) that the subset of the matrices whose entries are 0 or 1 are in bijection with matchings with no left-nestings and no right-crossings.

2 Matchings with no left-nestings

Let $\mathcal{M}_n$ be the set of matchings on $[2n]$, and let $M \in \mathcal{M}_n$. If $i < j$ and $\alpha = (i, j)$ is an arc of M we call i the *opener* of α, and we call j the *closer* of α. In what follows it will be convenient to order the arcs with respect to closer. In particular, "the last arc" refers to the arc with closer $2n$. In the introduction we defined what left- and right-nestings are, and by $\mathrm{lne}(M)$ and $\mathrm{rne}(M)$ we shall denote the number of left- and right-nestings, respectively. Let

$$\mathcal{N}_n = \{\, M \in \mathcal{M}_n : \mathrm{lne}(M) = 0 \,\}$$

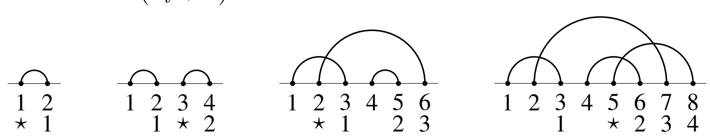
and $\mathcal{N} = \cup_{n\geq 0}\mathcal{N}_n$. Define $\mathcal{I}_n$ as the Cartesian product $\mathcal{I}_n = [0,0] \times [0,1] \times \cdots \times [0, n-1]$, where $[i,j] = \{i, i{+}1, \ldots, j\}$. In other words, $\mathcal{I}_n$ is the set of inversion tables of length n. Also, let $\mathcal{I} = \cup_{n\geq 0}\mathcal{I}_n$.

Theorem 1 *Matchings of* $[2n]$ *with no left-nestings are in bijection with inversion tables of length* n*, and thus* $|\mathcal{N}_n| = n!$.

Proof: Using recursion we define a bijection $f : \mathcal{I} \to \mathcal{N}$. Let $f(\epsilon) = \emptyset$, that is, let the empty inversion table map to the empty matching. Let $w = (a_1, \ldots, a_n)$ be any inversion table in $\mathcal{I}_n$ with $n > 0$. Let $w' = (a_1, \ldots, a_{n-1})$ and let $M' = f(w')$. Now create a matching M in $\mathcal{N}_n$ by inserting a new last arc in M' whose opener is immediately to the left of the $(a_n + 1)$st closer of M' if $a_n < n - 1$ and immediately to the left of its own closer if $a_n = n - 1$. Set $f(w) = M$. Note that the opener of the last arc has to be immediately to the left of some closer, otherwise a left-nesting would be created. Also note that removing the last arc from a matching in $\mathcal{N}_n$ cannot create a left-nesting. From a simple induction argument it thus follows that the described map is a bijection.

It is also easy to give a direct, non-recursive, description of the inverse of f. Indeed, $f^{-1}(M) = (a_1, \ldots, a_n)$ where a_i is the number of closers to the left of the opener of the ith arc; here, as before, arcs are ordered by closer. ☐

As an example, let $w = (a_1, a_2, a_3, a_4) = (0, 1, 0, 1)$. To construct the matching corresponding to that inversion table we insert the arcs one at the time, so that—as in the proof—the opener of the new last arc is immediately to the left of the $(a_i + 1)$st closer:



Here the star marks the opener of the new arc. Reading the number to the right of the star we get $(1, 2, 1, 2)$ and subtracting one from each coordinate we recover the inversion table $(0, 1, 0, 1)$.

3 Factorial posets

A poset P of cardinality n is said to be labeled if its elements are identified with the integers $1, \ldots, n$. A poset P is *naturally labeled* if $i < j$ in P implies $i < j$ in the usual order.

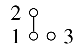
Definition 2 *We call a naturally labeled poset* P *on* $[n]$ *such that, for* $i, j, k \in [n]$,

$$i < j <_P k \implies i <_P k$$

a factorial poset*, and by* $\mathcal{F}_n$ *we denote the set of factorial posets on* $[n]$*. Similarly, we call a naturally labeled poset* P *on* $[n]$ *such that, for* $i, j, k \in [n]$,

$$i > j >_P k \implies i >_P k$$

a dually factorial poset.

There are 6 factorial posets on $\{1,2,3\}$, and we listed them on page 1. It is easy to check that of those posets, exactly one is not dually factorial: With $P =$  we have $3 > 2 >_P 1$, but $3 \not>_P 1$.

Definition 3 *The* predecessor set *of $j \in P$ is* $\mathrm{Pred}(j) = \{i : i <_P j\}$*, and we denote by* $\mathrm{pred}(j) = \#\,\mathrm{Pred}(j)$ *the number of predecessors of j. Similarly we define* $\mathrm{Succ}(j) = \{i : i >_P j\}$ *as the* successor set *of j and* $\mathrm{succ}(j) = \#\,\mathrm{Succ}(j)$ *as the number of successors of j.*

Note that P is factorial if, and only if, for all k in P, there is a j in $[0, n-1]$, such that $\mathrm{Pred}(k) = [1, j]$. It is well known—see for example Bogart [1]—that a poset is $(\mathbf{2}+\mathbf{2})$-free if, and only if, the collection $\{\,\mathrm{Pred}(k) : k \in P\,\}$ of predecessor sets can be linearly ordered by inclusion; hence the following proposition.

Proposition 4 *Factorial posets are $(\mathbf{2}+\mathbf{2})$-free.*

Theorem 5 *Factorial posets on $[n]$ are in bijection with inversion tables of length n, and thus $|\mathcal{F}_n| = n!$.*

Proof: Define $g : \mathcal{F}_n \to \mathcal{I}_n$ by $g(P) = (a_1, \ldots, a_n)$ where $a_k = \mathrm{pred}(k)$. To see that g is a bijection we describe its inverse. Given an inversion table $w = (a_1, \ldots, a_n)$ in $\mathcal{I}_n$ we construct a factorial poset $P = P(w)$ by postulating that $i <_P k$ precisely when $1 \le i \le a_k$. That this definition is consistent is easily seen by building P recursively. $\square$

We now have two bijections, f from inversion tables to matchings with no left-nestings, and g from factorial posets to inversion tables. Let $h = f \circ g$ be their composition:

$$\begin{tikzcd} \mathcal{F}_n \arrow[rr, "h"] \arrow[dr, "g"'] & & \mathcal{N}_n \\ & \mathcal{I}_n \arrow[ur, "f"'] & \end{tikzcd}$$

Let $P \in \mathcal{F}_n$. From the proofs of Theorems 5 and 1 it is immediate that to build $M = h(P)$ we insert the arcs one at the time so that, in the ith step, the opener of the new last arc is immediately to the left of the $(\mathrm{pred}(i)+1)$st closer.

Next we describe the inverse map, h^{-1}. Take $M \in \mathcal{N}_n$ and let $\alpha_1, \ldots, \alpha_n$ be its arcs ordered by closer. Then $i < j$ in $P = h^{-1}(M)$ if and only if the closer of α_i is to the left of the opener of α_j.

An *interval order* is a poset with the property that each element x can be assigned an interval $I(x)$ of real numbers so that $x < y$ in the poset if and only if every point in $I(x)$ is less than every point in $I(y)$. Such an assignment is called an *interval representation* of the poset. In 1970, Fishburn [8] showed that a poset is $(\mathbf{2}+\mathbf{2})$-free precisely when it has an interval representation. Let us for a moment identify the arcs of a matching with intervals of the real line. Then the function h, above, gives an interval representation of each factorial poset.

4 A unique labeling

Let $M \in \mathcal{N}_n$ and let $\alpha_1, \ldots, \alpha_n$ be its arcs ordered by closer. Let $P = h^{-1}(M)$. Assume that $1 \le i < j \le n$ in the usual order. Note that if α_i and α_j form a nesting, then we cannot have $\mathrm{pred}(i) = \mathrm{pred}(j)$ since then it would be a left-nesting which can never occur by the definition of g^{-1}. Thus α_i and α_j form a nesting precisely when $\mathrm{pred}(i) > \mathrm{pred}(j)$. If, in addition, $j = i+1$ and $\mathrm{succ}(i) = \mathrm{succ}(j)$ then α_i and

α_j form a right-nesting. Thus M is non-neighbor-nesting precisely when for each $i \in [n-1]$ we have $\mathrm{pred}(i) \le \mathrm{pred}(i+1)$ or $\mathrm{succ}(i) > \mathrm{succ}(i+1)$. By applying the bijection of Bousquet-Mélou et al. [2] from non-neighbor-nesting matchings to unlabeled $(\mathbf{2}+\mathbf{2})$-free posets we get the following result.

Proposition 6 *Factorial posets on* $[n]$ *such that for each* $i \in [n-1]$ *we have*

$$\mathrm{pred}(i) \le \mathrm{pred}(i+1) \;\; \textit{or} \;\; \mathrm{succ}(i) > \mathrm{succ}(i+1) \tag{1}$$

are in bijection with unlabeled $(\mathbf{2}+\mathbf{2})$*-free posets on* n *nodes; hence there are exactly* $\mathfrak{f}_n$ *such posets.*

An alternative way to see the above result is that given an unlabeled $(\mathbf{2}+\mathbf{2})$-free poset P there is exactly one way to label P so that the resulting poset is factorial and satisfies (1). The key observation to such a labeling is that if P is factorial and (1) holds then the pairs $(\mathrm{succ}(1), \mathrm{pred}(1)), \ldots, (\mathrm{succ}(n), \mathrm{pred}(n))$ are ordered weakly decreasing with respect to the first coordinate, and on equal first coordinate weakly increasing with respect to the second coordinate. We, however, omit the details of this argument.

5 Crossings versus nestings

A *crossing* of a matching M is a pair of arcs (i,k) and (j,ℓ) with $i<j<k<\ell$, and we can define left- and right-crossings analogously to how it was defined for nesting arcs. With A and B as in the table below there are bijections between

$$\{M \in \mathcal{M}_n : M \text{ is non } A\} \text{ and } \{M \in \mathcal{M}_n : M \text{ is non } B\}.$$

A	B
nesting	crossing
neighbor nesting	neighbor crossing
left-nesting	left-crossing

The first case is well known: for bijections between non-nesting matchings and non-crossing matchings see for instance [4, 5, 10]. We give bijections for the two remaining cases in this section. There exist a more complicated bijection [3] that can explain all three levels at once; see comment at the end of this section.

The second case is the most challenging, so let us look at the third case first. The proof of Theorem 1 gives a bijection f from inversion tables to non-left-nesting matchings. That bijection can be modified to give a bijection f_{nc} from inversion tables to non-left-crossing matchings (Theorem 7), and so $f_{\mathrm{nc}} \circ f^{-1}$ is a bijection from non-left-nesting to non-left-crossing matchings.

Theorem 7 *Matchings of* $[2n]$ *with no left-crossing are in bijection with inversion tables of length* n*; hence there are exactly* $n!$ *such matchings.*

Proof: As in the proof of Theorem 1 we define a bijection f_{nc} recursively. The difference is that this time the opener of the new last arc is immediately to the right of the a_nth closer if $a_n > 0$, or to the extreme left if $a_n = 0$. $\square$

For the second case, we shall give a bijection via matrices of a certain kind. Let $\mathcal{T}_n$ be the set of upper triangular matrices with non-negative integer entries, such that no row or column has only zeros and the total sum of the entries is n. These matrices have recently been studied by Dukes and Parviainen [6,

§2]; they gave a recursive encoding of the matrices in $\mathcal{T}_n$ as ascent sequences, and thus they showed that $|\mathcal{T}_n| = \mathfrak{f}_n$. This fact seems to have been first observed by Vladeta Jovovic [11]. We shall give a surjection ψ from the set of matchings of $[2n]$ to $\mathcal{T}_n$. Further, we shall show that if ψ is restricted to non-neighbor-nesting matchings, or non-neighbor-crossing matchings, then ψ is a bijection.

Before we describe ψ we need a few definitions. Let M be a matching and let $O(M)$ and $C(M)$ be the set of openers and closers of M, respectively. Write

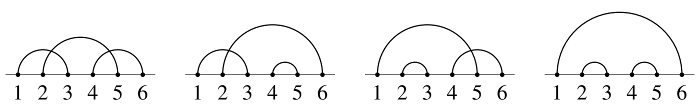
$$O(M) = O_1 \cup \cdots \cup O_k \ \text{ and } \ C(M) = C_1 \cup \cdots \cup O_\ell$$

as disjoint unions of maximal intervals. Clearly, $k = \ell$; we denote this number $\mathrm{int}(M)$. As an example, for $M = \{(1,2),(3,5),(4,6)\}$ we have $O(M) = [1,1] \cup [3,4]$, $C(M) = [2,2] \cup [5,6]$ and $\mathrm{int}(M) = 2$.

We are now in a position to define the promised map from matchings to matrices. Assume that M is a matching and that its intervals of openers and closers are ordered in the natural order. Let $\psi(M) = T$ where $T = (t_{ij})$ is an $\mathrm{int}(M) \times \mathrm{int}(M)$ matrix and

$$t_{ij} = |M \,\cap\, O_i \times C_j|.$$

In other words, t_{ij} is the number of arcs whose opener is in O_i and closer in C_j. For instance, the preimage of $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ under ψ consists of the following 4 matchings:



Note that of these matchings exactly one has no neighbor nestings and exactly one has no neighbor crossings. We shall see that this is no coincidence. (For brevity the proofs of the fllowing results have been excluded from this extended abstract.)

Theorem 8 *When restricted to matchings of* $[2n]$ *with no neighbor nestings* ψ *is a bijection onto* $\mathcal{T}_n$*.*

Theorem 9 *When restricted to matchings of* $[2n]$ *with no neighbor crossings* ψ *is a bijection onto* $\mathcal{T}_n$*.*

We have now explained the hierarchy of nesting and crossing conditions that we set out to explain in the beginning of this section. As we pointed out, the bijections for the more general cases do not specialize to give bijections between the smaller sets. Indeed, if we specialize the map ψ to matchings with no nestings we get the subset of matrices $(t_{ij}) \in \mathcal{T}_n$ such that for all $i, j, x, y > 0$, at least one of $t_{i,j}$ and $t_{i-x,j+y}$ must be zero. The non-zero entries in such a matrix will form a "path" with the entries as vertices, which can be seen to be equivalent to a Motzkin path. Thus, the matrices just described are in bijection with Motzkin paths with positive integer weights on the vertices of the path such that the sum of the weights is n. If we on the other hand specialize ψ to matchings with no crossings we get the somewhat odd constraint that for all $i < i + x \leq j < j + y$ at least one of $t_{i,j}$ and $t_{i+x,j+y}$ must be zero.

Corollary 10 *The two subsets of* $\mathcal{T}_n$ *mentioned above are enumerated by the Catalan numbers.*

Before we close this section we give one more result that is almost for free given the map ψ. Let $\mathcal{T}_n^{01} \subset \mathcal{T}_n$ be the set of zero-one matrices in $\mathcal{T}_n$. For instance,

$$\mathcal{T}_3^{01} = \left\{ \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0&0\\0&1&0\\0&0&1\end{smallmatrix}\right) \right\}.$$

Dukes and Parviainen [6, §4] showed that the matrices in $\mathcal{T}_n^{01}$ correspond to those ascent sequences that have no two equal consecutive entries. We offer the following proposition.

Proposition 11 *When restricted to matchings of* $[2n]$ *with no left-nestings and no right-crossings* ψ *is a bijection onto* $\mathcal{T}_n^{01}$.

An important remark is that there exist a recent bijection by Chen et.al. [3], via certain walks in the Youngs lattice called vacillating tableaux, that uniformly shows all three cases above. For readers familiar with this bijection let us briefly describe why this is the case.

Let M be a matching of $[2n]$ in which j and $j+1$ are two consecutive closers. In the notation of [3], let $\phi(M) = (\lambda^0, \ldots, \lambda^{4n})$ be the vacillating tableau corresponding to M. The assumption that j and $j+1$ are closers means that $\lambda^{2j+2} \subsetneq \lambda^{2j}$ and $\lambda^{2j} \subsetneq \lambda^{2j-2}$. Let r_j be the row where λ^{2j+2} has fewer elements than λ^{2j} and let r_{j-1} be the row where λ^{2j} has fewer elements than λ^{2j-2}. The arcs ending in j and $j+1$ form a right-nesting if and only if $r_{j-1} \leq r_j$, and thus they form a right-crossing if and only if $r_{j-1} > r_j$.

Now, consider the involution M^* obtained by conjugating each partition in $\phi(M)$ and then applying ϕ^{-1}. A moment of thought gives that the arcs ending in j and $j+1$ form a right-nesting in M if and only if they form a right-crossing in M^*.

Similarly, if i and $i+1$ are two consecutive openers of M, then $\lambda^{2i-1} \subsetneq \lambda^{2i}$ and $\lambda^{2i} \subsetneq \lambda^{2i+2}$. This time let r_x be the row in which λ^{2x} is greater than λ^{2x-2}. Then the arcs with openers i and $i+1$ form a left-crossing if and only if $r_i < r_{i+1}$. Hence i and $i+1$ form a left-nesting in M if and only if they form a left-crossing in M^*.

This shows that the bijection in [3] may be used to explain all three levels discussed here at once. It also shows that using the above restrictions we get two different subets of all vacillating tableaux enumerated by $n!$ and one subset, satisfying both restrictions, that is enumerated by the Fishburn numbers.

6 Ascent and descent correcting sequences

It is easy to see that condition (1) in Proposition 6 is equivalent to

$$i >_P k \text{ and } i+1 \not>_P k \implies i = \mathrm{pred}(\ell) \text{ for some } \ell \text{ in } P. \tag{2}$$

Let a *descent correcting sequence* be an inversion table $(a_1, \ldots, a_n)$ such that

$$a_i > a_{i+1} \implies a_\ell = i \text{ for some } \ell > i.$$

That is, if there is a descent at position i then this has to be "corrected" by the value i occurring later in the sequence. Condition (2) translates directly to the condition for a descent correcting sequence, and thus we have the following Proposition (in which $\mathfrak{f}_n$ is the nth Fishburn number).

Proposition 12 *There are exactly* $\mathfrak{f}_n$ *descent correcting sequences of length* n.

We may similarly use the map f_{nc} from matchings with no left-crossings to inversion tables. We then get that the sequences corresponding to matchings with no neighbor crossings are the inversion tables $(a_1, \ldots, a_n)$ such that

$$a_i < a_{i+1} \neq i+1 \implies a_\ell = i \text{ for some } \ell > i.$$

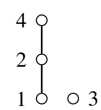
We call them *ascent correcting sequences*. Using Theorem 9 we arrive at the following result.

Proposition 13 *There are exactly* $\mathfrak{f}_n$ *ascent correcting sequences of length* n.

7 Posets that are both factorial and dually factorial

Note that being dually factorial entails the condition in Proposition 6. So, under h, matchings corresponding to dually factorial posets have no right-nestings. In fact, they do not have any nestings at all. To see this, assume that $M \in \mathcal{N}_n$ and let $\alpha_1, \ldots, \alpha_n$ are its arcs ordered by closer. Also, assume that $1 \le i < j \le n$. Recall that the arcs α_i and α_j form a nesting precisely when $\mathrm{pred}(i) > \mathrm{pred}(j)$, which is equivalent to there being a $k <_P i$ such that $k \not<_P j$; this cannot happen in a dually factorial poset. It is easy to see that this argument works both ways, so $M = h(P)$ is non-nesting if and only if P is dually factorial. It is well known that non-nesting matchings are counted by the Catalan numbers. See for instance Stanley [13, Ex. 6.19uu]. One way to associate a given non-nesting matching with a Dyck path is to map its openers to up-steps and its closers to down-steps.

Proposition 14 *Exactly $C_n = \binom{2n}{n}/(n+1)$ posets on $[n]$ are both factorial and dually factorial.*



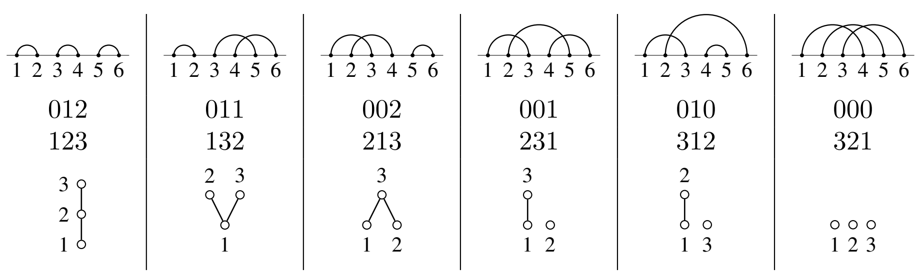
Let us mention an alternative proof. To the right is the smallest example of a factorial poset that is not dually factorial but satisfies the condition of Proposition 6: As stated by Proposition 4, factorial posets are $(\mathbf{2}+\mathbf{2})$-free; those that, in addition, are dually factorial are $(\mathbf{3}+\mathbf{1})$-free. However, we omit the details of this argument.

Proposition 15 *If P is a factorial poset satisfying* (1) *from Proposition 6, then P is dually factorial if and only if P is $(\mathbf{3}+\mathbf{1})$-free.*

Since posets that are both factorial and dually factorial have a unique labeling we can regard them as unlabeled. Further, unlabeled posets that are both $(\mathbf{2}+\mathbf{2})$- and $(\mathbf{3}+\mathbf{1})$-free (also called semiorders) are known to be enumerated by the Catalan numbers; see [13, Ex. 6.19ddd] and [15].

8 Statistics and equidistributions

One question we shall consider in this section is what statistics are respected by the bijections f, g and h. For reference, we list the size 3 matchings, inversion tables, permutations and posets that correspond to each other under those bijections:



There are several well known ways of translating between permutations and inversion tables. Here we have chosen the following way: Given $\pi \in \mathcal{S}_n$, we build the corresponding inversion table w from right to left. The right most letter of w is $\pi^{-1}(n) - 1$. The remaining letters of w are obtained by repeating this procedure on the length $n-1$ permutation that results from π by deleting n.

We shall now define the relevant statistics, and we start with statistics on posets. The ordinal sum [12, §3.2] of two posets P and Q is the poset $P \oplus Q$ on the union $P \cup Q$ such that $x \leq y$ in $P \oplus Q$ if $x \leq_P y$ or $x \leq_Q y$, or $x \in P$ and $y \in Q$. Let us say that P has k *components*, and write $\mathrm{comp}(P) = k$, if P is the ordinal sum of k, but not $k+1$, nonempty posets. The number of minimal elements of a poset P is denoted $\min(P)$. The number of levels of P—in other words, the number of distinct predecessor sets in P—is denoted $\mathrm{lev}(P)$. A pair of elements x and y in P are said to be *incomparable* if $x \not\leq_P y$ and $y \not\leq_P x$. The number of incomparable pairs in P we denote by $\mathrm{ip}(P)$.

Let π be a permutation. An *ascent* in π is a letter followed by a larger letter; a *descent* in π is a letter followed by a smaller letter. The number of ascents and descents are denoted $\mathrm{asc}(\pi)$ and $\mathrm{des}(\pi)$, respectively. An inversion is a pair $i < j$ such that $\pi(i) > \pi(j)$. The number of inversions is denoted $\mathrm{inv}(\pi)$. A *left-to-right minimum* of π is a letter with no smaller letter to the left of it; the number of left-to-right minima is denoted $\mathrm{lmin}(\pi)$. The statistics *right-to-left minima* (rmin), *left-to-right maxima* (lmax), and *right-to-left maxima* (rmax) are defined similarly. For permutations π and σ, let $\pi \oplus \sigma = \pi\sigma'$, where σ' is obtained from σ by adding $|\pi|$ to each of its letters, and juxtaposition denotes concatenation. We say that π has k components, and write $\mathrm{comp}(\pi) = k$, if π is the sum of k, but not $k+1$, nonempty permutations. Let $\mathrm{dent}(\pi)$ denote the number of distinct entries of the inversion table associated with π.

For M a matching on $[2m]$ and N a matching on $[2n]$, let $M \oplus N = M \cup N'$, where N' is the matching on $[2m+1, 2m+2n]$ obtained from N by adding $2m$ to all of its openers and closers. Let us say that M has k *components*, and write $\mathrm{comp}(M) = k$, if M is the sum of k, but not $k+1$, nonempty matchings. Let $\min(M) = j - 1$ where j is the smallest closer of M; for a factorial poset, j is the closer of the arc with opener 1. Let $\mathrm{last}(M)$ be the number of closers that are smaller than the opener of the last arc. Recall from Section 5 that $\mathrm{int}(M)$ denotes the number of intervals in the list of openers of M. Let us assume that k is the closer of some arc of M, and let $\alpha = (i, j)$ be another arc of M. If $i < k < j$ we say that k is *embraced* by α, and by $\mathrm{emb}(M)$ we denote the number of pairs (k, α) in M such that the closer k is embraced by α.

Proposition 16 *Let f and g be as in the proofs of Theorems 1 and 5. Let P be a factorial poset on $[n]$. Let $w = g(P)$ and $M = f(w)$ be the corresponding inversion table and matching, respectively. Let π be the permutation corresponding to w. Then*

$$\begin{array}{llllll}
(\ \mathrm{comp}(P), & \min(P), & \mathrm{pred}(n), & \mathrm{lev}(P), & \mathrm{ip}(P) &) \ = \\
(\ \mathrm{comp}(\pi), & \mathrm{lmin}(\pi), & \pi^{-1}(n) - 1, & \mathrm{dent}(\pi), & \mathrm{inv}(\pi) &) \ = \\
(\ \mathrm{comp}(M), & \min(M), & \mathrm{last}(M), & \mathrm{int}(M), & \mathrm{emb}(M) &)
\end{array}$$

Proof: For brevity the proof of this theorem has been excluded from this extended abstract. □

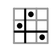
Let us note a few direct consequences of the above proposition.

Corollary 17 *The statistic* ip *is Mahonian on $\mathcal{F}_n$. That is, it has the same distribution as* inv *on $\mathcal{S}_n$. Also, the statistic* emb *is Mahonian on $\mathcal{N}_n$.*

Corollary 18 *The statistic* lev *is Eulerian on the set $\mathcal{F}_n$. That is, it has the same distribution as* des *on $\mathcal{S}_n$. Also, the statistic* int *is Eulerian on $\mathcal{N}_n$.*

Proof: It suffices to show that the statistic dent is Eulerian. The following proof is due to Emeric Deutsch (personal communication, May 2009). Let $d(n, k)$ be the number of inversion tables of length n with k

distinct entries. Clearly, $d(n,0) = 0$ for $n > 0$ and $d(n,k) = 0$ for $k > n$. We shall show that, for $0 < k \leq n$, $d(n,k) = kd(n-1,k) + (n-k+1)d(n-1,k-1)$ (the Eulerian recursion). Inversion tables of length n with k distinct entries fall into two disjoint classes: Those whose last entry is equal to at least one of the preceding $n-1$ entries; there are $kd(n-1,k)$ such inversion tables. Those whose last entry is different from the preceding $n-1$ entries; there are $(n-(k-1))d(n-1,k-1)$ such inversion tables. $\square$

Recall that $\mathrm{lne}(M)$ and $\mathrm{rne}(M)$ denote the number of left- and right-nestings, respectively. Let $\mathrm{lcr}(M)$ and $\mathrm{rcr}(M)$ denote the number of left- and right-crossings, respectively. The bijections $f : \mathcal{I}_n \to \mathcal{N}_n$ and $g : \mathcal{F}_n \to \mathcal{I}_n$ that we have presented do not specialize to the bijections presented by Bousquet-Mélou et al. [2]. If one were to find bijections that do specialize in the desired way, then one could also hope to prove the following conjecture (checked by computer for $n \leq 7$). Here we view $p =$  as a function counting the occurrences of the pattern p.

Conjecture 19 *These three triples of statistics are equidistributed.*

$$
\begin{array}{llllll}
(& \mathrm{rne}, & \mathrm{comp}, & \mathrm{min} &) & \textit{on } \mathcal{F}_n, \\
(& p, & \mathrm{comp}, & \mathrm{lmin} &) & \textit{on } \mathcal{S}_n, \\
(& \mathrm{rne}, & \mathrm{comp}, & \mathrm{min} &) & \textit{on } \mathcal{N}_n.
\end{array}
$$

9 Two additional conjectures and a generalization

Conjecture 20 *Assume that $i < j < k < \ell$. Let us say that the arcs (i,ℓ) and (j,k) are m-left-nesting if $j - i \leq m$. Note that a 1-left-nesting is the same as a left-nesting. This conjecture claims that among all the matchings on $[2n]$ there are exactly $\mathfrak{f}_n$ that have no 2-left-nestings.*

Conjecture 21 *The distribution of* lne *over the set of all matchings on $[2n]$ is given by the "Second-order Eulerian triangle", entry A008517 in OEIS [11].*

Conjectures 20 and 21 have been checked by computer for $n \leq 7$.

Problem 22 *Consider the following generalization of factorial posets. Let P and Q be labeled posets on $[n]$ such that $i <_P j \implies i <_Q j$. If, in addition,*

$$i <_Q j <_P k \implies i <_P k$$

*then we say that P is Q-*factorial. *Note that* $\mathbf{n}$*-factorial coincides with factorial, where* $\mathbf{n}$ *is the n-chain. Note also that Q itself is always a Q-factorial poset and it is the only one if Q is an antichain. Is this generalization useful? How many Q-factorial posets are there?*

Note added in proof: Paul Levande has found proofs for Conjectures 20 and 21.

Acknowledgment

Thanks to Mark Dukes for helpful suggestions on the presentation.

References

[1] K. P. Bogart, An obvious proof of Fishburn's interval order theorem, *Discrete Mathematics* 118 no. 1-3 (1993) 239–242.

[2] M. Bousquet-Mélou, A. Claesson, M. Dukes and S. Kitaev, $(2+2)$-free posets, ascent sequences and pattern avoiding permutations, *Journal of Combinatorial Theory Series A* 117 (2010) 884–909.

[3] W. Y. C. Chen, E. Y. P. Deng, R. R. X. Du, R. P. Stanley and C. H. Yan, Crossings and nestings of matchings and partitions, *Trans. Amer. Math. Soc.* 359 (2007) 1555-1575.

[4] A. de Médicis, X. G. Viennot, Moments des q-polynômes de Laguerre et la bijection de Foata-Zeilberger, *Advances in Applied Mathematics*, **15** (1994), 262–304.

[5] M. de Sainte-Catherine, Couplage et Pfaffiens en combinatoire, physique et informatique, Thèse du 3me cycle, Université de Bordeaux I, 1983.

[6] M. Dukes, R. Parviainen, Ascent sequences and upper triangular matrices containing non-negative integers, *Electronic Journal of Combinatorics* **17**(1) (2010), #R53 (16pp).

[7] P. C. Fishburn, *Interval Graphs and Interval Orders*, Wiley, New York, 1985.

[8] P. C. Fishburn, Intransitive indifference in preference theory: a survey, *Operational Research* **18** (1970) 207–208.

[9] P. C. Fishburn, Intransitive indifference with unequal indifference intervals, *Journal of Mathematical Psychology* **7** (1970) 144–149.

[10] A. Kasraoui, J. Zeng, Distribution of crossings, nestings and alignments of two edges in matchings and partitions, *Electronic Journal of Combinatorics* **13** no. 1 (2006) 12 pp.

[11] S. Plouffe and N. J. A. Sloane, The Encyclopedia of Integer Sequences, Academic Press Inc., San Diego, 1995. Electronic version available at `www.research.att.com/~njas/sequences/`.

[12] R. P. Stanley, *Enumerative combinatorics Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, 1997.

[13] R. P. Stanley, *Enumerative combinatorics Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, 1999.

[14] A. Stoimenow, Enumeration of chord diagrams and an upper bound for Vassiliev invariants, *J. Knot Theory Ramifications* **7** no. 1 (1998) 93–114.

[15] R.L. Wine, J.E. Freund, On the enumeration of decision patterns involving n means, Annals of Mathematical Statistics **28** (1957) 256–259.

[16] D. Zagier, Vassiliev invariants and a strange identity related to the Dedeking eta-function, *Topology*, **40** (2001) 945–960.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 522–533

The Frobenius Complex

Eric Clark and Richard Ehrenborg

Department of Mathematics, University of Kentucky, Lexington, KY 40506-0027

Abstract. Motivated by the classical Frobenius problem, we introduce the Frobenius poset on the integers $\mathbb{Z}$, that is, for a sub-semigroup Λ of the non-negative integers $(\mathbb{N}, +)$, we define the order by $n \leq_\Lambda m$ if $m - n \in \Lambda$. When Λ is generated by two relatively prime integers a and b, we show that the order complex of an interval in the Frobenius poset is either contractible or homotopy equivalent to a sphere. We also show that when Λ is generated by the integers $\{a, a+d, a+2d, \dots, a+(a-1)d\}$, the order complex is homotopy equivalent to a wedge of spheres.

Résumé. Motivé par le problème de Frobenius classique, nous introduisons l'ensemble partiellement ordonné de Frobenius sur les entiers $\mathbb{Z}$, c.à.d. que pour un sous-semigroupe Λ de les entiers non-négatifs $(\mathbb{N}, +)$ nous définissons l'ordre par $n \leq_\Lambda m$ si $m - n \in \Lambda$. Quand le Λ est engendré par deux nombres a et b, relativement premiers entre eux, noux montrons que le complexe des chaînes d'un intervalle quelquonque dans l'ensemble partiellement ordonné de Frobenius est soit contractible soit homotopiquement équivalent à une sphère. Nous montrons aussi que dans le cas où Λ est engendré par les entiers $\{a, a+d, a+2d, \dots, a+(a-1)d\}$, le complexe des chaînes a le type de homotopie d'un bouquet de sphères.

Keywords: order complex, homotopy type, Morse matching, cylindrical posets

1 Introduction

The classical Frobenius problem is to find the largest integer for which change cannot be made using coins with the relatively prime denominations $a_1, a_2, \dots, a_d$; see for instance [2, Section 1.2]. We will reformulate this question by introducing the following poset.

Let Λ be a sub-semigroup of the non-negative integers $\mathbb{N}$, that is, Λ is closed under addition and the element 0 lies in Λ. We define the Frobenius poset $P = (\mathbb{Z}, \leq_\Lambda)$ on the integers $\mathbb{Z}$ by the order relation $n \leq_\Lambda m$ if $m - n \in \Lambda$. We denote by $[n, m]_\Lambda$ the interval from n to m in the Frobenius poset, that is,

$$[n, m]_\Lambda = \{i \in [n, m] \;:\; i - n, m - i \in \Lambda\}.$$

Observe that the interval $[n, m]_\Lambda$ in the Frobenius poset is isomorphic to the interval $[n+i, m+i]_\Lambda$, that is, the interval $[n, m]_\Lambda$ only depends on the difference $m - n$. Also note that each interval is self-dual by sending i in $[0, n]_\Lambda$ to $n - i$.

In this form, the original Frobenius problem would be to find the largest integer n that is not comparable to zero in the Frobenius poset when Λ is generated by $\{a_1, a_2, \dots, a_d\}$. The largest such integer is known as the Frobenius number. In general, calculating the Frobenius number is difficult. However, in the case where the semigroup is generated by two relatively prime integers a and b, it is well known that

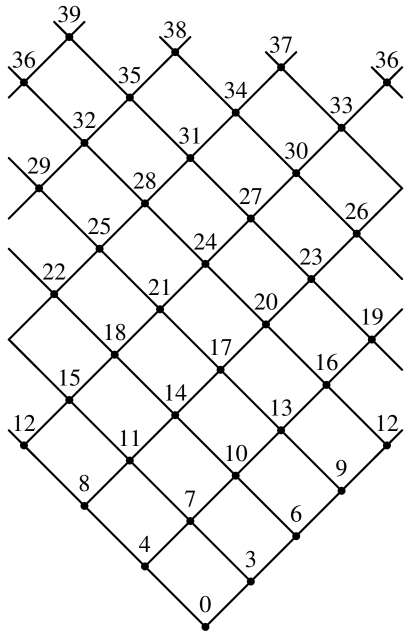


Fig. 1: The filter generated by 0 in the Frobenius poset corresponding to the semigroup Λ generated by $a = 3$ and $b = 4$, that is, $\Lambda = \mathbb{N} - \{1, 2, 5\}$. Note that you get a better picture by rolling the page into a cylinder.

the Frobenius number is given by $ab - a - b$. Also, when the semigroup is generated by the arithmetic sequence $\{a, a + d, \ldots, a + sd\}$, the Frobenius number was shown by Roberts [17] to be

$$\left(\left\lfloor \frac{a-2}{s} \right\rfloor + 1\right) \cdot a + (d-1)(a-1) - 1. \tag{1}$$

We study the topology of the order complex of intervals of this poset in the two generator case and when the generators form an arithmetic sequence where $s = a - 1$.

The technique we use is discrete Morse theory which was developed by Forman [8, 9]. Thus we construct an acyclic partial matching on the face poset of the order complex by using the Patchwork Theorem. We then identify the unmatched, or critical, cells. These tell us the number and dimension of cells in a CW-complex to which our order complex is homotopy equivalent. Using extra structure about the critical cells, we can determine exactly what the homotopy type is.

A more general situation is to consider a semigroup Λ of $\mathbb{N}^d$ and define a partial order on $\mathbb{Z}^d$ by $\mu \leq_\Lambda \lambda$ if $\lambda - \mu \in \Lambda$. Define the semigroup algebra $k[\Lambda]$ as the linear span of the monomials whose powers belong to Λ, that is, $k[\Lambda] = \mathrm{span}\{x^\lambda = x_1^{\lambda_1} \cdots x_d^{\lambda_d} : \lambda \in \Lambda\}$. Laudal and Sletsjøe [14] makes the connection between the homology of the order complex of intervals in this partial order and the semigroup algebra $k[\Lambda]$.

Theorem 1.1 (Laudal and Sletsjøe) *For Λ a sub-semigroup of $\mathbb{N}^d$ with the associated monoid Λ, the following equality holds*

$$\dim_k \mathrm{Tor}_i^{k[\Lambda]}(k, k)_\lambda = \dim_k \widetilde{H}_{i-2}(\Delta([0, \lambda]_\Lambda), k),$$

for all $\lambda \in \Lambda$ *and* $i \geq 0$.

The papers [5, 10, 16] continue to study the topology of the intervals in this partial order. Hersh and Welker [10] give bounds on the indices of the non-vanishing homology groups of the order complex of the intervals. Peeva, Reiner, and Sturmfels [16] show that the semigroup ring $k[\Lambda]$ is Koszul if and only if each interval in Λ is Cohen-Macaulay.

We end this paper with some open questions and concluding remarks.

2 Discrete Morse theory

We recall the following definitions and theorems from discrete Morse theory. See [8, 9, 12] for further details.

Definition 2.1 *A* partial matching *in a poset* P *is a partial matching in the underlying graph of the Hasse diagram of* P*, that is, a subset* $M \subseteq P \times P$ *such that* $(x, y) \in M$ *implies* $x \prec y$ *and each* $x \in P$ *belongs to at most one element of* M*. For* $(x, y) \in M$ *we write* $x = d(y)$ *and* $y = u(x)$*, where* d *and* u *stand for down and up, respectively.*

Definition 2.2 *A partial matching* M *on* P *is* acyclic *if there does not exist a cycle*

$$z_1 \succ d(z_1) \prec z_2 \succ d(z_2) \prec \cdots \prec z_n \succ d(z_n) \prec z_1,$$

in P *with* $n \geq 2$*, and all* $z_i \in P$ *distinct. Given a partial matching, the unmatched elements are called* critical*. If there are no critical elements, the acyclic matching is* perfect*.*

We now state the main result from discrete Morse theory. For a simplicial complex Δ, let $\mathcal{F}(\Delta)$ denote the poset of faces of Δ ordered by inclusion.

Theorem 2.3 *Let* Δ *be a simplicial complex. If* M *is an acyclic matching on* $\mathcal{F}(\Delta) - \{\widehat{0}\}$ *and* c_i *denotes the number of critical* i*-dimensional cells of* Δ*, then the complex* Δ *is homotopy equivalent to a CW complex* Δ_c *which has* c_i *cells of dimension* i*.*

For us it will be convenient to work with the reduced discrete Morse theory, that is, we include the empty set.

Corollary 2.4 *Let* Δ *be a simplicial complex and let* M *be an acyclic matching on* $\mathcal{F}(\Delta)$*. Then the space* Δ *is homotopy equivalent to a CW complex* Δ_c *which has* $c_0 + 1$ *cells of dimension* 0 *and* c_i *cells of dimension* i *for* $i > 0$*.*

In particular, if the matching from Corollary 2.4 is perfect, then Δ_c is contractible. Also, if the matching has exactly one critical cell then Δ_c is a combinatorial d-sphere where d is the dimension of the cell.

Given a set of critical cells of differing dimension, in general it is impossible to conclude that the CW complex Δ_c is homotopy equivalent to a wedge of spheres. See Kozlov [13] for an example. However, in certain cases, this is possible.

Theorem 2.5 *Let* M *be a Morse matching on* $\mathcal{F}(\Delta)$ *such that all* c_i *critical cells of dimension* i *are maximal. Then*

$$\Delta \simeq \bigvee_{i} \bigvee_{j=1}^{c_i} \mathbb{S}^i.$$

Proof: By the above statement, the complex Δ without the critical cells is contractible. In particular, the boundary of each of the critical cells contracts to a point. Since all of the critical cells are maximal, they can be independently added back into the complex. $\square$

Kozlov [13] gives a more general sufficient condition on an acyclic Morse matching for the complex to be homotopy equivalent to a wedge of spheres enumerated by the critical cells.

We are interested in finding an acyclic matching on the face poset of the Frobenius complex. The Patchwork Theorem [12] gives us a way of constructing one.

Theorem 2.6 *Assume that $\varphi : P \rightarrow Q$ is an order-preserving poset map, and assume that there are acyclic matchings on the fibers $\varphi^{-1}(q)$ for all $q \in Q$. Then the union of these matchings is itself an acyclic matching on P.*

3 Two generators

With two generators, the associated Frobenius poset can be embedded on a cylinder. By Bezout's identity there are two integers p and q such that $p \cdot a + q \cdot b = 1$. Define a group morphism $\gamma : \mathbb{Z} \longrightarrow \mathbb{Z}_{2ab} \times \mathbb{Z}$ by $\gamma(x) = ((p \cdot a - q \cdot b) \cdot x, x)$, that is, the first coordinate is modulo $2 \cdot a \cdot b$ which corresponds to encircling the cylinder. Observe that $\gamma(a) = ((p \cdot a - q \cdot b) \cdot a, a) = ((p \cdot a + q \cdot b) \cdot a, a) = (a, a)$ and $\gamma(b) = ((p \cdot a - q \cdot b) \cdot b, b) = ((-p \cdot a - q \cdot b) \cdot b, b) = (-b, b)$. Hence the two cover relations $x \prec x + a$ and $x \prec x + b$ in the Frobenius poset translates to $\gamma(x) + (a, a) = \gamma(x + a)$ and $\gamma(x) + (-b, b) = \gamma(x + b)$. In other words, to take an a step we make the step (a, a) on the cylinder and a b step corresponds to the step $(-b, b)$. As an example, see Figure 1 where $a = 3$ and $b = 4$.

In general, the Frobenius poset is not a lattice. When Λ is generated by two relatively prime integers a and b, we have the four relations $a <_\Lambda a + b$, $b <_\Lambda a + b$, $a <_\Lambda ab$, and $b <_\Lambda ab$. However, since $ab - a - b$ is the Frobenius number we have $a + b \not\leq_\Lambda ab$, showing that the poset is not a lattice. In Figure 1, we see that 3 and 4 are both lower bounds for 7 and 12.

Let $c_k(n)$ denote the number of chains in the Frobenius interval $[0, n]_\Lambda$ of length k. Using multiplication of generating functions, we have

$$\sum_{n \geq k} c_k(n) \cdot q^n = \left(\sum_{n \geq 1} c_1(n) \cdot q^n \right)^k .$$

By taking the alternating sum over k and using Philip Hall's expression for the Möbius function, we have

$$\sum_{n \geq 0} \mu(n) \cdot q^n = \frac{1}{1 + \sum_{n \geq 1} c_1(n) \cdot q^n}, \tag{2}$$

where $\mu(n)$ denotes the Möbius function of the interval $[0, n]_\Lambda$. Now assuming that Λ is generated by two relatively prime positive integers a and b, we have that

$$1 + \sum_{n \geq 1} c_1(n) \cdot q^n = \frac{1 - q^{ab}}{(1 - q^a) \cdot (1 - q^b)};$$

see [1, Exercise VIII.1.5]. Hence the Möbius function is given by

$$\begin{aligned}\sum_{n\geq 0}\mu(n)\cdot q^n &= \frac{(1-q^a)\cdot(1-q^b)}{1-q^{ab}} \\ &= 1-q^a-q^b+q^{a+b}+q^{ab}-q^{ab+a}-q^{ab+b}+q^{ab+a+b}+\cdots.\end{aligned}$$

Note that the coefficients are all ± 1 or 0. We will consider this fact in a topological setting. Recall that the *order complex* $\Delta(P)$ of a bounded poset P is the collection of chains in P, that is,

$$\Delta(P) = \{\{x_1, x_2, \ldots, x_k\} : \widehat{0} < x_1 < x_2 < \cdots < x_k < \widehat{1}\}$$

ordered by inclusion. Also, the reduced Euler characteristic of the order complex $\Delta(P)$ is given by the Möbius function of P. We call the order complex of the face poset of a Frobenius interval the *Frobenius complex*. We wish to study the homotopy type of the Frobenius complex. Since the reduced Euler characteristic of the Frobenius complex takes on the values $+1$, -1, or 0, we are lead to the following main theorem.

Theorem 3.1 *Let the sub-semigroup Λ be generated by two relatively prime positive integers a and b with $1 < a < b$. The order complex of the associated Frobenius interval $[0,n]_\Lambda$, for $n \geq 1$, is homotopy equivalent to either a sphere or contractible, according to*

$$\Delta([0,n]_\Lambda) \simeq \begin{cases} \mathbb{S}^{2n/ab-2} & \text{if } n \equiv 0 \bmod a\cdot b, \\ \mathbb{S}^{2(n-a)/ab-1} & \text{if } n \equiv a \bmod a\cdot b, \\ \mathbb{S}^{2(n-b)/ab-1} & \text{if } n \equiv b \bmod a\cdot b, \\ \mathbb{S}^{2(n-a-b)/ab} & \text{if } n \equiv a+b \bmod a\cdot b, \\ \textit{point} & \textit{otherwise.} \end{cases}$$

Observe that if n does not belong to the sub-semigroup Λ then we consider the order complex $\Delta([0,n]_\Lambda)$ to be the empty set which we view as contractible. This is distinct from the case when n equals a or b, that is, when the order complex $\Delta([0,n]_\Lambda)$ only contains the empty set. In this case, we view this as a sphere of dimension -1.

In the case where the two generators are 2 and 3, the semigroup is $\mathbb{N} - \{1\}$ and the order complex $\Delta([0,n]_\Lambda)$ consists of all subsets of the interval $[2, n-2]$ that do not contain two consecutive integers. This is known as the complex of sparse subsets. Its homotopy type was first determined by Kozlov [11]. See also [7] where it appears as the independence complex of a path. Billera and Myers [4] showed this complex is non-pure shellable.

As a corollary to Theorem 1.1, we obtain

Corollary 3.2 *Let a and b be relatively prime integers such that $1 < a < b$. Let R denote the ring $k[y,z]/(y^b - z^a)$. Then the multigraded Poincaré series*

$$P_k^R(t,q) = \sum_{n\in\Lambda}\sum_{i\geq 0}\dim_k\left(\mathrm{Tor}_i^R(k,k)_n\right)t^i q^n$$

is given by the rational function

$$\frac{1+tq^a+tq^b+t^2q^{a+b}}{1-t^2q^{ab}}.$$

Proof: Let Λ the semigroup generated by a and b. Observe that the ring R is isomorphic to the semigroup ring $k[\Lambda]$. By combining Theorems 1.1 and 3.1 the multigraded Poincaré series is given by

$$P_k^R(t,q) = 1 + tq^a + tq^b + t^2q^{a+b} + t^2q^{ab} + t^3q^{ab+a} + t^3q^{ab+b} + t^4q^{ab+a+b} + \cdots ,$$

which is the sought after rational generating function. $\square$

We now turn our attention to the proof of Theorem 3.1. Let Λ be generated by two relatively prime positive integers a and b with $1 < a < b$. Consider the Frobenius interval $[0,n]_\Lambda$. Define the three sets B_ℓ, C_ℓ and D_ℓ as follows:

$$\begin{aligned} B_\ell &= \{\ell ab + 2b, \ell ab + 3b, \ldots, \ell ab + (a-1)b\}, \\ C_\ell &= \{b, ab, ab+b, 2ab, 2ab+b, 3ab, \ldots, (\ell-1)ab + b, \ell ab\}, \\ D_\ell &= C_\ell \cup \{\ell ab + b\}. \end{aligned}$$

Note that $C_0 = \emptyset$, $D_0 = \{b\}$, and $C_{\ell+1} = D_\ell \cup \{(\ell+1)ab\}$.

Let Q be the infinite chain $\{a < a+b < ab+a < ab+a+b < 2ab+a < \cdots\}$ adjoined with a new maximal element $\widehat{1}_Q$, that is,

$$Q = \{m \in \mathbb{N} \;:\; m \equiv a, a+b \bmod ab\} \cup \{\widehat{1}_Q\}.$$

We now define a map φ from the face poset of the order complex $\Delta([0,n]_\Lambda)$ to the poset Q. We will later show that φ is an order-preserving poset map with natural matchings on the fibers. Let φ be defined by

$$\varphi(x) = \begin{cases} \ell ab + a & \text{if } \ell ab + a <_\Lambda n, \\ & C_\ell \subseteq x, \\ & B_t \cap x = \emptyset \text{ for } 0 \le t \le \ell, \\ & \text{and } \ell ab + b \notin x; \\ \ell ab + a + b & \text{if } \ell ab + a + b <_\Lambda n, \\ & D_\ell \subseteq x, \\ & B_t \cap x = \emptyset \text{ for } 0 \le t \le \ell, \\ & \text{and } \ell ab + ab \notin x; \\ \widehat{1}_Q & \text{otherwise.} \end{cases}$$

In order to make acyclic pairings on the fibers of φ, it will be useful to have a description of the chains that are mapped to the maximal element $\widehat{1}_Q$ and their structure. Let Γ denote this collection of chains in the Frobenius poset, that is, $\Gamma = \varphi^{-1}(\widehat{1}_Q)$.

Lemma 3.3 *The collection Γ consists of the chains x that satisfy one of the following four conditions:*

1. *There exists a non-negative integer λ such that $C_\lambda \subseteq x$, $\lambda ab + b \notin x$, $B_\lambda \cap x \neq \emptyset$, and $B_t \cap x = \emptyset$ for $0 \le t \le \lambda - 1$.*

2. *There exists a non-negative integer λ such that $D_\lambda \subseteq x$, $B_\lambda \cap x \neq \emptyset$, and $B_t \cap x = \emptyset$ for $0 \le t \le \lambda - 1$.*

3. *There exists a non-negative integer λ such that $x = C_\lambda$ and $\lambda ab + a \not<_\Lambda n$.*

4. *There exists a non-negative integer λ such that $x = D_\lambda$ and $\lambda ab + a + b \not<_\Lambda n$.*

We will refer to the condition met by a chain as its *type* and the associated λ as its *parameter*. The structure of Γ is given in the following lemma.

Lemma 3.4 *The following four conditions hold for the collection Γ.*

(i) Let x be a chain of type 1 with parameter λ in Γ. Then $x \cup \{\lambda ab + b\}$ is a chain in Γ of type 2 with the same parameter λ.

(ii) Let y be a chain of type 2 with parameter λ in Γ. Then $y - \{\lambda ab + b\}$ is a chain in Γ of type 1 with the same parameter λ.

(iii) Let x be a chain of type 1 with parameter λ and y be a chain of type 2 with parameter μ such that $x \prec y$. Then $\lambda \geq \mu$ holds with equality if and only if $y = x \cup \{\lambda ab + b\}$.

(iv) If z is an element of type 4, then z does not cover any element of type 1 or 2.

We now turn our attention to the map φ.

Lemma 3.5 *The map $\varphi : \mathcal{F}(\Delta([0, n]_\Lambda)) \longrightarrow Q$ is an order-preserving poset map.*

Lemma 3.6 *For $m <_Q \widehat{1}_Q$, the collection $\{(x, x \cup \{m\}) : m \notin x \in \varphi^{-1}(m)\}$ is a perfect acyclic matching on the fiber $\varphi^{-1}(m)$.*

Thus we have reduced the problem to finding an acyclic matching on the fiber $\Gamma = \varphi^{-1}(\widehat{1}_Q)$.

Lemma 3.7 *The collection $\{(x, x \cup \{\lambda ab + b\}) : x$ is a chain of type 1 with parameter $\lambda\}$ is an acyclic matching on Γ where the critical cells are the chains of type 3 and 4.*

Proof: We have seen from parts (i) and (ii) of Lemma 3.4 that to every element x of type 1 there exists a corresponding element y of type 2 with the same parameter and vice-versa. In other words, this is a perfect matching on chains of type 1 and 2. Chains of type 3 and 4 are left unmatched.

We must now show that this matching is acyclic, that is, a directed cycle of the form described in Definition 2.2 cannot exist. Let z_1 be a chain of type 2 with parameter λ. Then $d(z_1) = z_1 - \{\lambda ab + b\}$ is an element of type 1 with the same λ. Part (iii) of Lemma 3.4 tells us that any z_2 different from z_1 will have a smaller parameter. Therefore, we cannot return to z_1 using our matching. Hence the matching is acyclic. □

Lemma 3.8 *Let $n = kab + r$ for $0 \leq r < ab$. If $r = 0, a, b,$ or $a + b$, then the matching given in Lemma 3.7 has exactly one critical cell. If $r = jb$ for $2 \leq j \leq a - 1$, there are exactly two unmatched chains of Γ. Otherwise, there are no critical cells in Γ. More precisely, the critical cells of Γ are given by*

$$\begin{cases} \{D_{k-1}\} & \text{if } n = kab, \\ \{C_k\} & \text{if } n = kab + a, \\ \{C_k\} & \text{if } n = kab + b, \\ \{D_k\} & \text{if } n = kab + a + b, \\ \{C_k, D_k\} & \text{if } n = kab + ib, 2 \leq i \leq a - 1, \\ \emptyset & \text{otherwise.} \end{cases}$$

Proof: The only elements of Γ that were not matched are those of type 3 and 4 in Lemma 3.3. Thus, we need to determine the number of type 3 and 4 elements in Γ, that is, the number of integers λ such that $\lambda ab <_\Lambda n$ and $\lambda ab + a \not<_\Lambda$ (type 3) and integers λ such that $\lambda ab + b <_\Lambda n$ and $\lambda ab + a + b \not<_\Lambda n$ (type 4).

Using the Frobenius number, we know that every integer smaller than $n - (ab - a - b) = (k-1)ab + a + b + r$ is comparable with n with respect to the order $<_\Lambda$. We do not need to check $\ell ab + a$ or $\ell ab + a + b$ for $0 \leq \ell < k$ because this number is always comparable to n (unless $r = 0$ when we must check $(k-1)ab + a + b$). We also do not need to consider $\ell ab + a$ or $\ell ab + a + b$ for $\ell \geq k+1$ because we would have $\ell ab + a, \ell ab, \ell ab + a + b$, and $\ell ab + b$ all not contained in $[0, n]_\Lambda$. Thus, we only need to check $(k-1)ab + a + b$ (if $r = 0$), $kab + a$, and $kab + a + b$

There are nine cases to consider.

– $r \notin \Lambda$. Then we have both $kab+a \not<_\Lambda n$ and $kab \not<_\Lambda n$, and also $kab+a+b \not<_\Lambda$ and $kab+b \not<_\Lambda n$. Therefore, there are no critical cells.

Otherwise, r belongs to the semigroup Λ and we can write $r = ia + jb$, where i and j are unique non-negative integers.

– $(i, j) = (0, 0)$. We see that $kab+a \not<_\Lambda n$, but also $kab \not<_\Lambda n$. Similarly, we have $kab+a+b \not<_\Lambda n$ and $kab+b \not<_\Lambda n$. Finally, we check and see that $(k-1)ab+a+b \not<_\Lambda n$ because $kab-(k-1)ab+a+b = ab-a-b \notin \Lambda$. Also $(k-1)ab+b <_\Lambda n$ because $kab-(k-1)ab+b = ab-b = (a-1)b \in \Lambda$. Thus we have the one critical cell D_{k-1}.

– $(i, j) = (1, 0)$. We can easily see that $kab + a + b \not<_\Lambda n$ and $kab + b \not<_\Lambda n$. However, we have $kab + a \not<_\Lambda n$ while $kab <_\Lambda n$. Therefore, we have the one critical cell C_k.

– $(i, j) = (0, 1)$. In this case we again see that $kab + a + b \not<_\Lambda n$ and $kab + b \not<_\Lambda n$. However, we still have $kab + a \not<_\Lambda n$, while $kab <_\Lambda n$. Thus we have the one critical cell C_k.

– $(i, j) = (1, 1)$. First we note that $kab + a <_\Lambda n$. Thus we only check to see that $kab + a + b \not<_\Lambda n$ and $kab + b <_\Lambda n$. This is easily true, so there is one critical cell D_k.

– $i = 0, 2 \leq j \leq a - 1$. Clearly $kab + a \not<_\Lambda n$ while $kab <_\Lambda n$. Also, we see that $kab + a + b \not<_\Lambda n$ while $kab + b <_\Lambda n$. Thus the unmatched cells are C_k and D_k.

– $i \geq 1, j \geq 2$. Both $kab + a$ and $kab + a + b$ are both comparable with n. Therefore there are no critical cells.

– $i \geq 2, j = 0$. We see that $kab + a$ is comparable with n. Also, both $kab + a + b$ and $kab + b$ are not comparable with n. Therefore there are no critical cells.

– $i \geq 2, j = 1$. Then $kab + a$ and $kab + a + b$ are both comparable with n. Therefore there are no critical cells.

□

Proof of Theorem 3.1.: By applying the Patchwork Theorem to the function φ we see that the homotopy type of $\Delta([0, n]_\Lambda)$ depends only on the fiber $\varphi^{-1}(\widehat{1}_Q) = \Gamma$. Applying Lemmas 3.7 and 3.8, there is only one critical cell when $n \equiv 0, a, b, a + b \mod ab$ and no critical cells in every other case except when

$i = 0$ and $2 \leq j \leq a - 1$. However, we claim in this last case we can add the pair (C_k, D_k) to the matching on Γ and still be left with an acyclic matching.

Lemma 3.4 (iv) shows that a chain of type 4 does not cover any chain of type 1. Hence, when adding the edge (C_k, D_k) to the Morse matching of Γ, it will not create any directed cycles through the chain D_k. Hence the matching is still acyclic and there are no critical cells in this case.

The critical cells for $n \equiv 0, a, b, a+b \mod ab$ can be easily seen to be of dimension $2n/ab - 2, 2(n - a)/ab - 1, 2(n-b)/ab - 1$, and $2(n-a-b)/ab$, respectively. Therefore, applying the main theorem of reduced discrete Morse theory, Corollary 2.4, proves the result. $\square$

4 Generators in an arithmetic sequence

Recall that the q-analogue is defined as follows: $[a]_{q^d} = 1 + q^d + (q^d)^2 + \cdots + (q^d)^{a-1}$.

Theorem 4.1 *Let Λ be the sub-semigroup generated by the integers $\{a, a+d, a+2d, \ldots, a+(a-1)d\}$ where a and d are relatively prime. The order complex of the associated Frobenius interval $[0, n]_\Lambda$ is homotopy equivalent to a wedge of spheres where the ith Betti number satisfies*

$$\sum_{n \geq 0} \widetilde{\beta}_i q^n = q^{a+(i+1)(a+d)} \cdot [a]_{q^d} \cdot [a-1]_{q^d}^{i+1}.$$

Example 4.2 For the generators $\{4, 5, 6, 7\}$, that is $a = 4$ and $d = 1$, we have

$$\begin{aligned}
\sum_{n \geq 0} \widetilde{\beta}_1 q^n &= q^{14} \cdot [4] \cdot [3]^2 = q^{14} + \cdots + 3q^{20} + q^{21}, \\
\sum_{n \geq 0} \widetilde{\beta}_2 q^n &= q^{19} \cdot [4] \cdot [3]^3 = q^{19} + 4q^{20} + \cdots + q^{28},
\end{aligned}$$

and no other generating polynomial contains the q^{20} term. Hence the Frobenius complex $\Delta([0, 20]_\Lambda)$ is homotopy equivalent to a wedge of three circles and four 2-spheres.

The Frobenius number of an arithmetic sequence was given in equation (1). Therefore, for the generators $\{a, a+d, a+2d, \ldots, a+(a-1)d\}$, we have the Frobenius number

$$\left(\left\lfloor \frac{a-2}{a-1} \right\rfloor + 1 \right) \cdot a + (d-1)(a-1) - 1 \quad = \quad (a-1)d.$$

We will proceed as before and use Discrete Morse theory and the Patchwork theorem. Let A be the set $\{a+d, a+2d, \ldots, a+(a-1)d\}$.

Definition 4.3 *Given n, let R be the chain $\{1, 2, 3, 4, \ldots, n-a\}$ with a maximal element $\widehat{1}_R$ adjoined. That is,*

$$R = \{1, 2, 3, \ldots, n-a\} \cup \{\widehat{1}_R\}.$$

If $x = \{x_1, x_2, \ldots, x_k\} \in \Delta([0,n]_\Lambda)$ *and we define* $x_0 = 0$*, let* $\psi : \mathcal{F}(\Delta([0,n]_\Lambda)) \to R$ *be a map defined by*

$$\psi(x) = \begin{cases} x_{i-1} + a, & x_i - x_{i-1} \notin A, \\ & x_j - x_{j-1} \in A \\ & \text{for } 1 \le j \le i-1, \\ x_k + a, & n - x_k \notin \{a\} \cup A, \\ & x_j - x_{j-1} \in A \\ & \text{for } 1 \le j \le k, \\ \widehat{1}_R, & \text{otherwise.} \end{cases}$$

Lemma 4.4 *The element* $m \cdot d$ *is not contained in* Λ *for* $1 \le m \le (a-1)$.

Proof: Suppose $m \cdot d \in \Lambda$ for $1 \le m \le (a-1)$. Then for $0 \le s_i \le (a-1)$, we have

$$\begin{aligned} m \cdot d &= (a + s_1 d) + (a + s_2 d) + \cdots + (a + s_k d) \\ &= ka + (s_1 + \cdots + s_k)d. \end{aligned}$$

The fact that d and a are relatively prime implies that d divides k. That is, $k = \ell \cdot d$. Therefore, $m = \ell \cdot a + (s_1 + \cdots + s_k)$ which implies that $m \ge a$. This is a contradiction. □

Lemma 4.5 *Let* x_i *and* x_j *be elements of a chain* x *such that* $x_i - x_j \in \{a\} \cup A$*. Then the open interval* $(x_i, x_j)_\Lambda$ *is empty.*

The following lemma is an immediate consequence of Lemma 4.5 and the definition of the function ψ.

Lemma 4.6 *If* x *and* y *are chains such that* $x \subseteq y$ *and* $\psi(x) = x_{i-1} + a$ *then* $x_j = y_j$ *for* $1 \le j \le i-1$*. In particular, if* $\psi(x) = \widehat{1}_R$ *then* $x = y$*.*

We can finally give a few properties of the map ψ.

Lemma 4.7 *The map* $\psi : \mathcal{F}(\Delta([0,n]_\Lambda)) \to R$ *is an order preserving poset map.*

Lemma 4.8 *For* $m <_R \widehat{1}_R$*, the collection* $\{(x, x \cup \{m\}) : m \notin x \in \psi^{-1}(m)\}$ *is a perfect acyclic matching on the fiber* $\psi^{-1}(m)$*.*

Proof: Suppose $\psi(x) = x_{i-1} + a$ and $x_{i-1} + a \in x$. That is, $x_i = x_{i-1} + a$. It is clear that $d(x) = x - \{x_i\}$ is a valid chain in the Frobenius complex since we are simply removing an element. We need to check that $\psi(d(x)) = x_{i-1} + a$. We know that $d(x)_j - d(x)_{j-1} = x_j - x_{j-1} \in A$ for $1 \le j \le (i-1)$. Suppose $d(x)_i - d(x)_{i-1} = x_{i+1} - x_{i-1} \in A$. Then, by Lemma 4.5, $(x_{i-1}, x_{i+1})_\Lambda$ would have to be empty. This contradicts the fact that $x_i \in (x_{i-1}, x_{i+1})_\Lambda$ in the chain x. Since $d(x)_i - d(x)_{i-1} \notin A$ and $d(x)_j - d(x)_{j-1} \in A$ for $1 \le j \le (i-1)$, we have $\psi(d(x)) = d(x)_{i-1} + a = x_{i-1} + a$.

Now suppose that $\psi(x) = x_{i-1} + a$ and $x_{i-1} + a \notin x$. It is clear that $u(x) = x \cup \{x_{i-1} + a\}$ would be mapped to $x_{i-1} + a$. Thus, it must be shown that $u(x)$ is a valid chain, that is, $x_{i-1} + a$ is comparable to x_i. We know that $x_i - x_{i-1} \notin A$. Suppose

$$x_i - x_{i-1} = (a + s_1 d) + (a + s_2 d) + \cdots + (a + s_k d)$$

where $s_1 \leq s_2 \leq \cdots \leq s_k$ and $k \geq 2$. Then

$$x_i - (x_{i-1} + a) = (a + (s_1 + s_2)d) + (a + s_3 d) + \cdots + (a + s_k d).$$

If $s_1 + s_2 \leq a - 1$, then we have written this difference as a sum of generators. Therefore, x_i and $x_{i-1} + a$ are comparable.

If $s_1 + s_2 > a - 1$, then the difference is larger than $(a-1)d$, which is the Frobenius number of the generators. Thus, x_i and $x_{i-1} + a$ are comparable. Therefore, $u(x)$ is a valid chain.

Finally, the matching on the fiber is clearly acyclic since the same element is either added or removed from a chain. □

Using the Patchwork theorem, we have an acyclic matching on $\mathcal{F}(\Delta([0,n]_\Lambda))$ whose only critical cells are the elements of the fiber $\psi^{-1}(\widehat{1}_R)$. Note that Lemma 4.6 says that each of these cells are maximal. Therefore, due to Theorem 2.5, we will have that $\Delta([0,n]_\Lambda$ is homotopy equivalent to a wedge of spheres whose number and dimension corresponds to the number and dimension of the critical cells. Thus, we are interested in counting the number of chains that are mapped to $\widehat{1}_R$. The following lemma is straightforward from the definition of the function ψ.

Lemma 4.9 *The fiber $\psi^{-1}(\widehat{1}_R)$ consists of elements $x = \{x_1, x_2, \ldots, x_k\}$ where $x_i - x_{i-1} \in A$ for $1 \leq i \leq k$ and $n - x_k \in \{a\} \cup A$.*

Proof of Theorem 4.1: We know from Lemma 4.9 that the critical cells are in bijection with compositions of n where the last part belongs to the set $\{a\} \cup A$ and the remaining parts belong to the set A. Furthermore if such a composition has $i + 2$ parts, it will contribute to the i-dimensional homology. Hence, fixing i, we obtain the generating function

$$\begin{aligned}
\sum_{n \geq 0} \widetilde{\beta}_i q^n &= \left(\sum_{k=0}^{a-1} q^{a+kd} \right) \cdot \left(\sum_{\ell=1}^{a-1} q^{a+\ell d} \right)^{i+1} \\
&= q^{a+(i+1)(a+d)} \cdot \left(\sum_{k=0}^{a-1} q^{kd} \right) \cdot \left(\sum_{\ell=0}^{a-2} q^{\ell d} \right)^{i+1} \\
&= q^{a+(i+1)(a+d)} \cdot [a]_{q^d} \cdot [a-1]_{q^d}^{i+1}.
\end{aligned}$$

□

5 Concluding remarks

The Frobenius poset generated by two relatively prime integers can be embedded on a cylinder. There are many results (see, for example, [3, 6]) on posets that can be embedded in the plane. Can any of these results be extended to cylindrical posets?

There are other classes of generators, such as a geometric sequence, that have closed formulas for the Frobenius number, see [15]. Does the Frobenius complex have a nice topological representation in this case?

More generally, all computational evidence suggests that the Frobenius complex - even for randomly selected generators - has a relatively simple topology, that is, it is torsion-free. Is there a set of generators that creates torsion in the associated Frobenius complex?

Acknowledgements

The authors would like to thank Richard Stanley for pointing out the references [5, 10, 16] and Volkmar Welker for suggesting Corollary 3.2. We also thank Benjamin Braun and Margaret Readdy who read earlier versions of this paper.

References

[1] A. BARVINOK, "A Course in Convexity," American Mathematical Society, 2002.

[2] M. BECK AND S. ROBINS, "Computing the Continuous Discretely," Springer, 2007.

[3] L. J. BILLERA AND G. HETYEI, Decompositions of partially ordered sets, *Order* **17** (2000), 141–166.

[4] L. J. BILLERA AND A. N. MYERS, Shellability of interval orders, *Order* **15** (1999), 113–117.

[5] A. BJÖRNER AND V. WELKER, Segre and Rees products of posets, with ring-theoretic applications, *J. Pure Appl. Algebra* **198** (2005), 43–55.

[6] K. L. COLLINS, Planar lattices are lexicographically shellable, *Order* **8** (1992), 275–381.

[7] R. EHRENBORG AND G. HETYEI, The topology of the independence complex, *European J. Combin.* **27** (2006), 906–923.

[8] R. FORMAN, Morse theory for cell complexes, *Adv. Math.* **134** (1998), 90–145.

[9] R. FORMAN, A user's guide to Morse theory, *Séminaire Lotharingien de Combinatoire* **48** (2002), Article B48c.

[10] P. HERSH AND V. WELKER, Gröbner basis degree bounds on $\mathrm{Tor}^{k[\Lambda]}(k,k)$ and discrete Morse theory for posets, in *Integer points in polyhedra—geometry, number theory, algebra, optimization,* pp. 101–138, Contemp. Math., 374, Amer. Math. Soc., Providence, RI, 2005.

[11] D. N. KOZLOV, Complexes of directed trees, *J. Combin. Theory Ser. A* **88** (1999), 112–122.

[12] D. N. KOZLOV, "Combinatorial Algebraic Topology," Springer, 2008.

[13] D. N. KOZLOV, Discrete Morse theory and Hopf bundles, preprint 2009.

[14] O. A. LAUDAL AND A. SLETSJØE, Betti numbers of monoid algebras. Applications to 2-dimensional torus embeddings, *Math. Scand.* **56** (1985), 145–162.

[15] D. ONG AND V. PONOMARENKO, The Frobenius number of geometric sequences, *Integers* **8** (2008), A33.

[16] I. PEEVA, V. REINER, B. STURMFELS, How to shell a monoid, *Math. Ann.* **310** (1998), 379–393.

[17] J. B. ROBERTS, Note on linear forms, *Proc. Amer. Math Soc.* **7** (1956), 456–469.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 534–540

Extended Abstract for Enumerating Pattern Avoidance for Affine Permutations

Andrew Crites[†]

Department of Mathematics, University of Washington, Box 354350, Seattle, Washington, 98195-4350

Abstract. In this paper we study pattern avoidance for affine permutations. In particular, we show that for a given pattern p, there are only finitely many affine permutations in $\widetilde{S}_n$ that avoid p if and only if p avoids the pattern 321. We then count the number of affine permutations that avoid a given pattern p for each p in S_3, as well as give some conjectures for the patterns in S_4. This paper is just an outline; the full version will appear elsewhere.

Résumé. Dans cet œuvre, on étudie comment les permutations affines évitent les motifs. Spécifiquement, on peut dire que pour le motif p, il existe un nombre limité de permutations affines dans $\widetilde{S}_n$ qui évite p si et seulement si p évite le motif 321. Après, on compte le nombre de permutations affines qui évitent le motif p pour chaque p de S_3. Puis, on donne des conjectures pour les motifs de S_4. Ceci n'est qu'un aperçu; la version complète apparaîtra ailleurs.

Keywords: pattern avoidance, affine permutation, generating function, Catalan number

1 Introduction

Given a property Q, it is a natural question to ask if there is a simple characterization of all permutations with property Q. For example, in Lakshmibai and Sandhya (1990) the permutations corresponding to smooth Schubert varieties are exactly the permutations that avoid the two patterns 3412 and 4231. In Tenner (2007) it was shown that the permutations with Boolean order ideals are exactly the ones that avoid the two patterns 321 and 3412. A searchable database listing which classes of permutations avoid certain patterns can be found at Tenner (2009).

Since we know pattern avoidance can be used to describe useful classes of permutations, we might ask if we can enumerate the permutations avoiding a given pattern or set of patterns. For example, in Marcus and Tardos (2004) it was shown that if $S_n(p)$ is the number of permutations in the symmetric group, S_n, that avoid the pattern p, then there is some constant c such that $S_n(p) \leq c^n$. Thus the rate of growth of pattern avoiding permutations is bounded. This result was known as the Stanley-Wilf conjecture, now called the Marcus-Tardos Theorem.

We can express elements of the affine symmetric group, $\widetilde{S}_n$, as an infinite sequence of integers, and it is still natural to ask if there exists a subsequence with a given relative order. Thus we can extend the notion of pattern avoidance to these affine permutations and we can try to count how many $\omega \in \widetilde{S}_n$ avoid a given pattern.

[†]Andrew Crites acknowledges support from grant DMS-0800978 from the National Science Foundation.

For $p \in S_m$, let

$$f_n^p = \# \left\{ \omega \in \widetilde{S}_n : \omega \text{ avoids } p \right\} \tag{1}$$

and consider the generating function

$$f^p(t) = \sum_{n=2}^{\infty} f_n^p t^n. \tag{2}$$

For a given pattern p there could be infinitely many $\omega \in \widetilde{S}_n$ that avoid p. In this case, the generating function in (2) is not even defined. As a first step towards understanding $f^p(t)$, we will prove the following theorem.

Theorem 1 *Let $p \in S_m$. For any $n \geq 2$ there exist only finitely many $\omega \in \widetilde{S}_n$ that avoid p if and only if p avoids the pattern 321.*

It is worth noting that 321-avoiding permutations and 321-avoiding affine permutations appear as an interesting class of permutations in their own right. In (Billey et al., 1993, Theorem 2.1) it was shown that a permutation is fully commutative if and only if it is 321-avoiding. This means that every reduced expression for ω may be obtained from any other reduced expression using only relations of the form $s_i s_j = s_j s_i$ with $|i - j| > 1$. Moreover, a proof that this result can be extended to affine permutations as well appears in (Green, 2002, Theorem 2.7). For a detailed discussion of fully commutative elements in other Coxeter groups, see Stembridge (1996).

Even in the case where there might be infinitely many $\omega \in \widetilde{S}_n$ that avoid a pattern p, we can always construct the following generating function. Let

$$g_{m,n}^p = \# \left\{ \omega \in \widetilde{S}_n : \omega \text{ avoids } p \text{ and } \ell(\omega) = m \right\}. \tag{3}$$

Then set

$$g^p(x, y) = \sum_{n=2}^{\infty} \sum_{m=0}^{\infty} g_{m,n}^p x^m y^n. \tag{4}$$

Since there are only finitely many elements in $\widetilde{S}_n$ of a given length, we always have $g_{m,n}^p < \infty$. The generating function $g^{321}(x, y)$ is computed in (Hanusa and Jones, 2009, Theorem 3.2).

The outline of this abstract is as follows. In Section 2 we will review the definition of the affine symmetric group and list several of its useful properties. In Section 3 we will outline the proof of Theorem 1. Finally, in Section 4 we will give some basic results and conjectures for $f^p(t)$ for the patterns in S_3 and S_4. The full text of this paper has been submitted for publication and is currently available on the math arXiv:1002.1933.

2 Background

For $n \geq 2$, let $\widetilde{S}_n$ denote of the set of all bijections $\omega : \mathbb{Z} \to \mathbb{Z}$ with $\omega(i + n) = \omega(i) + n$ for all $i \in \mathbb{Z}$ and

$$\sum_{i=1}^{n} \omega(i) = \binom{n+1}{2}. \tag{5}$$

$\widetilde{S}_n$ is called the *affine symmetric group*, and the elements of $\widetilde{S}_n$ are called *affine permutations*. This definition of affine permutations first appeared in (Lusztig, 1983, §3.6) and was then developed in Shi (1986). Note that $\widetilde{S}_n$ also occurs as the affine Weyl group of type $\widetilde{A}_{n-1}$.

We can view an affine permutation in its one-line notation as the infinite string of integers

$$\cdots \omega_{-1}\omega_0\omega_1 \cdots \omega_n\omega_{n+1} \cdots ,$$

where, for simplicity of notation, we write $\omega_i = \omega(i)$. An affine permutation is completely determined by its action on $[n] := \{1, \ldots, n\}$. Thus we only need to record the base window $[\omega_1, \ldots, \omega_n]$ to capture all of the information about ω. Sometimes, however, it will be useful to write down a larger section of the one-line notation.

Given $i \not\equiv j \mod n$, let t_{ij} denote the affine transposition that interchanges $i + mn$ and $j + mn$ for all $m \in \mathbb{Z}$ and leaves all k not congruent to i or j fixed. Since $t_{ij} = t_{i+n,j+n}$ in $\widetilde{S}_n$, it suffices to assume $1 \leq i \leq n$ and $i < j$. Note that if we restrict to the affine permutations with $\{\omega_1, \ldots, \omega_n\} = [n]$, then we get a subgroup of $\widetilde{S}_n$ isomorphic to S_n, the group of permutations of $[n]$. Hence if $1 \leq i < j \leq n$, the above notion of transposition is the same as for the symmetric group.

Given a permutation $p \in S_k$ and an affine permutation $\omega \in \widetilde{S}_n$, we say that ω *contains the pattern* p if there is a subsequence of integers $i_1 < \cdots < i_k$ such that the subword $\omega_{i_1} \cdots \omega_{i_k}$ of ω has the same relative order as the elements of p. Otherwise, we say that ω *avoids* p. For example, if $\omega = [8, 1, 3, 5, 4, 0] \in \widetilde{S}_6$, then 8,1,5,0 is an occurrence of the pattern 4231 in ω, so that ω contains p. However, ω avoids the pattern 3412. A pattern can also come from terms outside of the base window $[\omega_1, \ldots, \omega_n]$. In the previous example, ω also has 2,8,6 as an occurrence of the pattern 132. Choosing a subword $\omega_{i_1} \cdots \omega_{i_k}$ with the same relative order as p will be referred to as *placing* p in ω.

2.1 Coxeter Groups

For a general reference on the basics of Coxeter groups, see Björner and Brenti (2005) or Humphreys (1990). Let $S = \{s_1, \ldots, s_n\}$ be a finite set, and let F denote the free group consisting of all words of finite length whose letters come from S. Here the group operation is concatenation of words, so that the empty word is the identity element. Let $M = (m_{ij})_{i,j=1}^n$ be any symmetric $n \times n$ matrix whose entries come from $\mathbb{Z}_{>0} \cup \{\infty\}$ with 1's on the diagonal and $m_{ij} > 1$ if $i \neq j$. Then let N be the normal subgroup of F generated by the relations

$$R = \{(s_i s_j)^{m_{ij}} = 1\}_{i,j=1}^n .$$

If $m_{ij} = \infty$, then there is no relationship between s_i and s_j. The Coxeter group corresponding to S and M is the quotient group $W = F/N$.

Any $w \in W$ can be written as a product of elements from S in infinitely many ways. Every such word will be called an *expression* for w. Any expression of minimal length will be called a *reduced expression*, and the number of letters in such an expression will be denoted $\ell(w)$, the *length* of w. Call any element of S a *simple reflection* and any element conjugate to a simple reflection, a *reflection*.

We graphically encode the relations in a Coxeter group via its *Coxeter graph*. This is the labeled graph whose vertices are the elements of S. We place an edge between two vertices s_i and s_j if $m_{ij} > 2$ and we label the edge m_{ij} whenever $m_{ij} > 3$. The Coxeter graphs of all the finite Coxeter groups have been classified. See, for example, (Humphreys, 1990, §2).

In (Björner and Brenti, 2005, §8.3) it was shown that $\widetilde{S}_n$ is the Coxeter group with generating set $S = \{s_0, s_1, \ldots, s_{n-1}\}$, and relations

$$R = \begin{cases} s_i^2 = 1, & \\ (s_i s_j)^2 = 1, & \text{if } |i-j| \geq 2, \\ (s_i s_{i+1})^3 = 1, & \text{for } 0 \leq i \leq n-1, \end{cases}$$

where all of the subscripts are taken mod n. Thus the Coxeter graph for $\widetilde{S}_n$ is an n-cycle, where every edge is unlabeled.

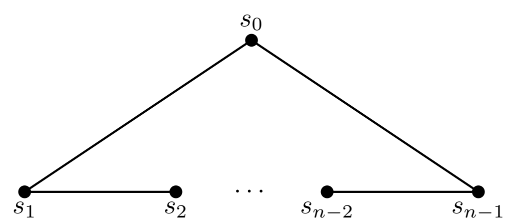


Fig. 1: Coxeter graph for $\widetilde{S}_n$.

If $J \subsetneq S$ is a proper subset of S, then we call the subgroup of W generated by just the elements of J a *parabolic subgroup*. Denote this subgroup by W_J. In the case of the affine symmetric group we have the following characterization of parabolic subgroups, which follows easily from the fact that when $J = S \backslash \{s_i\}$, $(\widetilde{S}_n)_J = \mathrm{Stab}([i, i+n-1])$ (Björner and Brenti, 2005, Proposition 8.3.4).

Proposition 2 *Let $J = S \backslash \{s_i\}$. Then $\omega \in \widetilde{S}_n$ is in the parabolic subgroup $(\widetilde{S}_n)_J$ if and only if there exists some integer $i \leq j \leq i+n-1$ such that $\omega_j \leq \omega_k < \omega_j + n$ for all $i \leq k \leq i+n-1$.*

2.2 Length Function for $\widetilde{S}_n$

For $\omega \in \widetilde{S}_n$, let $\ell(\omega)$ denote the length of ω when $\widetilde{S}_n$ is viewed as a Coxeter group. Recall that for a non-affine permutation $\pi \in S_n$ we can define an *inversion* as a pair (i, j) such that $i < j$ and $\pi_i > \pi_j$. For an affine permutation, if $\omega_i > \omega_j$ for some $i < j$, then we also have $\omega_{i+kn} > \omega_{j+kn}$ for all $k \in \mathbb{Z}$. Hence any affine permutation with a single inversion has infinitely many inversions. Thus we standardize each inversion as follows. Define an *affine inversion* as a pair (i, j) such that $1 \leq i \leq n$, $i < j$, and $\omega_i > \omega_j$. If we let $\mathrm{Inv}_{\widetilde{S}_n}(\omega)$ denote the set of all affine inversions in ω, then $\ell(\omega) = \#\mathrm{Inv}_{\widetilde{S}_n}(\omega)$, (Björner and Brenti, 2005, Proposition 8.3.1).

We also have the following characterization of the length of an affine permutation, which will be useful later.

Theorem 3 (Shi, 1986, Lemma 4.2.2) *Let $\omega \in \widetilde{S}_n$. Then*

$$\ell(\omega) = \sum_{1 \leq i < j \leq n} \left| \left\lfloor \frac{\omega_j - \omega_i}{n} \right\rfloor \right| = \mathrm{inv}(\omega_1, \ldots, \omega_n) + \sum_{1 \leq i < j \leq n} \left\lfloor \frac{|\omega_j - \omega_i|}{n} \right\rfloor, \tag{6}$$

where $\mathrm{inv}(\omega_1, \ldots, \omega_n) = \#\{1 \leq i < j \leq n : \omega_i > \omega_j\}$.

For $1 \leq i \leq n$ define $\mathrm{Inv}_i(\omega) = \#\{j \in \mathbb{N} : i < j, \omega_i > \omega_j\}$. Now let $\mathrm{Inv}(\omega) = (\mathrm{Inv}_1(\omega), \ldots, \mathrm{Inv}_n(\omega))$, which will be called the *affine inversion table* of ω. In (Björner and Brenti, 1996, Theorem 4.6) it was shown that there is a bijection between $\widetilde{S}_n$ and elements of $\mathbb{Z}^n_{\geq 0}$ containing at least one zero entry.

3 Outline of Proof of Theorem 1

The Proof of Theorem 1 is broken up into two parts. First, if $p \in S_m$ contains the pattern 321, then we exhibit an infinite family of affine permutations, all of which avoid 321 and hence avoid p. Second, if p avoids the pattern 321, then we show that there exists a constant L, depending on p, such that if $\ell(\omega) > L$, then ω must contain p as follows. Using the length formula in Theorem 3, if $\ell(\omega)$ is large, then there must be two indices $1 \leq i < j \leq n$ with $|\omega_i - \omega_j|$ large. Once $|\omega_i - \omega_j|$ is large enough, we then show how to use translates ω_{i+rn} and ω_{j+sn} of ω_i and ω_j to construct an occurrence of p in ω. Hence if ω avoids p, $\ell(\omega)$ must be bounded above, so that there can be only a finite number of such ω.

The algorithm for constructing an occurrence of p gives the length bound $\ell(\omega) \leq (m^{\ell+1} + 2)\binom{n}{2}$, where $p \in S_m$, $\omega \in \widetilde{S}_n$ and ℓ is the length of the sequence of left-to-right maxima in p. In general, this upper bound is much larger than needed. For example, let $p = 3412 \in S_4$. Then our algorithm gives that if $\omega \in \widetilde{S}_n$ avoids p, then $\ell(\omega) \leq 66\binom{n}{2}$. However, we can actually prove a tighter bound $\ell(\omega) \leq 3\binom{n}{2}$ for this particular pattern. Thus it would be nice to find an algorithm that gives a tighter upper bound on length.

4 Generating Functions for Patterns in S_3 and S_4

Let f_n^p and $f^p(t)$ be as in (1) and (2) in Section 1. Then by Theorem 1 we have $f_n^{321} = \infty$ for all n. However, for all of the other patterns $p \in S_3$ we can still compute $f^p(t)$.

Theorem 4 *Let $f^p(t)$ be as above. Then*

$$f^{123}(t) = 0, \tag{7}$$

$$f^{132}(t) = f^{213}(t) = \sum_{n=2}^{\infty} t^n, \tag{8}$$

$$f^{231}(t) = f^{312}(t) = \sum_{n=2}^{\infty} \binom{2n-1}{n} t^n. \tag{9}$$

The only tricky part in the proof of Theorem 4 is Equation 9. The proof involves using the affine inversion table of an affine permutation and some identities amongst the Catalan numbers.

We now look at pattern avoidance for patterns in S_4. There are 24 patterns to consider, although for all but three patterns, $f^p(t)$ is easy to compute. First let

$$P = \{1432, 2431, 3214, 3241, 3421, 4132, 4213, 4231, 4312, 4321\}.$$

By Theorem 1, if $p \in P$, then $f_n^p = \infty$, so $f^p(t)$ is not defined.

Theorem 5 *We have*

$$f^{1234}(t) = 0, \tag{10}$$

$$f^{1243}(t) = f^{1324}(t) = f^{2134}(t) = f^{2143}(t) = \sum_{n=2}^{\infty} t^n, \tag{11}$$

$$f^{1342}(t) = f^{1423}(t) = f^{2314}(t) = f^{3124}(t) = \sum_{n=2}^{\infty} \binom{2n-1}{n} t^n. \tag{12}$$

Based on some initial calculations, we also have the following conjectures for the remaining three patterns in S_4.

Conjecture 1 *The following equalities hold:*

$$f_n^{3142} = \sum_{k=0}^{n-1} \frac{(n-k)}{n} \binom{n-1+k}{k} 2^k \tag{13}$$

$$f_n^{3412} = f_n^{4123} = \frac{1}{3} \sum_{k=0}^{n} \binom{n}{k}^2 \binom{2k}{k}. \tag{14}$$

Note that (13) is sequence A064062 and (14) is sequence A087457 in Sloane (2009). It is also worth comparing (14) to the number of 3412-avoiding, non-affine permutations given in (Gessel, 1990, §7) as

$$u_3(n) = 2 \sum_{k=0}^{n} \binom{n}{k}^2 \binom{2k}{k} \frac{3k^2 + 2k + 1 - n - 2kn}{(k+1)^2(k+2)(n-k+1)}. \tag{15}$$

Acknowledgements

The author would like to thank Brant Jones for inspiring him to ask this question and Sara Billey for all of her guidance and helpful conversations. He is also indebted to Neil Sloane and his team, whose encyclopedia of integer sequences was invaluable for forming the conjectures in this paper.

References

S. Billey and S. A. Mitchell. Affine partitions and affine Grassmannians. *Electron. J. Combin.*, 16(2): Research Paper 18, 45 pp. (electronic), 2009.

S. Billey, W. Jockusch, and R. P. Stanley. Some combinatorial properties of Schubert polynomials. *J. Algebraic Combin.*, 2(4):345–374, 1993. ISSN 0925-9899.

A. Björner and F. Brenti. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York, 2005. ISBN 978-3540-442387; 3-540-44238-3.

A. Björner and F. Brenti. Affine permutations of type A. *Electron. J. Combin.*, 3(2):Research Paper 18, approx. 35 pp. (electronic), 1996. ISSN 1077-8926. The Foata Festschrift.

I. M. Gessel. Symmetric functions and P-recursiveness. *J. Combin. Theory Ser. A*, 53(2):257–285, 1990. ISSN 0097-3165. doi: 10.1016/0097-3165(90)90060-A. URL `http://dx.doi.org/10.1016/0097-3165(90)90060-A`.

R. M. Green. On 321-avoiding permutations in affine Weyl groups. *J. Algebraic Combin.*, 15(3):241–252, 2002. ISSN 0925-9899.

C. Hanusa and B. Jones. The enumeration of fully commutative affine permutations. preprint, math.CO/0907.0709v1, July 2009.

J. E. Humphreys. *Reflection groups and Coxeter groups*, volume 29 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990. ISBN 0-521-37510-X.

D. E. Knuth. *The art of computer programming. Volume 3*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. Sorting and searching, Addison-Wesley Series in Computer Science and Information Processing.

V. Lakshmibai and B. Sandhya. Criterion for smoothness of Schubert varieties in $\mathrm{Sl}(n)/B$. *Proc. Indian Acad. Sci. Math. Sci.*, 100(1):45–52, 1990. ISSN 0253-4142.

G. Lusztig. Some examples of square integrable representations of semisimple p-adic groups. *Trans. Amer. Math. Soc.*, 277(2):623–653, 1983. ISSN 0002-9947.

A. Marcus and G. Tardos. Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Combin. Theory Ser. A*, 107(1):153–160, 2004. ISSN 0097-3165.

J. Y. Shi. *The Kazhdan-Lusztig cells in certain affine Weyl groups*, volume 1179 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. ISBN 3-540-16439-1.

N. J. A. Sloane. Online encyclopedia of integer sequences. `http://www.research.att.com/~njas/sequences/`, 2009.

J. R. Stembridge. On the fully commutative elements of Coxeter groups. *J. Algebraic Combin.*, 5(4): 353–385, 1996. ISSN 0925-9899.

B. E. Tenner. Pattern avoidance and the Bruhat order. *J. Combin. Theory Ser. A*, 114(5):888–905, 2007. ISSN 0097-3165.

B. E. Tenner. Database of permutation pattern avoidance. `http://math.depaul.edu/~bridget/patterns.html`, 2009.

J. West. *Permutations with forbidden sequences; and, stack-sortable permutations*. PhD thesis, Massachusetts Institute of Technology, September 1990.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 541–552

Tropical secant graphs of monomial curves

María Angélica Cueto[1†] and Shaowei Lin[1‡]

[1] *Department of Mathematics, University of California, Berkeley, CA 94720, USA*

Abstract. We construct and study an embedded weighted balanced graph in $\mathbb{R}^{n+1}$ parameterized by a strictly increasing sequence of n coprime numbers $\{i_1, \ldots, i_n\}$, called the *tropical secant surface graph*. We identify it with the tropicalization of a surface in $\mathbb{C}^{n+1}$ parameterized by binomials. Using this graph, we construct the tropicalization of the first secant variety of a monomial projective curve with exponent vector $(0, i_1, \ldots, i_n)$, which can be described by a balanced graph called the *tropical secant graph*. The combinatorics involved in computing the degree of this classical secant variety is non-trivial. One earlier approach to this is due to K. Ranestad. Using techniques from tropical geometry, we give algorithms to effectively compute this degree (as well as its multidegree) and the Newton polytope of the first secant variety of any given monomial curve in $\mathbb{P}^4$.

Résumé. On construit et on étude un graphe plongé dans $\mathbb{R}^{n+1}$ paramétrisé par une suite strictement croissante de n nombres entiers $\{i_1, \ldots, i_n\}$, premiers entre eux. Ce graphe s'appelle *graphe tropical surface sécante*. On montre que ce graphe est la tropicalisation d'une surface dans $\mathbb{C}^{n+1}$ paramétrisé par des binômes. On utilise ce graphe pour construire la tropicalisation de la première sécante d'une courbe monomiale ayant comme vecteur d'exponents $(0, i_1, \ldots, i_n)$. On répresent ce variété tropical pour un graphe balancé (le *graphe tropical sécante*). La combinatoire qu'on utilise pour le calcul du degré de ces variétés sécantes classiques n'est pas triviale, et a été developé par K. Ranestad. En utilisant des techniques de la géométrie tropicale, on donne des algorithmes qui calculent le degré (même le multidegré) et le polytope de Newton de la première sécante d'une courbe monomiale de $\mathbb{P}^4$.

Keywords: monomial curves, secant varieties, resolution graphs, tropical geometry, Newton polytope

1 Introduction

In this paper, we define and study an abstract graph (the *abstract tropical secant surface graph*) which we embed in $\mathbb{R}^{n+1}$, assigning integer coordinates to each node. This graph is parameterized by a sequence of n coprime positive integers $i_1 < \ldots < i_n$. The abstract graph is constructed by gluing two caterpillar trees and several star trees, according to the combinatorics of the given integer sequence. Our embedding has a key feature: we can endow this graph with weights on all edges in such a way that it satisfies the balancing condition (Theorem 3). We call this weighted graph the *tropical secant surface graph* or *master graph* (Section 2). As the name suggests, this balanced graph is closely related to a tropical surface and it will be the cornerstone of our paper. More precisely, it is the building block for constructing the tropicalization of a threefold: the first secant variety of a monomial projective curve whose set of

† Supported by a UC Berkeley Chancellor's Fellowship.
‡ Supported by a Singapore A*STAR Fellowship.

exponents is $\{0, i_1, \ldots, i_n\}$. By definition, this secant variety is the closure of the union of lines that meet the curve in two distinct points. These varieties have been studied extensively in the literature (Cox and Sidman, 2007; Ranestad, 2006). We describe this tropical connection in Section 6.

The tropicalization of the first secant variety of a monomial projective curve *strictly contains*, as a *subfan*, the set of all tropical lines between any two points in the tropicalization of the monomial curve itself, i.e. points that are obtained as coordinatewise minima of two points in the classical plane spanned by the lattice $\Lambda = \langle \mathbf{1}(0, i_1, \ldots, i_n)\rangle$. The latter is the first tropical secant variety of the corresponding classical line in the n-dimensional tropical projective torus $\mathbb{TP}^n = \mathbb{R}^{n+1}/(1, \ldots, 1)$. The union of these tropical lines is precisely the cone from the classical line $\mathbb{R}\langle(0, i_1, \ldots, i_n)\rangle$ over the pure 1-dimensional subfan of the secondary fan of the point configuration $\{0, i_1, \ldots, i_n\} \subset \mathbb{R}$ consisting of all regular subdivisions with the property that two of its facets contains all $n + 1$ points. By (Theorem 3.1, Develin, 2006), we know that this subfan is precisely the cone from the plane $\mathbb{R} \otimes \Lambda$ over the complex of lower faces of the cyclic polytope $C(2, n-1)$ (i.e. $n-1$ points in dimension 2). This complex is the subgraph of the tropical secant graph consisting of the chain graph with $n - 1$ nodes $E_{i_1}, \ldots, E_{i_{n-1}}$, depicted in Figure 1.

In recent years, tropical geometry has provided a new approach to attack implicitization problems (Dickenstein et al., 2007; Sturmfels et al., 2007; Cueto et al., 2010). In particular, tropicalization interplays nicely with several classical constructions, such as Hadamard products of subvarieties of tori. Using such techniques, we can effectively compute the Chow polytope of these secant varieties, as we discuss in Section 7. In the case of the secants of monomial curves in $\mathbb{P}^4$, the Chow polytopes coincide with the Newton polytopes of these hypersurfaces. Interpolation techniques can then be used to obtain their defining equations.

As one may suspect, computing the tropicalization of an algebraic variety without information on its defining ideal is not an easy task. Such methods rely on a parametric representation of the variety and the characterization of tropical varieties in terms of valuations (Bieri and Groves, 1984), and they are known as *geometric tropicalization* (Theorem 7). As we explain in Section 4, the main difficulty lies in finding a suitable compactification of the variety such that its boundary has simple normal crossings, or combinatorial normal crossings in the case of surfaces. However, this geometric construction does not provide information about the tropical variety as a weighted set: the multiplicities are missing in the construction of Hacking et al. (2009) and they are essential for tropical implicitization methods. We give a formula to compute these numbers in Theorem 8. The combinatorics involved in the construction of such compactifications is non-trivial, since they are the combinatorial counterpart of the algebro-geometric process of *resolution of singularities*.

In the case of surfaces, the resolution can be achieved in theory by blowing up plane curves at finitely many points, as described in Section 5. We then use the rational parameterization of the original surface to obtain a resolution of this surface from the resolution of the arrangement of plane curves in $\mathbb{T}^2$. In practice, knowing which points to blow up and how the intersection multiplicities of proper transforms and exceptional divisors are carried along the various blow-ups can be a combinatorial challenge. However, the surfaces studied in this paper (binomial surfaces obtained from a dehomogeneization of the first secant of monomial projective curves) have very rich combinatorial structures, and we can make full use of this feature to compute their tropicalizations via resolutions. Indeed, our methods allow us to read off the intersection numbers of the boundary divisors directly from the master graphs, which encode the resolution diagrams of these surfaces (Figure 1). This is carried out in Section 3, in particular in Theorem 3.

Finally, we use this tropical surface to effectively compute the first secant variety of any monomial curve as a collection of 4-dimensional cones *with multiplicities* (Theorem 16). From this construction we recover the multidegree of this secant variety with respect to the rank-two lattice generated by the all-one's vector and the exponent vector parameterizing the curve. The degree of this variety was previously worked out in (Ranestad, 2006), and our work gives similar combinatorial formulas for this degree in terms of the exponent vector. But tropical methods enable us to obtain more information, namely the Chow polytope of the secant variety. We illustrate all our results in Example 18 which was inspired by (Ranestad, 2006).

2 The master graph

In this section, we describe the main object of this paper: the master graph. We start by defining an abstract graph, called the *abstract tropical secant surface graph*, parameterized by a list $(i_1,\dots,i_n)$ of n distinct, coprime, nonnegative integers. Throughout the paper, we set $n \geq 4$ and we call $i_0 = 0$ to simplify notation. We construct this abstract graph by gluing three different families of graphs along the common labeled nodes D_{i_j}, as depicted in Figure 1. The first two graphs $G_{E,D}$ and $G_{h,D}$ are caterpillar trees with $2n-1$ and $2n$ nodes, grouped in two levels, with labels $D_0, D_{i_1},\dots,D_{i_n}$, $E_{i_1},\dots,E_{i_{n-1}}$ and $h_{i_1},\dots,h_{i_{n-1}}$ respectively. The third family of graphs is parameterized by subsets of the index set $\{0,i_1,\dots,i_n\}$ of size at least two, which are obtained by intersecting an arithmetic progression of integers with the index set. Note that several arithmetic progressions can give the same subset of $\{0,i_1,\dots,i_n\}$ and all of them will give the same node $F_{\underline{a}}$ in the graph. If $\underline{a} = \{i_{j_1},\dots,i_{j_k}\}$ then the graph $G_{F_{\underline{a}},D}$ has $k+1$ nodes and k edges: a central node $F_{\underline{a}}$ and k nodes labeled $D_{i_{j_1}},\dots,D_{i_{j_k}}$. The central node is connected to the other k nodes in the graph.

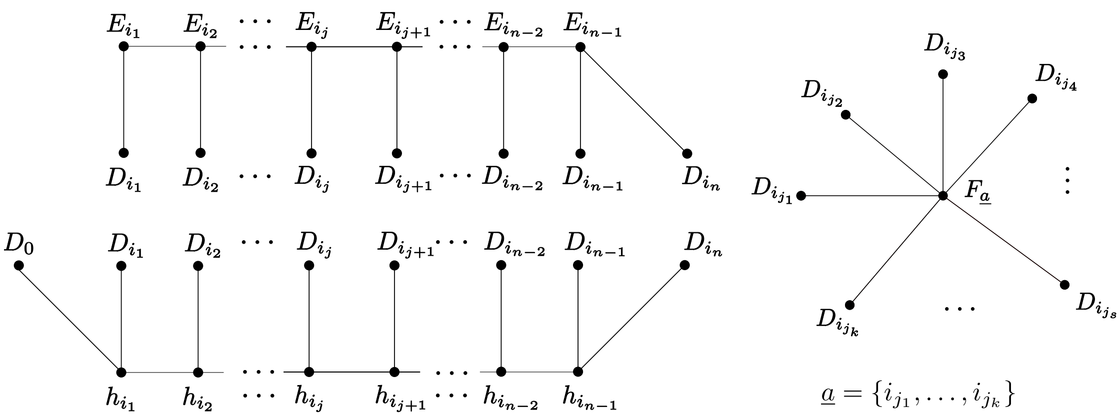


Fig. 1: The graphs $G_{E,D}$, $G_{h,D}$ and $G_{F_{\underline{a}},D}$ glue together to form the *abstract tropical secant surface graph*.

Next, we embed this graph in $\mathbb{R}^{n+1}$, mapping each node to an integer vector, as in Definition 1. Our chosen embedding has addition data: a weight on each edge that makes the graph *balanced*. We call this weighted graph the *tropical secant surface graph* or *master graph*. For a numerical example, see Figure 2.

Definition 1 *The* master graph *is a weighted graph in* $\mathbb{R}^{n+1}$ *parameterized by* $\{i_1,\dots,i_n\}$ *with nodes:*

(i) $D_{i_j} = e_j := (0,\dots,0,1,0,\dots,0)$ $\quad$ *(*$0 \leq j \leq n$*),*

(ii) $E_{i_j} = (0,i_1,\dots,i_{j-1},i_j,\dots,i_j)\,,\, h_{i_j} = (-i_j,-i_j,\dots,-i_j,-i_{j+1},\dots,-i_n)$ $\quad$ *(*$1 \leq j \leq n-1$*),*

(iii) $F_{\underline{a}} = \sum_{i_j \in \underline{a}} e_j$ *where* $\underline{a} \subseteq \{0, i_1, \dots, i_n\}$ *has size at least two and is obtained by intersecting an arithmetic progression of integers with the index set* $\{0, i_1, \dots, i_n\}$.

Its edges agree with the edges of the abstract tropical secant surface graph, and have weights:

(i) $m_{D_{i_0}, h_{i_1}} = 1$, $m_{D_{i_n}, E_{i_{n-1}}} = \gcd(i_1, \dots, i_{n-1})$, $m_{D_{i_n}, h_{i_{n-1}}} = i_n$,

(ii) $m_{D_{i_j}, E_{i_j}} = \gcd(i_1, \dots, i_j)$, $m_{D_{i_j}, h_{i_j}} = \gcd(i_j, \dots, i_n)$ $\quad (1 \leq j \leq n-1)$,

(iii) $m_{E_{i_j}, E_{i_{j+1}}} = \gcd(i_1, \dots, i_j)$, $m_{h_{i_j}, h_{i_{j+1}}} = \gcd(i_{j+1}, \dots, i_n)$ $\quad (1 \leq j \leq n-2)$,

(iv) $m_{F_{\underline{a}}, D_{i_j}} = \sum_r \varphi(r)$*, where we sum over the common differences* r *of all arithmetic progressions containing* i_j *and giving the same subset* $\underline{a}$*. Here,* φ *denotes Euler's phi function.*

Definition 2 *Let* $(G, m) \subset \mathbb{R}^N$ *be a weighted graph where each node has integer coordinates. Let* w *be a node in* G *and let* $\{w_1, \dots, w_r\}$ *be the set of nodes adjacent to* w*. Consider the primitive lattices* $\Lambda_w = \mathbb{R}\langle w \rangle \cap \mathbb{Z}^N$ *and* $\Lambda_{w,w_i} = \mathbb{R}\langle w, w_i \rangle \cap \mathbb{Z}^N$*. Then* $\Lambda_{w,w_i}/\Lambda_w$ *is a rank one lattice, and it admits a unique generator* u_i *lifting to the cone* $\mathbb{R}_{\geq 0}\langle w, w_i \rangle / \mathbb{R}\langle w \rangle$*. We say that the* node w *is* balanced *if* $\sum_{i=1}^r m(w_i, w) u_i = 0 \in \mathbb{R}^N / \mathbb{R}\langle w \rangle$*. If all nodes are balanced, then* G *satisfies the* balancing condition.

Theorem 3 *The master graph satisfies the* balancing condition.

Remark 4 *If the arithmetic progression* $\underline{a}$ *has two elements, then* $F_{\underline{a}}$ *is a bivalent node and we can safely eliminate it from the graph if desired, replacing its two adjacent edges by a single edge. Both edges have the same multiplicity, which we assign to the new edge. To simplify notation, we keep these bivalent nodes.*

3 The master graph is a tropical surface

In this section, we explain the suggestive name "tropical secant surface graph." More concretely, we show that the master graph is the tropicalization of a surface in $\mathbb{C}^{n+1}$ parameterized by the binomial map $(\lambda, w) \mapsto (1-\lambda, w^{i_1} - \lambda, \dots, w^{i_n} - \lambda)$. Before that, we review the basics of tropical geometry.

Definition 5 *Given a variety* $X \subset \mathbb{C}^N$ *with defining ideal* $I = I_X$*, we define the* tropicalization *of* X *as*

$$\mathcal{T}X = \mathcal{T}I = \{w \in \mathbb{R}^N : \mathit{in}_w(I) \text{ does not contain a monomial}\}.$$

Here, $\mathit{in}_w(I) = \langle \mathit{in}_w(f) : f \in I \rangle$*, and if* $f = \sum_\alpha c_\alpha \underline{x}^\alpha$ *where all* $c_\alpha \neq 0$*, then* $\mathit{in}_w(f) = \sum_{\alpha \cdot w = W} c_\alpha \underline{x}^\alpha$ *where* $W = \min\{\alpha \cdot w : c_\alpha \neq 0\}$*. In the case of an embedded projective variety* $X \subset \mathbb{P}^N$*, the tropicalization of* X *is defined as* $\mathcal{T}(X') \subset \mathbb{R}^{N+1}$ *where* X' *is the affine cone over* X *in* $\mathbb{C}^{N+1}$.

Although it may not be clear from Definition 5, tropicalizations are toric in nature. More precisely, let $\mathbb{T}^N = (\mathbb{C}^*)^N$ be the algebraic torus. Let Y be a subvariety of $\mathbb{T}^N$ with defining ideal $I_Y \subseteq \mathbb{C}[\mathbb{T}^N] = \mathbb{C}[y_1^{\pm}, \dots, y_N^{\pm}]$. We define the tropicalization of $Y \subset \mathbb{T}^N$ as

$$\mathcal{T}Y = \{v \in \mathbb{R}^N : 1 \notin \mathrm{in}_v(I_Y)\}.$$

Here, the initial ideal with respect to a vector v is the same as that in Definition 5. Consider the Zariski closure $\overline{Y}$ of Y in $\mathbb{C}^N$. It is easy to see that $\mathcal{T}Y$ equals $\mathcal{T}\overline{Y}$. Indeed, this follows from the fact that I_Y is the saturation ideal $\left(I_{\overline{Y}}\mathbb{C}[\mathbb{T}^N] : (y_1 \cdots y_N)^\infty\right)$ and $I_{\overline{Y}} = I_Y \cap \mathbb{C}[y_1, \dots, y_N]$. Therefore, if we start with an irreducible variety $X \subset \mathbb{C}^N$ not contained in a coordinate hyperplane, then we can consider the very affine variety $Y = X \cap \mathbb{T}^N$, which has the same dimension as X. The tropical variety $\mathcal{T}Y$ is a pure polyhedral subfan of the Gröbner fan of I and it preserves an important invariant of Y: both objects have the same dimension (Bieri and Groves, 1984). We can choose to study $\mathcal{T}Y$ or $\mathcal{T}X$, and both sets will give us equivalent information about X. This approach will be useful in subsequent sections.

Tropical implicitization is a recently developed technique to approach classical implicitization problems (Sturmfels and Tevelev, 2008). For instance, when Y is a codimension-one hypersurface, $I_Y = \langle g \rangle$ is principal and $\mathcal{T}Y$ is the union of non-maximal cones in the normal fan of the Newton polytope of g, so knowing $\mathcal{T}Y$ can help us in finding g. But to achieve this, we need to compute $\mathcal{T}Y$ without explicitly knowing I_Y. We show how to do this in Section 4.

A point $w \in \mathcal{T}X$ is called *regular* if $\mathcal{T}X$ is a linear space locally near w. We can assign a positive integer number to regular points of the tropical variety, to have good properties. More precisely, we define the *multiplicity* m_w of a regular point w as the sum of multiplicities of all minimal associated primes of the initial ideal $\mathrm{in}_w(I)$. For a given maximal cone σ in $\mathcal{T}X$, we define its multiplicity as the multiplicity at a regular point w in σ, that is, the multiplicity of any point in the relative interior. One can show that this assignment does not depend on the choice of w and that with these multiplicities, the tropical variety satisfies the *balancing condition* (Corollary 3.4, Sturmfels and Tevelev, 2008).

In the case of projective varieties, or in general, when we have a torus action, the tropical variety $\mathcal{T}X$ has a *lineality space*, that is, the maximal linear space contained in all cones of the fan $\mathcal{T}X$. For example, the lineality space of a tropical hypersurface $\mathcal{T}(g)$ will equal the orthogonal complement of the affine span of the Newton polytope of g, after appropriate translation to the origin. The extreme cases correspond to toric varieties globally parameterized by a monomial map with associated matrix A. Their tropicalizations $\mathcal{T}X$ will be classical linear spaces: the row span of A. In particular, $\mathcal{T}X$ coincides with its lineality space as sets with constant multiplicity *one* (Dickenstein et al., 2007).

We now realize the master graph as a tropical surface in $\mathbb{R}^{n+1}$:

Theorem 6 *Fix a strictly increasing sequence $(0, i_1, \dots, i_n)$ of coprime integers. Let Z be the surface in $\mathbb{C}^{n+1}$ parameterized by $(\lambda, \omega) \mapsto (1-\lambda, \omega^{i_1} - \lambda, \dots, \omega^{i_n} - \lambda)$. Then, the tropical surface $\mathcal{T}Z \subset \mathbb{R}^{n+1}$ coincides with the cone over the master graph as weighted polyhedral fans, with the convention that we assign the weight $m_{D_{i_1}, E_{i_1}} + m_{F_{\underline{e}}, D_{i_1}}$ to the cone over the edge $D_{i_1}E_{i_1}$ and we disregard the cone over the edge $D_{i_1}F_{\underline{e}}$, if the ending sequence $\underline{e} = \{i_1, \dots, i_n\}$ gives a node $F_{\underline{e}}$ in the master graph.*

The proof of this statement involves techniques from geometric tropicalization and resolution of singularities of plane curves. Beautiful combinatorics emerge from them, as we will see in the next sections.

4 Geometric Tropicalization

In this section, we present the basics of geometric tropicalization. The spirit of this approach relies on computing the tropicalization of subvarieties of tori by analyzing the combinatorics of their boundary in a suitable compactification of the torus and of the subvariety therein. In what follows, we describe the method and its applications to implicitizations of subvarieties of tori.

Let $f_1, \dots, f_N$ be Laurent polynomials in $\mathbb{C}[t_1^{\pm}, \dots, t_r^{\pm}]$ and consider the rational map $\mathbf{f}\colon \mathbb{T}^r \dashrightarrow \mathbb{T}^N$, $\mathbf{f} = (f_1, \dots, f_N)$. For simplicity, we will assume that the fiber of $\mathbf{f}$ over a generic point of $Y \subset \mathbb{T}^N$ is finite. Our goal is to compute the tropicalization $\mathcal{T}Y$ of the closure of the image of the map $\mathbf{f}$ inside the torus without knowledge of its defining ideal. When the coefficients of $f_1, \dots, f_N$ are generic with respect to their Newton polytopes, a method for constructing $\mathcal{T}Y$ was given in (Thm 2.1, Sturmfels et al., 2007) and proved in (Thm 5.1, Sturmfels and Tevelev, 2008). We describe an algorithm proposed in (§5, Sturmfels and Tevelev, 2008) which may be applied to maps $\mathbf{f}$ which are non-generic. For simplicity, we state it for the case of parametric *surfaces*, although the method generalizes to higher dimensions as well.

Theorem 7 (Geometric Tropicalization (Hacking et al., 2009, §2)) *Let $\mathbb{T}^N$ be the N-dimensional torus over $\mathbb{C}$ with coordinate functions $t_1, \dots, t_N$, and let Y be a closed* surface *in $\mathbb{T}^N$. Suppose Y is smooth and $\overline{Y} \supset Y$ is any compactification whose boundary $D = \overline{Y} \setminus Y$ is a smooth divisor with simple normal crossings. Let $D_1, \dots, D_m$ be the irreducible components of D, and write $\Delta_{Y,D}$ for the intersection complex of the boundary divisor D, i.e. the graph on $\{1, \dots, m\}$ defined by*

$$\{k_i, k_j\} \in \Delta_{Y,D} \iff D_{k_i} \cap D_{k_j} \neq \emptyset.$$

Define the integer vectors $[D_k] := (val_{D_k}(t_1), \dots, val_{D_k}(t_N)) \in \mathbb{Z}^N$ ($k = 1, \dots, m$), where $val_{D_k}(t_j)$ is the order of zero-poles of t_j along D_k. For any $\sigma \in \Delta_{Y,D}$, let $[\sigma]$ be the cone in $\mathbb{Z}^N$ spanned by $\{[D_k] : k \in \sigma\}$ and let $\mathbb{R}_{\geq 0}[\sigma]$ be the cone in $\mathbb{R}^N$ spanned by the same integer vectors. Then,

$$\mathcal{T}Y = \bigcup_{\sigma \in \Delta_{Y,D}} \mathbb{R}_{\geq 0}[\sigma].$$

We complement the previous result by a formula giving the multiplicities of regular points in tropical surfaces. A similar formula will hold in higher dimensions:

Theorem 8 *(Cueto, 2011) In the notation of Theorem 7, the multiplicity of a regular point w in the tropical surface equals:*

$$m_w = \sum_{\substack{\sigma \in \Delta_{Y,D} \\ w \in \mathbb{R}_{\geq 0}[\sigma]}} (D_{k_1} \cdot D_{k_2})\, \mathit{index}\big((\mathbb{R} \otimes_{\mathbb{Z}} [\sigma]) \cap \mathbb{Z}^N : \mathbb{Z}[\sigma]\big),$$

where $D_{k_1} \cdot D_{k_2}$ denotes the intersection number of these divisors and we sum over all two-dimensional cones σ whose associated rational cone $\mathbb{R}_{\geq 0}[\sigma]$ contains the point w.

To compute $\mathcal{T}Y$ using the previous theorems, we require a compactification $\overline{Y} \supset Y$ whose boundary has simple normal crossings (SNC). In words, all components of the divisor D should be smooth and they show intersect "as transversally as possible." One method for producing such a *tropical compactification* is taking the closure $\overline{Y}$ of Y in $\mathbb{P}^N \supset \mathbb{T}^N$ and finding a resolution of singularities for the boundary $\overline{Y} \setminus Y$. This latter step can be difficult. However, in the case of surfaces, it is enough to require the boundary to have *combinatorial normal crossings* (CNC), that is, "no three divisors intersect at a point" (Sturmfels and Tevelev, 2008). We describe the resolution process for our binomial surface Z in the next section.

5 Combinatorics of Monomial Curves

In this section, we compute the tropical variety of the surface Z described in Theorem 6. Let $f_{i_j} := \omega^{i_j} - \lambda$ ($0 \leq j \leq n$) and consider the parameterization $\mathbf{f}\colon \mathbb{C}^2 \to Z$ given by these $n+1$ polynomials. Since geometric tropicalization involves subvarieties of tori, we restrict our domain to $X = \mathbb{T}^2 \smallsetminus \bigcup_{j=1}^{n} (f_{i_j} = 0)$.

We give a compactification of X which, in turn, gives a *tropical compactification* of $Z \cap \mathbb{T}^{n+1}$ with CNC boundary via the map $\mathbf{f}$.

First, we naively compactify X inside $\mathbb{P}^2$. The components of the boundary divisor are $D_{i_j} = (f_{i_j}^h(\omega, \lambda, u) = 0)$ and $D_\infty = (u = 0)$, where $f_{i_j}^h$ is the homogenization of f_{i_j} with respect to the new variable u. We encounter three types of singularities: the origin, the point $(0 : 1 : 0)$ at infinity, and isolated singularities in $\mathbb{T}^2$. We resolve them by blowing up these points and contracting divisors with negative self-intersection (encoded by superfluous bivalent nodes), in a way that preserves the CNC condition. The resolutions diagrams will *precisely* be the graphs in Figure 1, where h_1 corresponds to the divisor D_∞. The nodes E_{i_j} $(1 \leq j \leq n-1)$ and h_{i_j} $(2 \leq j \leq n-1)$ will correspond to exceptional divisors. All intersection multiplicities will equal one, so to compute the multiplicities of the edges in $\mathcal{T}Z$ involving nodes h_{i_j} or E_{i_j}, we only need to calculate indices of suitable lattices associated to these edges.

We now describe the resolution process at each one of our three types of singular points. At the origin, all curves D_{i_j} (except for D_0) intersect and they are tangential to each other. For any j, the strict transform of a given D_{i_j}, after a single blow-up, equals D_{i_j-1}, so we can resolve this singularity after i_{n-1}-blow-ups. The exceptional divisors are labeled E_k $(1 \leq k \leq i_{n-1})$ and all of them give bivalent nodes in the resolution diagram, except for the $n-1$ nodes E_{i_j}. We eliminate the bivalent nodes by contraction. By induction, we see that the valuation of each exceptional divisor is the integer vectors E_{i_j} from Theorem 3.

At infinity, the resolution process is more delicate. Here, the singular point $p = (0 : 1 : 0)$ corresponds to the intersection of D_∞ and all divisors D_{i_j} with $i_j \geq 2$. However, we know that p is a singular point of all prime divisors D_{i_j}. Therefore, we first need to perform a blow-up to smooth them out. More precisely, if π denotes this blow-up and H is the exceptional divisor, we obtain $\pi^*(D_{i_j}) = D_{i_j} + (i_j - 1)H$, $\pi^*(D_\infty) = D'_\infty + H$, where $H = (t = 0)$, and $D'_{i_j} = (\omega - t^{i_j-1} = 0)$, $D'_\infty = (w = 0)$ are the strict transforms. Therefore, the new setting is very similar to the one we described before for the singularity of the boundary D at the origin. The main difference with the resolution at the origin is that along the series of blow-ups, the strict transform of H will continue to be tangential to the divisors intersecting at a "fat point", whereas H was not present in the resolution at the origin. All exceptional divisors will be denoted by h_k $(k = 2, \dots, i_n)$ and again we only keep the non-bivalent nodes h_{i_j} $(2 \leq j \leq n)$ after appropriate contractions. For simplicity, we denote the strict transform of D_∞ by h_1. At the end of the resolution process H gets contracted, explaining why we do not see it in the resolution diagram (Figure 1). As expected, we recover the integer vectors h_{i_j} from Theorem 3.

Finally, we come to multiple intersections among the divisors D_{i_j} in $\mathbb{T}^2$. If (λ, ω) satisfies $f_{i_j} = \lambda - \omega^{i_j} = 0$ and $f_{i_k} = \lambda - \omega^{i_k} = 0$, then $\omega^{i_j} = \lambda = \omega^{i_k}$, so ω is a primitive r-th root of unity for some $r \mid (i_k - i_j)$. Alternatively, $i_j \equiv i_k \equiv s \pmod r$, $\omega = e^{2\pi i p/r}$ and $\lambda = \omega^s$ for p coprime to r. All other curves $(f_{i_l} = 0)$ with $i_l \equiv s \pmod r$ will also meet at (λ, ω). We represent this crossing point (λ, ω) by $x_{p,r,s}$ and the index set of curves meeting at $x_{p,r,s}$ by $\underline{a}_{r,s}$, or $\underline{a}$ for short. That is,

$$x_{p,r,s} = (e^{2\pi i p s/r}, e^{2\pi i p/r}), \quad \underline{a} = \underline{a}_{r,s} := \{i_j | i_j \equiv s \pmod r\}.$$

Furthermore, the curves $D_{i_j} = (f_{i_j} = 0)$ meeting at $x_{p,r,s}$ intersect transversally.

If three or more curves meet at a point, we blow up this point to separate the curves. To simplify notations, we also blow up crossings with $|\underline{a}| = 2$. After a single blow-up at each crossing point $x_{p,r,s}$ we obtain a new divisor $F_{\underline{a}, x_{p,r,s}}$ (the exceptional divisor associated to the point $x_{p,r,s}$) which intersects the proper transform of all D_{i_j} normally, for $j \in \underline{a}$. After studying the coefficient of $F_{\underline{a}, x_{p,r,s}}$ in the pull-back of each character of the torus $\mathbb{T}^{n+1}$ under the map $\mathbf{f}$, we get the node $F_{\underline{a}} = [F_{\underline{a}, x_{p,r,s}}] = \sum_{i_j \in \underline{a}} e_j$, as desired. The resolution diagram will correspond to the graph in the right-hand side of Figure 1.

Finally, we use Theorem 8 to compute the multiplicity of the edge $F_{\underline{a}}D_{i_j}$ in $\mathcal{T}Z$. All summands equal one and so the multiplicity is just the number of such summands, that is, the number of points $x_{p,r,s}$ such that $F_{\underline{a}} = [F_{\underline{a},x_{p,r,s}}]$. This number equals the sum $\sum_l \varphi(l)$ over all common differences l giving $\underline{a}$.

6 The tropical secant graph is a Hadamard product

In this section, we use the master graph to effectively compute the tropicalization of the first secant variety of a monomial projective curve C. Without loss of generality, we may assume that the curve is parameterized as $(1 : t^{i_1} : \ldots : t^{i_n})$, where $0 < i_1 < \ldots < i_n$ are coprime integers. By definition,

$$Sec^1(C) = \overline{\{a \cdot p + b \cdot q : a, b \in \mathbb{C}, p, q \in C\}} \subset \mathbb{P}^n.$$

As discussed in Section 3, tropicalizations are toric in nature. Thus, for the rest of this section, instead of looking at the projective varieties C and $Sec^1(C)$, we study the corresponding very affine varieties which are intersections of their affine cones in $\mathbb{R}^{n+1}$ with the algebraic torus $\mathbb{T}^{n+1}$. To simplify notation, we will also denote them by C and $Sec^1(C)$ in a way that is clear from the context. Tropicalizations of projective varieties and their corresponding very affine varieties are the same.

We parameterize this secant variety by the *secant map* $\phi\colon \mathbb{T}^4 \to \mathbb{T}^{n+1}$, $\phi(a, b, s, t) = (as^{i_k} + bt^{i_k})_{0 \leq k \leq n}$. After a monomial change of coordinates $b = -\lambda a$ and $t = \omega s$, this map can be written as

$$\phi(a, s, \omega, \lambda) = \left(as^{i_k}\left(\omega^{i_k} - \lambda\right)\right)_{0 \leq k \leq n}.$$

From this observation, it is natural to consider the Hadamard product of subvarieties of tori:

Definition 9 *Let $X, Y \subset \mathbb{T}^N$ be two subvarieties of tori. The* Hadamard product *of X and Y equals*

$$X \centerdot Y = \overline{\{(x_1y_1, \ldots, x_Ny_N) \mid x \in X, y \in Y\}} \subset \mathbb{T}^N.$$

From the construction, we get the following characterization of our secant variety:

Proposition 10 *The first secant variety $Sec^1(C) \subset \mathbb{R}^{n+1}$ of the monomial curve C parameterized by $t \mapsto (1 : t^{i_1} : \ldots : t^{i_n}) \in \mathbb{P}^n$ equals $C \centerdot Z \subset \mathbb{T}^{n+1}$ where Z is the surface parameterized by $(\lambda, \omega) \mapsto (1 - \lambda, \omega^{i_1} - \lambda, \ldots, \omega^{i_n} - \lambda)$.*

We now explain the relationship between Hadamard products and their tropicalization:

Proposition 11 *(Corollary 13, Cueto et al., 2010) Given C, Z as in Proposition 10, then as* sets

$$\mathcal{T}Sec^1(C) = \mathcal{T}C + \mathcal{T}Z, \tag{1}$$

where the sum on the (RHS) denotes the Minkowski sum in $\mathbb{R}^{n+1}$.

As we mentioned earlier, $\mathcal{T}C = \mathbb{R}\langle \mathbf{1}, (0, i_1, \ldots, i_n)\rangle$ with constant weight one. By construction, the lineality space of $\mathcal{T}Z \subset \mathbb{R}^{n+1}$ is the origin, and the lineality space of $\mathcal{T}Sec^1(C) \subset \mathbb{R}^{n+1}$ equals $\mathcal{T}C$.

As occurs in general with Hadamard products and their tropicalizations, the right-hand side of (1) has no canonical fan structure. Some maximal cones can be subdivided, whereas others can be merged into bigger cones. Hence, we present this set as a collection of four-dimensional weighted cones in $\mathbb{R}^{n+1}$ obtained as a Minkowski sum of maximal cones in $\mathcal{T}C$ and $\mathcal{T}Z$. The multiplicity at a regular point would simply be the sum of multiplicities of all cones in the collection containing it. Moreover, we will be able to express this number in terms of the multiplicities in $\mathcal{T}Z$, using the following result from (Sturmfels and Tevelev, 2008) that shows the interplay between maps on tori and their tropicalization. Let $\alpha\colon \mathbb{T}^r \to \mathbb{T}^N$ be a homomorphism of tori, that is, a monomial map whose exponents are encoded in a matrix $A \in \mathbb{Z}^{N \times r}$.

Theorem 12 *(Sturmfels and Tevelev, 2008) Let $V \subset \mathbb{T}^r$ be a subvariety. Then $\mathcal{T}(\alpha(V)) = A(\mathcal{T}V)$.*

Moreover, if α induces a generically finite morphism of degree δ on V, then the multiplicity of $\mathcal{T}(\alpha(V))$ at a regular point w is

$$m_w = \frac{1}{\delta} \cdot \sum_v m_v \cdot \textit{ index } (\mathbb{L}_w \cap \mathbb{Z}^N : A(\mathbb{L}_v \cap \mathbb{Z}^r)), \tag{2}$$

where the sum is over all points $v \in \mathcal{T}V$ with $Av = w$. We also assume that the number of such v is finite, and that all of them are regular in $\mathcal{T}V$. In this setting, $\mathbb{L}_v, \mathbb{L}_w$ denote the linear spans of neighborhoods of $v \in \mathcal{T}V$ and $w \in A(\mathcal{T}V)$ respectively.

The key fact in the computation of multiplicities for $\mathcal{T}Sec^1(C)$ is that we can express the Hadamard product in terms of the monomial map $\alpha\colon \mathbb{T}^{2n+2} \to \mathbb{T}^{n+1}$ given by the matrix $A = (I_{n+1} \mid I_{n+1}) \in \mathbb{Z}^{(n+1)\times 2(n+1)}$. The subvariety $V \subset \mathbb{T}^{2n+2}$ is the Cartesian product $C \times Z$, where we consider each surface inside the torus. From (Cueto et al., 2010), we have $\mathcal{T}V = \mathcal{T}(C \times Z) = \mathcal{T}C \times \mathcal{T}Z$ and the multiplicity m_v at a regular point $v = (c, z)$ of V equals m_z. By dimension arguments, we see that α is generically finite when restricted to V, so we can use formula (2) to compute multiplicities in $\mathcal{T}Sec^1(C)$.

Lemma 13 *For $V = C \times Z$ and α as above, the generic fiber of $\alpha_{|V}$ has size 2, hence $\delta = 2$.*

Next, we compute the fiber of a regular point w in $\mathcal{T}(\alpha(V))$ under the linear map A. The strategy will be to pick all possible pairs of maximal cones σ, σ' in $\mathcal{T}Z$ and to compute the dimension of $(\mathbb{R}\sigma + \mathcal{T}C) \bigcap (\mathbb{R}\sigma' + \mathcal{T}C)$. If this dimension is strictly less than four, then we know that generic points in $\mathcal{T}C \times \sigma$ and $\mathcal{T}C \times \sigma'$ belong to different fibers of A. If it equals four, we compute the fiber of A at any point in the intersection. In particular, we conclude:

Lemma 14 *(i) The cones $\langle D_0, h_{i_1}\rangle + \mathcal{T}C$, $\langle F_{\{0,i_1,\dots,i_n\}}, D_{i_j}\rangle + \mathcal{T}C$ ($0 \le j \le n$), $\langle D_{i_n}, E_{i_{n-1}}\rangle + \mathcal{T}C$ and $\langle D_{i_n}, h_{i_{n-1}}\rangle + \mathcal{T}C$ are not maximal, so we disregard them together with the node $F_{\{0,i_1,\dots,i_n\}}$.*

(ii) For all $1 \le j \le n-2$, we have equalities $\langle E_{i_j}, D_{i_j}\rangle + \mathcal{T}C = \langle h_{i_j}, D_{i_j}\rangle + \mathcal{T}C$ and $\langle E_{i_j}, E_{i_{j+1}}\rangle + \mathcal{T}C = \langle h_{i_j}, h_{i_{j+1}}\rangle + \mathcal{T}C$ because $E_{i_j} \equiv h_{i_j}$ modulo $\mathcal{T}C$. Hence, we disregard all nodes h_{i_j}.

(iii) $i_1 \cdot F_{\underline{e}} = E_{i_1}$ and $(i_n - i_{n-1}) \cdot F_{\underline{b}} \equiv E_{i_{n-1}}$ modulo $\mathcal{T}C$, where $\underline{e} = \{i_1, \dots, i_n\}$ and $\underline{b} = \{0, i_1, \dots, i_{n-1}\}$. Thus, the maximal cones $\mathbb{R}\langle F_{\underline{e}}, D_{i_1}\rangle + \mathcal{T}C$ and $\mathbb{R}\langle E_{i_1}, D_{i_1}\rangle + \mathcal{T}C$ coincide, as well as $\mathbb{R}\langle F_{\underline{b}}, D_{i_{n-1}}\rangle + \mathcal{T}C$ and $\mathbb{R}\langle E_{i_{n-1}}, D_{i_{n-1}}\rangle + \mathcal{T}C$.

(iv) All other fibers have size one.

As a consequence of this lemma, in numerical examples we will identify the nodes E_{i_1} and $F_{\underline{e}}$, as well as $E_{i_{n-1}}$ and $F_{\underline{b}}$. In this identification, the nodes $F_{\underline{e}}$ and $F_{\underline{b}}$ are removed, and the edges adjacent to the nodes $F_{\underline{e}}$ and $F_{\underline{b}}$ are added to those of E_{i_1} and $E_{i_{n-1}}$. We also merge the corresponding edges $E_{i_1}D_{i_1}$ and $F_{\underline{e}}D_{i_1}$ (resp. $E_{i_{n-1}}D_{i_{n-1}}$ and $F_{\underline{b}}D_{i_{n-1}}$) in the tropical secant graph, assigning the sum of their weights to the new edge.

The indices involved in (2) are calculated as follows. Let $l_1 = \mathbf{1}$ and $l_2 = (0, i_1, \dots, i_n)$ be the generators of $\mathcal{T}C$. For each edge of $\mathcal{T}Z$, we pick its two end points x, y. The index in (2) associated to a point $v \in \mathcal{T}C + \mathbb{R}_{\ge 0}\langle x, y\rangle \subset \mathcal{T}C + \mathcal{T}Z$ is the quotient of the gcd of the 4-minors of the matrix $(x \mid y \mid l_1 \mid l_2)$ by the gcd of the 2-minors of the matrix $(x \mid y)$. These gcd's are computed as the product of the nonzero diagonal elements of the Smith normal form of each matrix. Here is our main result:

Definition 15 *The* tropical secant graph *is a weighted subgraph of the master graph in* $\mathbb{R}^{n+1}$*, with nodes:*

(i) $D_{i_j} = e_j := (0,\dots,0,1,0,\dots,0) \qquad (0 \le j \le n)$,

(ii) $E_{i_j} = (0, i_1, \dots, i_{j-1}, i_j, \dots, i_j) = \sum_{k<j} i_k \cdot e_k + i_j \cdot (\sum_{k \ge j} e_k) \qquad (1 \le j \le n-1)$,

(iii) $F_{\underline{a}} = \sum_{i_j \in \underline{a}} e_j$ *where* $\underline{a} \subsetneq \{0, i_1, \dots, i_n\}$ *varies among all proper subsets containing at least two elements that are obtained from an arithmetic progression.*

The edges are a subset of the edges of the master graph. Their positive weights are assigned as follows:

(i) $m_{E_{i_j}, E_{i_{j+1}}} = \gcd(i_1, \dots, i_j) \gcd\limits_{j<t<n} (i_n - i_t) \qquad (1 \le j \le n-2)$,

(ii) $m_{D_{i_j}, E_{i_j}} = \gcd\big(\gcd(i_1, \dots, i_{j-1}) \gcd\limits_{j<s\le n} (i_s - i_j)\,;\, \gcd\limits_{0 \le k < j} (i_j - i_k) \gcd(i_{j+1}, \dots, i_n)\big) \quad (1 \le j \le n-1)$,

(iii) $m_{F_{\underline{a}}, D_{i_j}} = \frac{1}{2} \sum\limits_r \varphi(r) \cdot \gcd\big(\gcd\limits_{i_l, i_k \notin \underline{a}} (|\, i_l - i_k\,|)\,;\, \gcd\limits_{\substack{i_l, i_k \in \underline{a} \\ l,k \ne j}} (|\, i_l - i_k\,|)\big)$ *($i_j \in \underline{a}$, where the sum runs over all common differences r of arithmetic progressions giving the subset $\underline{a}$).*

(By convention, a gcd over an empty set of indices is taken to be 0.)

Theorem 16 *Given a monomial curve C with primitive exponent vector $(0, i_1, \dots, i_n)$, $0 = i_0 < i_1 < \dots < i_n$, the tropicalization of the first secant variety of C can be characterized set-theoretically as a collection of 4-dimensional weighted cones (with no fan structure). Each cone has a 2-dimensional lineality space with basis given by the* intrinsic lattice $\Lambda = \langle (1,\dots,1), (0, i_1, \dots, i_n) \rangle$. *The collection is obtained as the cone from the subspace $\mathbb{R} \otimes_{\mathbb{Z}} \Lambda$ over the tropical secant graph, preserving all weights.*

7 The Newton polytope of the secant graph for $\mathbb{P}^4$

In this section, we focus our attention on the *inverse problem*. That is, given the tropical variety of an irreducible hypersurface, we wish to recover its defining equation. A first step towards a satisfactory answer would consist of computing the *Newton polytope* of the defining equation $f = \sum_a c_a \underline{x}^a$, i.e. the convex hull of integer vectors a such that $\underline{x}^a$ appears with a nonzero coefficient in f. This will let us find the defining equation via interpolation.

We now explain the connection between $\mathcal{T}(f)$ and NP(f) for an irreducible polynomial f in $n+1$ variables defined over $\mathbb{C}$. For a vector $w \in \mathbb{R}^{n+1}$, the initial form $\mathrm{in}_w(f)$ is a monomial if and only if w is in the interior of a maximal cone (chamber) of the normal fan of NP(f). The tropical variety of the hypersurface $(f = 0)$ is the union of codimension one cones of the normal fan of NP(f). The multiplicity of a maximal cone in $\mathcal{T}(f)$ is the lattice length of the edge of NP(f) normal to that cone.

A construction for the Newton polytope NP(f) from its normal fan $\mathcal{T}(f)$ equipped with multiplicities was developed in Dickenstein et al. (2007). We describe this *ray-shooting algorithm* in Theorem 17:

Theorem 17 *Suppose $w \in \mathbb{R}^{n+1}$ is a generic vector so that the ray $(w + \mathbb{R}_{>0}\, e_i)$ intersects $\mathcal{T}(f)$ only at regular points of $\mathcal{T}(f)$, for all i. Let $\mathcal{P}^w$ be the vertex of the polytope $\mathcal{P} = NP(f)$ that attains the maximum of $\{w \cdot x : x \in \mathcal{P}\}$. Then the i^{th} coordinate of $\mathcal{P}^w$ equals $\sum_v m_v \cdot |l_i^v|$, where the sum is taken over all points $v \in \mathcal{T}(f) \cap (w + \mathbb{R}_{>0} e_i)$, m_v is the multiplicity of v in $\mathcal{T}(f)$, and l_i^v is the i^{th} coordinate of the primitive integral normal vector l^v to the maximal cone in $\mathcal{T}(f)$ containing v.*

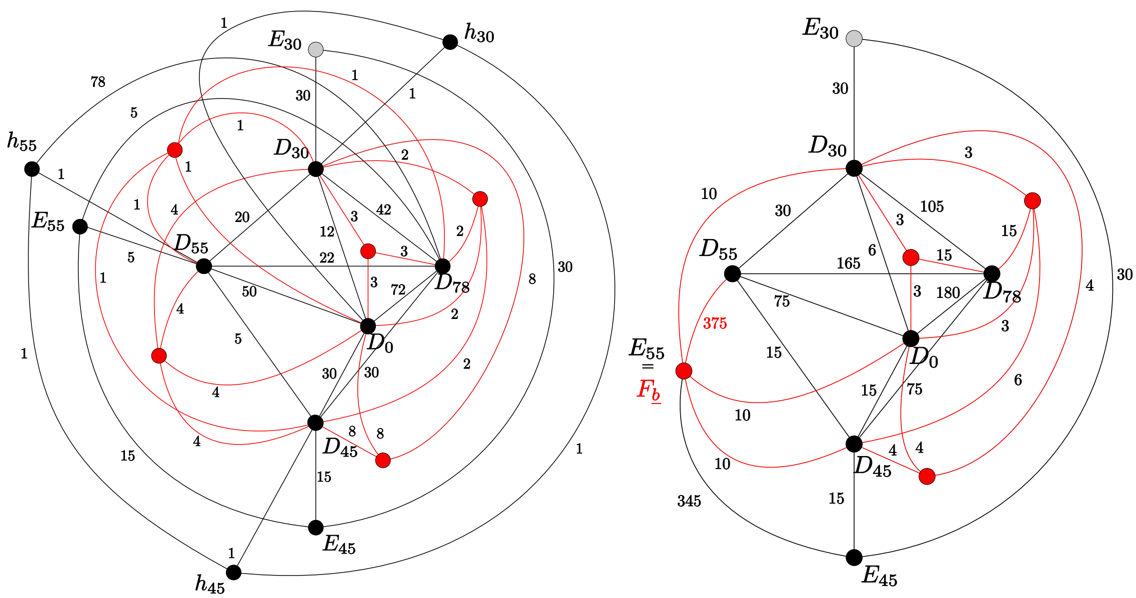


Fig. 2: The master graph and the tropical secant graph of the monomial curve $(1 : t^{30} : t^{45} : t^{55} : t^{78})$.

Note that we do not need a fan structure on $\mathcal{T}(f)$ to use Theorem 17. A description of $\mathcal{T}(f)$ as a set, together with a way to compute the multiplicities at regular points, gives us enough information to compute vertices of NP(f) in any generic directions. Computing a single vertex using Theorem 17 will give us the multidegree of f with respect to the grading given by the intrinsic lattice Λ from Theorem 16.

The entire polytope NP(f) can be computed by iterating the ray-shooting algorithm with different objective vectors (one per chamber). A method to choose these vectors appropriately was developed in (Algorithm 2, Cueto et al., 2010): the *walking algorithm*. The core of the method is to keep track of the cones that we meet while ray-shooting from a given objective vector, to use the list of such cones to walk from chamber to chamber in the normal fan of NP(f), picking objective vectors along the way, and to repeat the shooting algorithm with these new vectors. We illustrate these methods with an example.

Example 18 *The first secant variety of the monomial curve* $t \mapsto (1 : t^{30} : t^{45} : t^{55} : t^{78})$ *in* $\mathbb{P}^4$ *is known to be a hypersurface of degree 1820 (Example 3.3, Ranestad, 2006). We use geometric tropicalization to compute the tropicalization of this variety. By Theorems 6 and 16, we construct the two graphs in Figure 2: the leftmost picture corresponds to the master graph, whereas the rightmost picture is the tropical secant graph. The ten nodes in the tropical secant graph have coordinates* $D_0 = e_0$, $D_{30} = e_1$, $D_{45} = e_2$, $D_{55} = e_3$, $D_{78} = e_4$, $E_{30} = (0,30,30,30,30)$, $E_{45} = (0,30,45,45,45)$, $F_{\{0,30,45,55\}} \equiv E_{55} = (0,30,45,55,55)$, $F_{\{0,30,78\}} = (1,1,0,0,1)$, $F_{\{0,30,45,78\}} = (1,1,1,0,1)$, *and* $F_{\{0,30,45\}} = (1,1,1,0,0)$. *The master graph has the five extra nodes* $h_{30} = (-30,-30,-45,-55,-78)$, $h_{45} = (-45,-45,-45,-55,-78)$, $h_{55} = (-55,-55,-55,-55,-78)$, $F_{\{0,30,45,55,78\}} = (1,1,1,1,1)$, *and* $F_{\{0,30,45,55\}} = (1,1,1,1,0)$. *The unlabeled nodes in Figure 2 indicate nodes of type* $F_{\underline{a}}$, *where the subset* $\underline{a}$ *consists of the indices of all nodes* D_{i_j} *adjacent to the unlabeled node. Notice that the nodes* E_{55} *and* $F_{\underline{b}}$ *coincide in the tropical secant graph, as predicted by Lemma 14.*

Finally, we apply the ray-shooting and walking algorithms to recover the Newton polytope of this hypersurface. Its multidegree with respect to the lattice $\Lambda = \mathbb{Z}\langle \mathbf{1}, (0,30,45,55,78)\rangle$ *is* $(1\,820, 76\,950)$. *The polytope has 24 vertices and* f*-vector* $(24,38,16)$. *Using* `LattE` *we see that it contains* $7\,566\,849$ *lattice points, which gives an upper bound for the number of monomials in the defining equation.*

The implicitization methods discussed in this section can be generalized to monomial curves in higher dimensional projective spaces, where the first secant has no longer codimension one. In this case, one can recover the *Chow polytope* of the secant variety by a natural generalization of the ray-shooting method: the *orthant-shooting algorithm* (Theorem 2.2, Dickenstein et al., 2007). Instead of shooting rays, we shoot orthants (i.e. cones spanned by vectors in the canonical basis of $\mathbb{R}^{n+1}$) of dimension equal to the codimension of our variety. A formula similar to the one described in Theorem 17 will give us the vertex of the Chow polytope associated to the input objective vector. However, an analog to the *walking algorithm* still needs to be developed, since there is, a priori, no canonical way of ordering the intersection points for walking along the complement of the tropical variety. We hope to pursue this direction in the near future.

Acknowledgements

We thank Bernd Sturmfels for suggesting this problem to us and for inspiring discussions. We also thank Melody Chan, Alex Fink and Jenia Tevelev for fruitful conversations.

References

R. Bieri and J. Groves. The geometry of the set of characters induced by valuations. *J. Reine Angew. Math.*, 347:168–195, 1984. ISSN 0075-4102.

D. Cox and J. Sidman. Secant varieties of toric varieties. *J. Pure Appl. Algebra*, 209(3):651–669, 2007. ISSN 0022-4049.

M. A. Cueto. *Tropical Implicitization*. PhD thesis, University of California - Berkeley, 2011.

M. A. Cueto, E. Tobis, and J. Yu. An implicitization challenge for binary factor analysis. Contribution MEGA'09 (Barcelona, Spain). Accepted for publication in *J. Symbolic Comput.*, Special Issue, 2010.

M. Develin. Tropical secant varieties of linear spaces. *Discrete Comput. Geom.*, 35(1):117–129, 2006. ISSN 0179-5376.

A. Dickenstein, E. M. Feichtner, and B. Sturmfels. Tropical discriminants. *J. Amer. Math. Soc.*, 20(4): 1111–1133 (electronic), 2007. ISSN 0894-0347.

P. Hacking, S. Keel, and J. Tevelev. Stable pair, tropical, and log canonical compactifications of moduli spaces of del Pezzo surfaces. *Invent. Math.*, 178(1):173–227, 2009. ISSN 0020-9910.

K. Ranestad. The degree of the secant variety and the join of monomial curves. *Collect. Math.*, 57(1): 27–41, 2006. ISSN 0010-0757.

B. Sturmfels and J. Tevelev. Elimination theory for tropical varieties. *Math. Res. Lett.*, 15(3):543–562, 2008. ISSN 1073-2780.

B. Sturmfels, J. Tevelev, and J. Yu. The Newton polytope of the implicit equation. *Mosc. Math. J.*, 7(2): 327–346, 351, 2007. ISSN 1609-3321.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 553–564

A note on moments of derivatives of characteristic polynomials

Paul-Olivier Dehaye

ETH Zürich, Department of Mathematics, 8092 Zürich, Switzerland, pdehaye@math.ethz.ch

---

Abstract. We present a simple technique to compute moments of derivatives of unitary characteristic polynomials. The first part of the technique relies on an idea of Bump and Gamburd: it uses orthonormality of Schur functions over unitary groups to compute matrix averages of characteristic polynomials. In order to consider derivatives of those polynomials, we here need the added strength of the Generalized Binomial Theorem of Okounkov and Olshanski. This result is very natural as it provides coefficients for the Taylor expansions of Schur functions, in terms of shifted Schur functions. The answer is finally given as a sum over partitions of functions of the contents. One can also obtain alternative expressions involving hypergeometric functions of matrix arguments.

Résumé. Nous introduisons une nouvelle technique, en deux parties, pour calculer les moments de dérivées de polynômes caractéristiques. La première étape repose sur une idée de Bump et Gamburd et utilise l'orthonormalité des fonctions de Schur sur les groupes unitaires pour calculer des moyennes de polynômes caractéristiques de matrices aléatoires. La deuxième étape, qui est nécessaire pour passer aux dérivées, utilise une généralisation du théorème binomial due à Okounkov et Olshanski. Ce théorème livre les coefficients des séries de Taylor pour les fonctions de Schur sous la forme de "shifted Schur functions". La réponse finale est donnée sous forme de somme sur les partitions de fonctions des contenus. Nous obtenons aussi d'autres expressions en terme de fonctions hypergéométriques d'argument matriciel.

Keywords: random matrix theory, hook-content formula, moment of characteristic polynomials, shifted Schur function, generalized Pochhammer symbol, hypergeometric function of matrix argument

---

1 Introduction

We take for the characteristic polynomial of a $N \times N$ unitary matrix U

$$Z_U(\theta) \quad := \quad \prod_{j=1}^{N} \left(1 - e^{\mathrm{i}(\theta_j - \theta)}\right), \tag{1}$$

where the θ_js are the eigenangles of U and set

$$V_U(\theta) \quad := \quad e^{\mathrm{i}N(\theta+\pi)/2} e^{-\mathrm{i}\sum_{j=1}^{N}\theta_j/2} Z_U(\theta). \tag{2}$$

It is easily checked that for real θ, $V_U(\theta)$ is real and that $|V_U(\theta)|$ equals $|Z_U(\theta)|$.

For k and r integers, with $0 \leq r \leq 2k$, we will investigate the averages (with respect to Haar measure)

$$(\mathcal{M})_N(2k, r) \quad := \quad \left\langle |Z_U(0)|^{2k} \left(\frac{Z_U'(0)}{Z_U(0)}\right)^r \right\rangle_{\mathrm{U}(N)}, \tag{3}$$

$$|\mathcal{V}|_N(2k, r) \quad := \quad \left\langle |V_U(0)|^{2k} \left|\frac{V_U'(0)}{V_U(0)}\right|^r \right\rangle_{\mathrm{U}(N)}. \tag{4}$$

This is notation we already used in [Deh08] (where a $|\mathcal{M}|$ was also present but is not needed here), and is only notation in the LHS: the $(\mathcal{M})$ and $|\mathcal{V}|$ are thus meant each as *one* symbol and are supposed to mnemotechnically remind the reader of what is in the RHS. We immediately state the following easy lemma.

Lemma 1 *For $k \geq h$ non-negative integers, we have the relation*

$$|\mathcal{V}|_N(2k, 2h) \quad = \quad \sum_{i=0}^{2h} \binom{2h}{i} (\mathcal{M})_N(2k, i) \left(\frac{\mathrm{i}N}{2}\right)^{2h-i}. \tag{5}$$

Proof: This is available (in the same notation) in [Deh08] and a consequence of Equation (2), which leads to the polynomial relations between $\frac{Z_U'(0)}{Z_U(0)}, \frac{V_U'(0)}{V_U(0)}$ and their norms. These relations give

$$|\mathcal{V}|_N(2k, 2h) \quad = \quad \sum_{j=0}^{h} \binom{h}{j} \left(\frac{-N^2}{4}\right)^{h-j} \sum_{l=0}^{j} (\mathrm{i}N)^{j-l} \binom{j}{l} (\mathcal{M})_N(2k, j+l), \tag{6}$$

which is easily deduced from [Deh08]. □

We are actually more concerned with the renormalizations

$$(\mathcal{M})(2k, r) \quad = \quad \lim_{N\to\infty} \frac{(\mathcal{M})_N(2k, r)}{N^{k^2+r}} \tag{7}$$

$$|\mathcal{V}|(2k, r) \quad = \quad \lim_{N\to\infty} \frac{|\mathcal{V}|_N(2k, r)}{N^{k^2+r}}. \tag{8}$$

Theorem 2 will show that these normalizations are appropriate.

The random matrix theory problem of evaluating $(\mathcal{M})(2k, r)$ and $|\mathcal{V}|(2k, r)$ has applications in number theory (see [Deh08] for a more detailed exposition of these ideas). Indeed, these values are related to the factor $g(k, h)$ in the formula

$$\lim_{T\to\infty} \frac{1}{T} \frac{1}{(\log \frac{T}{2\pi})^{k^2+2h}} \int_0^T \left|\zeta\left(\frac{1}{2} + \mathrm{i}t\right)\right|^{2k-2h} \left|\zeta'\left(\frac{1}{2} + \mathrm{i}t\right)\right|^{2h} \mathrm{d}t \quad = \quad a(k)g(k, h), \tag{9}$$

where $a(k)$ is a (known) factor defined as a product over primes. This, along with a discrete moment version due to Hughes, is the principal underlying motivation for the all the random matrix theory analysis that occurs in [HKO00, Hug01, Hug05, Mez03, CRS06, FW06].

Another application is tied to the work of Hall [Hal02a, Hal02b, Hal04, Hal08], where results on the objects studied here can be used to hint towards optimizations of rigorous arguments in number theory, and serve as (conjectural) inputs on theorems there. The number theory statements concern average spacings between zeroes of the Riemann zeta function. The works of Steuding [Ste05] and Saker [Sak09] follow similar approaches.

In [Deh08], the author investigated for fixed r ratios of those quantities and established they were a rational function. We reprove this result here, but with a much simpler method leading to a much simpler result. In particular, rationality of the RHS is transparent from the following statement (definitions are given in Section 2), since the RHS sums are finite sums over partitions of r:

Theorem 2 *For $0 \leq r \leq 2k$, with $r, k \in \mathbb{N}$,*

$$\frac{(\mathcal{M})_N(2k,r)}{(\mathcal{M})_N(2k,0)}\mathfrak{i}^r = \sum_{\mu \vdash r} \frac{r!}{h_\mu^2} \frac{(N \uparrow \mu)((-k) \uparrow \mu)}{(-2k) \uparrow \mu}, \tag{10}$$

$$\frac{(\mathcal{M})(2k,r)}{(\mathcal{M})(2k,0)}\mathfrak{i}^r = \sum_{\mu \vdash r} \frac{r!}{h_\mu^2} \frac{k \uparrow \mu}{(2k) \uparrow \mu}, \tag{11}$$

while the denominators on the left are known ([BG06],[KS00]):

$$(\mathcal{M})_N(2k,0) = \mathfrak{s}_{\langle N^k \rangle}\left(\{1\}^{2k}\right) = \frac{G(N+2k+1)G(N+1)}{G(N+k+1)^2} \frac{G(k+1)^2}{G(2k+1)} \tag{12}$$

and

$$(\mathcal{M})(2k,0) = \lim_{N\to\infty} \frac{\mathfrak{s}_{\langle N^k \rangle}\left(\{1\}^{2k}\right)}{N^{k^2}} = \frac{G(k+1)^2}{G(2k+1)}, \tag{13}$$

where $G(\cdot)$ is the Barnes G-function.

We now briefly discuss the technique used to obtain this theorem. The first idea will be similar to an idea of Bump and Gamburd [BG06] of using orthonormality of Schur functions to efficiently compute matrix averages. The new idea here is to combine this with the Generalized Binomial Theorem (23) of Okounkov and Olshanski [OO97] in order to obtain information about the moments of derivatives instead of moments of polynomials directly.

This paper is structured as follows. We give in Section 2 the basic definitions needed. In Section 3, we explain the Generalized Binomial Theorem. We prove Theorem 2 in Section 4. We use this result to deduce in Section 5 further properties of the rational functions obtained. We present in Section 6 an alternative interpretation of these results in terms of hypergeometric functions of a matrix argument. Finally, we announce briefly in Section 7 further results.

2 Definitions

Since Theorem 2 presents its result as a sum over partitions, we first need to define some classical objects associated to them. We follow conventions of [Sta99] throughout.

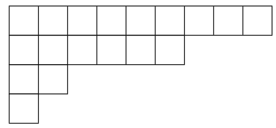


Fig. 1: The Ferrers diagram of partition $(9, 6, 2, 1)$.

0	1	2	3	4	5	6	7	8
-1	0	1	2	3	4			
-2	-1							
-3								

12	10	8	7	6	5	3	2	1
8	6	4	3	2	1			
3	1							
1								

Fig. 2: The contents and hook lengths of the partition $(9, 6, 2, 1)$. Its hook number is thus 4180377600.

Partitions are weakly decreasing sequences $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_{l(\lambda)}$ of positive integers, its *parts*. The integer $l = l(\lambda)$ is called the *length* of the partition λ. We call the sum $\sum_i \lambda_i$ of its parts the *size* $|\lambda|$ of the partition λ. We sometimes say that λ partitions $|\lambda|$, which is written $\lambda \vdash |\lambda|$. If the partition has k parts of equal size N, we simplify notation to $\langle N^k \rangle$.

We always prefer to think of partitions graphically. To each partition we associate a *Ferrers diagram*, *i.e.* the Young diagram of the partition presented in the English convention. We only give one example (Figure 1) as it should be clear from it how the diagram is constructed: the parts λ_i indicate how many standard boxes to consider on each row.

There exists an involution acting on partitions, which we denote by λ^t. Its action on diagrams amounts to a reflection along the main diagonal.

Partitions can be indexed in many different ways. Indeed, we have already seen that finite sequences of (decreasing) part sizes can be used as an index set. Shifted part lengths are essential for the work of Okounkov and Olshanski underlying Section 3, but we do not need to define that system of coordinates.

Define the *content* of a box $\square$ located at position (i, j) in a partition λ as $c(\square) = j - i$ (see Figure 2). We use this to define the symbol

$$k \uparrow \mu \quad := \quad \prod_{\square \in \mu} (k + c(\square)). \tag{14}$$

This definition leads to $k \uparrow (n) = k(k+1) \cdots (k+n-1)$ and $k \uparrow (1^n) = k(k-1) \cdots (k-n+1)$. We sometimes abbreviate the first $k \uparrow n$ and the second $k \downarrow n$. This is clearly a generalization of the Pochhammer symbol. Indeed, we even adopt the convention that $N \uparrow (-k) = 1/((N+1) \uparrow k)$, which guarantees

$$(N + a - 1) \downarrow (a + b) \quad = \quad (N \uparrow a) \cdot ((N-1) \downarrow b). \tag{15}$$

To get back to the generalization to partitions, we have an immediate relation under conjugation:

$$k \uparrow \mu^t \quad = \quad (-1)^{|\mu|} ((-k) \uparrow \mu). \tag{16}$$

Given a box $\square \in \lambda$, define its hook (set)

$$h_\square = h_{(i,j)} \quad := \quad \{(i,j') \in \lambda : j' \geq j\} \cup \{(i',j) \in \lambda : i' \geq i\}\,. \tag{17}$$

Remark that the box $\square$ itself is in its hook. Define the *hook length* $|h_{(i,j)}|$ as the cardinality of the hook (see Figure 2) and call their product the *hook number* of a partition λ:

$$h_\lambda \quad := \quad \prod_{\square \in \lambda} h_\square. \tag{18}$$

Hook numbers are of importance thanks to the hook length formula of Frame, Robinson and Thrall [FRT54]. This counts the number f_λ of standard tableaux of shape λ, which is also the dimension $\dim \chi^\lambda = \chi^\lambda(1)$ of the character associated to the partition λ for the symmetric group $\mathcal{S}_{|\lambda|}$ (see [Sag01, Sta99]):

$$f_\lambda \quad := \quad \dim \chi^\lambda_{\mathcal{S}_{|\lambda|}} = \frac{|\lambda|!}{h_\lambda}. \tag{19}$$

The last classical combinatorial object we need is the *Schur functions*. To each partition λ we associate an element $\mathfrak{s}_\lambda$ of degree $|\lambda|$ of the ring $\Lambda_{\mathbb{Q}}[X]$ of polynomials symmetric in a countable set of variables $X = \{x_i\}$. We refer the reader to [Bum04] for definitions, and only state a few properties.

Let $r \in \mathbb{N}$, and take a partition λ of size r. One can define a map from χ^λ_N from $\mathrm{U}(N)$ to $\mathbb{C}$ in the following way:

$$\chi^\lambda_N(g) := \mathfrak{s}_\lambda(g) := \mathfrak{s}_\lambda(e^{\mathrm{i}\theta_1}, \cdots, e^{\mathrm{i}\theta_N}, 0, 0, 0, \cdots), \tag{20}$$

with the $e^{\mathrm{i}\theta_j}$ the eigenvalues of $g \in \mathrm{U}(N)$. When $l(\lambda) > N$, $\chi^\lambda_N(g) \equiv 0$, but once $N \geq l(\lambda)$, the $\chi^\lambda_N(g)$ become (different) irreducible characters of $\mathrm{U}(N)$. We thus have the formula

$$\left\langle \chi^\lambda_N, \chi^\mu_N \right\rangle_{\mathrm{U}(N)} = \langle \mathfrak{s}_\lambda(\cdot), \mathfrak{s}_\mu(\cdot) \rangle_{\mathrm{U}(N)} = \begin{cases} 1 & \text{if } \lambda = \mu \text{ and } N \geq |\lambda|, \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

The last formula we need concerns the evaluation of Schur polynomials, at repeated values of the arguments. Denote by $\{a\}^R$ the multiset consisting of the union of R copies of a and countably many copies of 0. The hook-content formula (see [Sta99, Bum04]), a consequence of the Weyl Dimension Formula, states then that

$$\mathfrak{s}_\lambda\left(\{1\}^k\right) \quad = \quad \frac{k \uparrow \lambda}{h_\lambda}. \tag{22}$$

3 Shifted Schur Functions and Generalized Binomial Theorem

The Generalized Binomial Theorem as formulated in [OO97, Theorem 5.1] can be interpreted as a Taylor expansion of the character $\chi^\lambda_n = \mathfrak{s}_\lambda(\cdot)$ of $\mathrm{U}(n)$ around the identity $\mathrm{Id}_{n\times n}$. It says explicitly that

$$\frac{\mathfrak{s}_\lambda(1+x_1, \cdots, 1+x_n)}{\mathfrak{s}_\lambda\left(\{1\}^n\right)} \quad = \quad \sum_{\substack{\mu \\ l(\mu) \leq n}} \frac{\mathfrak{s}^*_\mu(\lambda_1, \cdots, \lambda_n)\mathfrak{s}_\mu(x_1, \cdots, x_n)}{n \uparrow \mu}, \tag{23}$$

where the $\mathfrak{s}^*_\mu$ are shifted Schur functions. Those were introduced by Okounkov and Olshanski in [OO97] and need not be defined here as we only need their values on a very limited set of arguments. This is given by the following lemma, which is very elegant and seems to be new. Note the (almost-)symmetry between k and N.

Lemma 3

$$\begin{aligned}\mathfrak{s}^*_\mu\left(\{N\}^k\right) &= h_\mu \cdot \mathfrak{s}_{\mu^t}\left(\{1\}^N\right)\cdot \mathfrak{s}_\mu\left(\{1\}^k\right) && (24)\\ &= (-1)^{|\mu|}\frac{((-N)\uparrow\mu)(k\uparrow\mu)}{h_\mu} && (25)\end{aligned}$$

Proof: We will need two equations from [OO97]. Equation (11.28) tells us that

$$\mathfrak{s}^*_\mu(x_1,\cdots,x_n) = \det\left[\mathfrak{h}^*_{\mu_i-i+j}(x_1+j-1,\cdots,x_n+j-1)\right]^{R,R}_{i,j=1} \tag{26}$$

for large enough R (it is then stable in R). Equation (11.22) from [OO97] deals precisely with those $\mathfrak{h}^*$:

$$\mathfrak{h}^*_r\left(\{N\}^k\right) = (N\downarrow r)\cdot \mathfrak{h}_r\left(\{1\}^k\right). \tag{27}$$

Combining these equations, we obtain

$$\begin{aligned}\mathfrak{s}^*_\mu\left(\{N\}^k\right) &= \det\left[((N+j-1)\downarrow(\mu_i-i+j))\cdot\mathfrak{h}^*_{\mu_i-i+j}\left(\{1\}^k\right)\right]^{R,R}_{i,j=1} && (28)\\ &= \det\left[(N\uparrow j)((N-1)\downarrow(\mu_i-i))\cdot\mathfrak{h}^*_{\mu_i-i+j}\left(\{1\}^k\right)\right]^{R,R}_{i,j=1} && (29)\\ &= \left(\prod_{i=1}^R((N-1)\downarrow(\mu_i-i))\right)\left(\prod_{j=1}^R(N\uparrow j)\right)\det\left[\mathfrak{h}^*_{\mu_i-i+j}\left(\{1\}^k\right)\right]^{R,R}_{i,j=1} && (30)\\ &= \left(\prod_{i=1}^R((N+i-1)\downarrow\mu_i)\right)\mathfrak{s}_\mu\left(\{1\}^k\right) && (31)\\ &= \left(\prod_{\square\in\mu}N-c(\square)\right)\mathfrak{s}_\mu\left(\{1\}^k\right) && (32)\\ &= h_\mu\mathfrak{s}_{\mu^t}\left(\{1\}^N\right)\mathfrak{s}_\mu\left(\{1\}^k\right). && (33)\end{aligned}$$

The fourth line follows from Equation (15), the fifth from reorganizing a product over rows into a product over boxes, and the sixth from Equation (22). □

We are now ready to launch into the proof of Theorem 2.

4 Proof of Theorem 2

The method of proof will be very similar to the technique presented in [BG06]. In particular, both the Cauchy identity [Bum04]

$$\prod_{i,j} 1 + x_i y_j \quad = \quad \sum_{\lambda} \mathfrak{s}_{\lambda^t}(x_i)\mathfrak{s}_{\lambda}(y_j), \tag{34}$$

where x_i and y_j are finite sets of variables, and the asymptotic orthonormality of the Schur functions will again play a crucial role. However, the power of their technique is now supplemented by the Generalized Binomial Theorem, which will provide for a dramatic simplification of the arguments and results in [Deh08].

Proof of Theorem 2: We have

$$\begin{aligned} \overline{Z_U(0)} &= \prod_{j=1}^{N} \left(1 - e^{-\mathrm{i}\theta_j}\right) && (35) \\ &= \prod_{j=1}^{N} -e^{-\mathrm{i}\theta_j} \left(1 - e^{\mathrm{i}\theta_j}\right) && (36) \\ &= (-1)^N \overline{\det U} Z_U(0) && (37) \end{aligned}$$

and thus

$$\begin{aligned} \overline{Z_U(0)}^k &= (-1)^{kN} \overline{\det U}^k Z_U(0)^k && (38) \\ &= (-1)^{kN} \overline{\mathfrak{s}_{\langle k^N \rangle}(U)} Z_U(0)^k. && (39) \end{aligned}$$

We use the Cauchy Identity from Equation (34) to obtain

$$\begin{aligned} Z_U(a_1)\cdots Z_U(a_r) &= \sum_{\lambda} \mathfrak{s}_{\lambda^t}(U)\mathfrak{s}_{\lambda}\left(-e^{-\mathrm{i}a_1}, \cdots, -e^{-\mathrm{i}a_r}\right) && (40) \\ &= \sum_{\lambda} (-1)^{|\lambda|} \mathfrak{s}_{\lambda^t}(U)\mathfrak{s}_{\lambda}\left(e^{-\mathrm{i}a_1}, \cdots, e^{-\mathrm{i}a_r}\right). && (41) \end{aligned}$$

To the first order in small a, we have $e^{-\mathrm{i}a} \approx 1 - \mathrm{i}a$, so

$$Z_U'(0)^r \quad = \quad \sum_{\lambda} (-1)^{|\lambda|} \mathfrak{s}_{\lambda^t}(U)\ \partial_1 \cdots \partial_r \mathfrak{s}_{\lambda}\left(1 - \mathrm{i}a_1, \cdots, 1 - \mathrm{i}a_r\right)\big|_{a_1 = \cdots = a_r = 0}, \tag{42}$$

where $\partial_i := \partial_{a_i}$.

Putting everything together, we obtain

$$|Z_U(0)|^{2k} \left(\frac{Z_U'(0)}{Z_U(0)}\right)^r = (-1)^{(kN)} \overline{\mathfrak{s}_{\langle k^N \rangle}(U)} \cdot \\ \sum_{\lambda} (-1)^{|\lambda|} \mathfrak{s}_{\lambda^t}(U)\ \partial_1 \cdots \partial_r \mathfrak{s}_{\lambda}\left((2k - r) \times \{1\} \cup \{1 - \mathrm{i}a_1, \cdots, 1 - \mathrm{i}a_r\}\right)\big|_{a_1 = \cdots = a_r = 0}. \tag{43}$$

Just as in the original proof of Bump and Gamburd, orthogonality of the Schur polynomials kills all terms in the sum but one (where λ^t equals $\langle k^N \rangle$) under averaging over $\mathrm{U}(N)$. Hence this simplifies to

$$\left\langle |Z_U(0)|^{2k} \left(\frac{Z_U'(0)}{Z_U(0)}\right)^r \right\rangle_{\mathrm{U}(N)} = \partial_1 \cdots \partial_r \mathfrak{s}_{\langle N^k \rangle}\left((2k-r)\times\{1\} \cup \{1-\mathrm{i}a_1, \cdots, 1-\mathrm{i}a_r\}\right)\Big|_{a_1=\cdots=a_r=0}. \quad (44)$$

This is the perfect opportunity to apply the Generalized Binomial Theorem. We wish to set $n = 2k$, and

$$x_i = \begin{cases} -\mathrm{i}a_i & \text{for} \quad 1 \le i \le r \\ 0 & \text{for} \quad r+1 \le i \le 2k \end{cases} \quad (45)$$

in Equation (23) to get

$$\left\langle |Z_U(0)|^{2k} \left(\frac{Z_U'(0)}{Z_U(0)}\right)^r \right\rangle_{\mathrm{U}(N)} = \mathfrak{s}_{\langle N^k \rangle}\left(\{1\}^{2k}\right) \times (-\mathrm{i})^r \sum_{\mu \vdash r} \frac{\mathfrak{s}^*_\mu\left(\{N\}^k\right)\ \partial_1 \cdots \partial_r \mathfrak{s}_\mu(a_1, \cdots, a_r)\big|_{a_1=\cdots=a_r=0}}{(2k) \uparrow \mu}. \quad (46)$$

The (additional) restriction on μ is obtained because of the derivatives: since the Schur functions are evaluated at $a_i = 0$, and $\mathfrak{s}_\mu$ is of total degree $|\mu|$ in the a_i, we must have $|\mu| = r$ for something to survive $\partial_1 \cdots \partial_r$.

When $\mu \vdash r$, we have[(i)]

$$\partial_1 \cdots \partial_r \mathfrak{s}_\mu(a_1, \cdots, a_r)\big|_{a_1=\cdots=a_r=0} = \left\langle \mathfrak{s}_\mu, \mathfrak{p}_{\langle 1^r \rangle} \right\rangle = \dim \chi^\mu_{\mathcal{S}_{|\mu|}} = \frac{|\mu|}{h_\mu} \quad (47)$$

since derivation and multiplication by power sums are adjoint.

Combined with Lemma (3), Equation (16) and Equation (22), this gives Equation (10), which then quickly implies (11). □

5 Properties of the Rational Functions

In the RHS of formulas (10) and (11), we have a sum of rational multiples of ratios of polynomials in k (and N), hence rational functions of k. We will now explain some of the properties of these functions, which are easily deduced from Equation (11) and the limiting version of Lemma 1.

Proposition 4 *For a fixed $r, h \in \mathbb{N}$, there exists sequences of even polynomials X_r, Y_r and $\tilde{X}_{2h}$ such that for all $k \ge r, 2h$,*

$$\frac{(\mathcal{M})(2k,r)}{(\mathcal{M})(2k,0)} = \left(-\frac{\mathrm{i}}{2}\right)^r \frac{X_r(2k)}{Y_r(2k)}, \quad (48)$$

$$\frac{|\mathcal{V}|\,(2k,2h)}{|\mathcal{V}|\,(2k,0)} = \frac{\tilde{X}_{2h}(2k)}{Y_{2h}(2k)}, \quad (49)$$

(i) Observe that higher derivatives require knowledge of values of symmetric group characters at other group elements than the identity.

and such that X_r and Y_r are of the same degree, monic, and with integer coefficients. We also have $\deg \tilde{X}_{2h} \leq \deg Y_{2h}$ *and that* $\tilde{X}_{2h}$ *has integer coefficients.*

This Proposition leads or relies on a few easy facts.

Remarks.

- All polynomials are even due to Equation (16).
- We emphasize that we are not quite taking the simplest possible form of $X_r(2k)$, $\tilde{X}_r(2k)$ and $Y_r(2k)$ here. For some (presumably small) rs, some spurious cancellations *will* occur. However, an explicit expression, given in [Deh08], can be obtained for the Y_r such that all the previous statements are true. Let us just say that we are taking for Y_r the common denominator not of every term in the sum over partitions λ, but of every (1-or-2-terms-)subsum over orbits of the involution on partitions, *after* all the simplifications of the type $\frac{k+i}{2k+2i} = \frac{1}{2}$ that occur in $\frac{k\uparrow\lambda}{(2k)\uparrow\lambda}$.
- With this convention, the zeroes of Y_r are exactly at the odd integers between $1-r$ and $r-1$.
- Since the squares of dimension of characters of a finite group G sum to the order of G, or due to the existence of the Robinson-Schensted-Knuth correspondence between permutations of n and pairs of Young tableaux of the same shape partitioning n, we have starting from Equation (19) the identity

$$\sum_{\lambda\vdash n} \frac{n!}{h_\lambda^2} = 1, \tag{50}$$

  which defines the *Plancherel measure* on partitions of n. The polynomial X_r is monic thanks to this last identity.
- The sum appearing in Equation (11) is a special case of a problem studied by Jonathan Novak in Equation (9.21) of his thesis [Nov09].

6 Hypergeometric Functions of Matrix Arguments

For completeness, we now discuss an alternative way to approach the expression in (10). Hypergeometric functions of a $N \times N$ matrix argument M are generalizations of hypergeometric functions of a complex variable. Originally, this extension is defined on multisets of complex numbers, which can then be seen as eigenvalues of a matrix using the trick of Equation (20). There is extensive literature on those functions, most of it tied to multivariate statistical analysis. For a recent, accessible presentation, consider [FW08].

For M a $N \times N$ matrix, we define [DGR96, GR91] a function of M as follows:

$${}_pF_q(a_i, b_j; M) \quad = \quad \sum_\lambda \frac{\prod_{i=1}^p a_i \uparrow \lambda}{\prod_{j=1}^q b_j \uparrow \lambda} \cdot \frac{\mathfrak{s}_\lambda(M)}{h_\lambda} \tag{51}$$

where $\mathfrak{s}_\lambda(M)$ is the Schur polynomial evaluated at the eigenvalues of M. This generalizes the classical hypergeometric functions of a complex variable z to (the multiset of eigenvalues of) a square matrix variable M, via the following substitutions:

- the sum over integers is replaced by a sum over partitions;
- the generalized Pochhammer symbol replaces the rising factorial;
- the extra factorial that is always introduced for classical hypergeometric functions (by convention then) is replaced by a hook number;
- powers of z are replaced by Schur functions of the eigenvalues of M.
- They admit integral representations, closely related to Selberg integrals (see [Kan93]).

By forming an exponential generating series of Equation (10) and with the help of Equation (22), we are able to obtain the (confluent) hypergeometric function

$$\sum_{r\geq 0}\frac{(\mathcal{M})_N(2k,r)}{(\mathcal{M})_N(2k,0)}\frac{(\mathrm{i}z)^r}{r!} \quad = \quad \sum_{\mu}\frac{1}{h_\mu^2}\frac{(N\uparrow\mu)((-k)\uparrow\mu)}{(-2k)\uparrow\mu}z^{|\lambda|} = {}_1F_1(-k;-2k;z\,\mathrm{Id}_{N\times N}). \quad (52)$$

We can then substitute for the RHS of this last equation many different expressions: the theory of hypergeometric functions of matrix arguments also involves integral expressions, differential equations and recurrence relations. However, since we are interested in asymptotics of these expressions for large N, for which little theory is developed, this is unfortunately of no real use at the moment. Note though that the hypergeometric function that appears is special, as it is only evaluated at scalar matrices. In that special case, $N\times N$ determinantal formulas have been developed [GR85] (but not in the confluent case).

7 Further Work

The occurrence of the Plancherel measure is not a coincidence. In fact, much can be derived from this, and this will be the basis of further work: we will prove in a subsequent paper that the leading term of $\tilde{X}_{2h}$ has coefficient $\frac{2h!}{h!2^{3h}}$ and is of degree $2h$ lower than X_{2h}.

Acknowledgements

The author wishes to acknowledge helpful discussions with Daniel Bump, Richard Hall and Chris Hughes, as well as Alexei Borodin, whose advice was crucial at some stages.

References

[BG06] D. Bump and A. Gamburd. On the averages of characteristic polynomials from classical groups. *Comm. Math. Phys.*, 265(1):227–274, 2006.

[Bum04] D. Bump. *Lie groups*, volume 225 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2004.

[CRS06] J. B. Conrey, M. O. Rubinstein, and N. C. Snaith. Moments of the derivative of characteristic polynomials with an application to the Riemann zeta function. *Comm. Math. Phys.*, 267(3):611–629, 2006.

[Deh08] P.-O. Dehaye. Joint moments of derivatives of characteristic polynomials. *Algebra & Number Theory*, 2(1):31–68, 2008.

[DGR96] H. Ding, K. I. Gross, and D. St. P. Richards. Ramanujan's master theorem for symmetric cones. *Pacific J. Math.*, 175(2):447–490, 1996.

[FRT54] J. S. Frame, G. de B. Robinson, and R. M. Thrall. The hook graphs of the symmetric groups. *Canadian J. Math.*, 6:316–324, 1954.

[FW06] P. J. Forrester and N. S. Witte. Boundary conditions associated with the Painlevé III′ and V evaluations of some random matrix averages. *J. Phys. A*, 39(28):8983–8995, 2006.

[FW08] P. J. Forrester and S. O. Warnaar. The importance of the Selberg integral. *Bull. Amer. Math. Soc. (N.S.)*, 45(4):489–534, 2008.

[GR85] R. D. Gupta and D. St. P. Richards. Hypergeometric functions of scalar matrix argument are expressible in terms of classical hypergeometric functions. *SIAM J. Math. Anal.*, 16(4):852–858, 1985.

[GR91] K. I. Gross and D. St. P. Richards. Hypergeometric functions on complex matrix space. *Bull. Amer. Math. Soc. (N.S.)*, 24(2):349–355, 1991.

[Hal02a] R.R. Hall. Generalized Wirtinger inequalities, random matrix theory, and the zeros of the Riemann zeta-function. *J. Number Theory*, 97(2):397–409, 2002.

[Hal02b] R.R. Hall. A Wirtinger type inequality and the spacing of the zeros of the Riemann zeta-function. *J. Number Theory*, 93(2):235–245, 2002.

[Hal04] R.R. Hall. Large spaces between the zeros of the Riemann zeta-function and random matrix theory. *J. Number Theory*, 109(2):240–265, 2004.

[Hal08] R. R. Hall. Large spaces between the zeros of the Riemann zeta-function and random matrix theory. II. *J. Number Theory*, 128(10):2836–2851, 2008.

[HKO00] C. P. Hughes, J. P. Keating, and N. O'Connell. Random matrix theory and the derivative of the Riemann zeta function. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 456(2003):2611–2627, 2000.

[Hug01] C. P. Hughes. *On the Characteristic Polynomial of a Random Unitary Matrix and the Riemann Zeta Function*. PhD thesis, University of Bristol, 2001.

[Hug05] C. P. Hughes. Joint moments of the Riemann zeta function and its derivative. Personal communication, 2005.

[Kan93] J. Kaneko. Selberg integrals and hypergeometric functions associated with Jack polynomials. *SIAM J. Math. Anal.*, 24(4):1086–1110, 1993.

[KS00] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2 + it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.

[Mez03] F. Mezzadri. Random matrix theory and the zeros of $\zeta'(s)$. *J. Phys. A*, 36(12):2945–2962, 2003. Random matrix theory.

[Nov09] J. Novak. *Topics in combinatorics and random matrix theory*. PhD thesis, Queen's University, Ontario, September 2009.

[OO97] A. Okunkov and G. Olshanskiĭ. Shifted Schur functions. *Algebra i Analiz*, 9(2):73–146, 1997.

[Sag01] B. E. Sagan. *The symmetric group*, volume 203 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2001. Representations, combinatorial algorithms, and symmetric functions.

[Sak09] S.H. Saker. Large spaces between the zeros of the Riemann zeta-function. arXiv:0906.5458v1, June 2009.

[Sta99] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Ste05] J. Steuding. The Riemann zeta-function and predictions from Random Matrix Theory. February 2005.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 565–572

f-vectors of subdivided simplicial complexes (extended abstract)

Emanuele Delucchi[1] and Aaron Pixton[2] and Lucas Sabalka[1]

[1] *Department of Mathematical Sciences, Binghamton University, Binghamton NY 13902-6000,*
[2] *Department of Mathematics, Princeton University, Princeton NJ 08544-1000*

Abstract. We take a geometric point of view on the recent result by Brenti and Welker, who showed that the roots of the f-polynomials of successive barycentric subdivisions of a finite simplicial complex X converge to fixed values dependinig only on the dimension of X.
We show that these numbers are roots of a certain polynomial whose coefficients can be computed explicitely. We observe and prove an interesting symmetry of these roots about the real number -2. This symmetry can be seen via a nice realization of barycentric subdivision as a simple map on formal power series. We then examine how such a symmetry extends to more general types of subdivisions. The generalization is formulated in terms of an operator on the (formal) ring on the set of simplices of the complex.

Résumé. On applie un point de vue géométrique à un récent résultat de Brenti et Welker, qui ont montré que les racines des polynômes f de subdivisions barycentriques successives d'un complexe simplicial X convergent vers des valeurs fixes, ne dépendant que de la dimension de X.
On preuve que ces nombres sont en effet eux-mêmes racines d'un polynôme dont les coefficients peuvent être calculés explicitement. De plus, on observe et on démontre une symétrie particulière de ces nombres autour du numéro -2. Cette symétrie se révèle en exprimant l'opération de subdivision barycentrique par une fonction sur des séries de puissances formelles. Une symétrie pareille existe pour des méthodes de subdivision plus générales, où elle s'exprime par des operateurs sur l'anneau des sommes formelles de simplexes du complexe.

Keywords: subdivisions of simplicial complexes, f-vectors, f-polynomials

1 Motivation and setup

This is an extended abstract of our paper Delucchi et al. (2009), to which we refer for a full exposition and the proofs of the statements. Let us begin here by stating the theorem which motivated our work.

Let X be an arbitrary finite simplicial complex of dimension $d-1$, and for convenience assume that all vectors and matrices are indexed by rows and columns starting at 0. We are interested in roots of the f-*polynomial* of X, defined as follows. Let f_i^X denote the number of i-dimensional faces of X. We declare that $f_{-1}^X = 1$, where the (-1)-dimensional face corresponds to the empty face, $\emptyset$. The *face vector*, or f-*vector* of X is the vector

$$f^X := (f_{-1}^X, f_0^X, \dots, f_{d-1}^X).$$

Let $\underline{t}$ denote the column vector of powers of t, $(t^d, t^{d-1}, \dots t^0)^T$. The *$f$-polynomial* $f^X(t)$ encodes the f-vector as a polynomial:

$$f^X(t) := \sum_{j=0}^{d} f_{j-1}^X t^{d-j} = f^X \underline{t}$$

We now focus on the recent result of Brenti and Welker Brenti and Welker (2008) that motivated our investigations. Let X' denote the barycentric subdivision of X, and more generally let $X^{(n)}$ denote the n^{th} barycentric subdivision of X.

Theorem 1.1 *Brenti and Welker (2008) Let X be a d-dimensional simplicial complex. Then, as n grows, the roots of $f^{X^{(n)}}$ converge to $d-1$ negative real numbers which depend only on d, not on X.*

This theorem may be surprising at first: there is no dependence on the initial complex X, only on the dimension d. However, geometrically this makes perfect sense. Barycentrically subdividing a simplicial complex X over and over again causes the resulting complex $X^{(n)}$ to have far more cells than the original X. Because higher-dimensional cells contribute more new cells (in every dimension) upon subdividing than lower-dimensional ones, the top-dimensional cells begin to dominate in their 'number of contributions' to subdivisions.

More precisely, each of the f_{d-1}^X top-dimensional cells of X contribute the same amount of cells to $X^{(n)}$. Since these cells eventually dominate contributions from smaller-dimensional cells, the f-polynomial for $X^{(n)}$ can be approximated by f_{d-1}^X times the f-polynomial associated to the n-fold subdivision of a single top-dimensional cell, $\sigma_d^{(n)}$. Since the roots of a polynomial are unaffected by multiplication by constants, the roots of $f^{X^{(n)}}$ converge to the roots of $f^{\sigma_d^{(n)}}$ as n increases.

We will see that both these sequences converge to the roots of a specific polynomial, and these roots satisfy an interesting symmetry.

We begin by observing the effect on f-vectors of barycentric subdivision. One key observation is that barycentric subdivision multiplies f-vectors by a fixed matrix, F_d:

Definition 1.2 *Define $\mathring{f}_i^X$ to be the number of interior i-faces of X for $i \geq 0$. We set $\mathring{f}_{-1}^X = 1$ if the dimension of X is -1, and 0 otherwise. Let σ_d denote the standard $(d-1)$-dimensional simplex. Define F_d to be the $(d+1)\times(d+1)$ matrix determined by the interior $(j-1)$-faces of the subdivided i-simplex:*

$$F_d := [\mathring{f}_{j-1}^{\sigma_i'}].$$

With this notation in place, we have the following.

Corollary 1.3 *For any $n \geq 0$,*

$$f^{X^{(n)}} = f^X F_d^n.$$

Thus, to understand barycentric subdivision, we need to understand the matrix F_d. We will compute the entries in F_d more explicitly later, but for now we simply observe a formula which follows from Inclusion-Exclusion:

Lemma 1.4 *If $j > i$ then $\mathring{f}_j^{\sigma'_i} = 0$. If $j \leq i$, then*

$$\mathring{f}_j^{\sigma'_i} = \sum_{k=0}^{i} (-1^k) \binom{i}{k} f_j^{\sigma'_{i-k}}.$$

By this lemma, F_d is lower triangular with diagonal entries $\mathring{f}_i^{\sigma'_i} = f_i^{\sigma'_i} = i!$. Thus, the eigenvalues of F_d are $0!, 1!, 2!, 3!, \ldots, d!$.

2 Main results

2.1 The limit polynomial

The goal of this section is to prove that the limit values of the roots of the $f^{\sigma_d^n}$ are themselves roots of a polynomial of which we can explititely compute the roots. The geometric intuition behing this fact is obtaining by noticing that, by definition, the coefficients of $f^{\sigma_d^{(n)}}$ record the number of cells of each dimension occurring in $\sigma_d^{(n)}$. Moreover, the number of cells in each dimension is bounded by a constant times the number of top-dimensional cells. Thus, if we normalize $f^{\sigma_d^{(n)}}$ by dividing by the number of top-dimensional cells, we have coefficients which, for each k, record the density of k-cells relative to the number of top-dimensional cells. As this density is positive but strictly decreases upon subdividing, there is a limiting value for the coefficient. Thus, there should be a limiting polynomial, with well-defined roots. Let us make this precise.

By Corollary 1.3, $f^{X^{(n)}}(t) = f^X F_d^n \underline{t}$. As the greatest eigenvalue of F_d is $d!$, we normalize $f^{X^{(n)}}(t)$ by dividing by $(d!)^n$ - let $p_n^X(t)$ denote the result:

$$p_n^X(t) := \frac{1}{(d!)^n} f^{X^{(n)}}(t).$$

Note this normalization does not alter the roots. It will also often be convenient to reverse the order of the coefficients of $p_n^X(t)$, with the effect of inverting the roots of $p_n^X(t)$ (that is, the roots of $f^{X^{(n)}}(t)$) about the unit circle in the extended complex plane:

$$q_n^X(t) := t^d p_n^X(t^{-1}).$$

To take powers of F_d, we diagonalize,

$$F_d = P_d D_d P_d^{-1},$$

where D_d is the diagonal matrix of eigenvalues $0!, 1!, \ldots, d!$ and P_d is the (lower triangular) diagonalizing matrix of eigenvectors. Thus, $F_d^n = P_d D_d^n P_d^{-1}$.

Now, let $\widetilde{D}_d := \frac{1}{d!} D_d$. Let $\bar{t}$ denote the column vector $\underline{t}$ in reverse order, $\bar{t} = (t^0, t^1, \ldots t^d)^T$. For any simplicial complex X, we thus have the following equations:

$$f^{X^{(n)}}(t) = f^X P_d D_d^n P_d^{-1} \underline{t} = (d!)^n \left(f^X P_d\right) \left(\widetilde{D}_d\right)^n \left(P_d^{-1}\right) \underline{t}$$

$$p_n^X(t) = \left(f^X P_d\right) \left(\widetilde{D}_d\right)^n \left(P_d^{-1}\right) \underline{t}, \quad q_n^X(t) = \left(f^X P_d\right) \left(\widetilde{D}_d\right)^n \left(P_d^{-1}\right) \bar{t}$$

As the eigenvalues of F_d are $0!, 1!, \ldots, d!$, for large n, D_d^n is dominated by its d^{th} diagonal entry, $(d!)^n$. In the limit, the powers of the matrix $\widetilde{D}_d = \frac{1}{d!} D_d$ converge to the matrix

$$M_{d,d} := \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}.$$

Thus, as n grows, the polynomials p_n^X and q_n^X respectively approach the polynomials

$$p_\infty^X(t) := \left(f^X P_d\right) M_{d,d} \left(P_d^{-1}\right) \underline{t}\,, \quad q_\infty^X(t) := \left(f^X P_d\right) M_{d,d} \left(P_d^{-1}\right) \overline{t}$$

in the sense that each sequence converges coefficient-wise in the vector space of polynomials of degree at most d.

By Corollary 1.3 and Lemma 1.4, we know the leading and trailing coefficients of $p_n^X(t)$ and $q_n^X(t)$: $p_n^X(t) = (d!)^{-n} t^d + \ldots + f_{d-1}^X$ and $q_n^X(t) = (d!)^{-n} + \ldots + f_{d-1}^X t^d$. Hence, in the limit, $p_\infty^X(t)$ does not have 0 as a root, but has degree less than d (one root of the p_n^X diverges to $-\infty$), while $q_\infty^X(t)$ is of degree d with 0 as a root. Because the polynomials $q_n^X(t)$ converge coefficient-wise to the polynomial $q_\infty^X(t)$ of the same degree, their roots also converge (see for instance Tyrtyshnikov (1997)):

Because the matrix P_d is lower triangular and $M_{d,d}$ has only one nonzero entry in position (d, d), we have

$$\left(f^X P_d\right) M_{d,d} = c_{X,d} e_d^T,$$

where e_d is the unit vector with a 1 in the d^{th} row, and $c_{X,d}$ is a constant depending on f^X and P_d. As both f^X and P_d do not depend on the amount of subdivision n, the roots of p_∞^X and q_∞^X do not depend on the value of $c_{X,d}$, and thus do not depend on *any* coefficient of f_d^X. This leads us to the following definition:

Definition 2.1 *Define the* limit p-polynomial *and the* limit q-polynomial *by*

$$p_d(t) := e_d^T P_d \underline{t}\,, \qquad q_d(t) := e_d^T P_d \overline{t}.$$

To summarize:

1. The roots of $f^{X^{(n)}}(t)$ are equal to the roots of $p_n^X(t)$.
2. The roots of $q_n^X(t)$ (resp. $p_n^X(t)$ converge to the roots of $q_d(t)$ (resp. $p_d^X(t)$), and depend only on the dimension of X.
3. The coefficient of t^i in the polynomial $p_d(t)$ (resp. $q_d(t)$) is the $(d-i)^{th}$ (resp. the i^{th}) entry in last row of P_d^{-1}.

In the full paper Delucchi et al. (2009) we derive explicit formulas for the computation of the coefficients of the matrix P_d^{-1}. We reproduce the result of some of these computations in our last section here.

Using the fact, proved by Brenti and Welker, that the limit of the roots of the f-polynomial are distinct and all real, we can summarize as follows.

Theorem A *Let X be a d-dimensional simplicial complex. Then, as n increases, the roots of $f^{X^{(n)}}$ converge to the $d-1$ (distinct) roots of a polynomial $p_d(t)$, whose coefficients can be explicitely computed and depend only on d, not on X.*

2.2 Symmetry of the limit values

Our result about the symetry of the 'limit roots' is the following.

Theorem B *For any dimension d, the $d-1$ 'limit' roots are invariant under the map $x \mapsto \frac{-x}{x+1}$.*

We will prove the corresponding symmetry for the roots of q_d instead of p_d, as it becomes a mirror symmetry instead of a Möbius invariance.

Theorem 2.2 *For every dimension d,*

$$q_\infty(t) = (-1)^d q_\infty(-1-t).$$

In particular, the roots of $q_\infty(t)$ are (linearly) symmetric with respect to $-\frac{1}{2}$.

As a first step, note that Lemma 1.4 gives the following expressions.

Lemma 2.3 *Let X be a simplicial complex. The f-polynomial of its barycentric subdivision $f^{X'}(t)$ and the corresponding $q_1^X(t)$ are given by*

$$f^{X'}(t) = \sum_{j=0}^{d} \Delta^j \{f^X(l)\}_l t^{d-j}, \qquad (d!) q_1^X(t) = \sum_{k=0}^{d} \Delta^k \{q_0^X(l)\}_l t^k.$$

This prompts us to consider barycentric subdivision as a function on polynomials in t defined by

$$b : \mathbb{Z}[t] \to \mathbb{Z}[t], \quad g(t) \mapsto \sum_{k \geq 0} \Delta^k \{g(l)\}_l t^k,$$

so that, for a simplicial complex X of dimension d we have then $b(q_j^X(t)) = d! q_{j+1}^X(t)$. The function b is linear, and thus it is given by its values on monomials, which we arrange in a formal power series in the variable x over the ring $\mathbb{Z}[t]$. We thus consider a function B on the ring $\mathbb{Z}[t][[x]]$ defined as

$$B : \quad \sum_{k \geq 0} g_k(t) x^k \quad \longmapsto \quad \sum_{k \geq 0} b\big(g_k(t)\big) x^k.$$

Theorem C *In $\mathbb{Z}[t][[x]]$, barycentric subdivision satisfies the identity*

$$B(e^{tx}) = \frac{1}{1-(e^x-1)t}.$$

To investigate the stated symmetry, we consider the following map

$$\iota : \mathbb{Z}[t] \to \mathbb{Z}[t], \quad g(t) \mapsto g(-1-t).$$

One readily checks by explicit calculation that $\iota B(e^{tx}) = B\iota(e^{tx})$. This suffices to prove the follwing key fact.

Lemma 2.4 *The map ι is an involution, and it satisfies*

$$\iota b \iota = b.$$

Recall that barycentric subdivision has the effect on each p- and q-polynomial of multiplying on the right by F before the $\underline{t}$ and $\overline{t}$, respectively, and rescaling by dividing by $d!$. In the limit, the limit p- and q-polynomials are invariant under barycentric subdivision up to this scaling: thus $b\big(q_\infty(t)\big) = d! q_\infty(t)$. Moreover, since the eigenvalues of F are all distinct, q_∞ is characterized by this identity and by having leading coefficient f^X_{d-1}.

A computation based on Lemma 2.4 shows $b\big(q_\infty(-1-t)\big) = d!\big(q_\infty(-1-t)\big)$ and since the lead coefficient of $q_\infty(-1-t)$ is $(-1)^d f^X_{d-1}$, the stated symmetry holds.

3 Symmetry for Other Subdivision Methods

In general, given any polynomial $g(t) \in \mathbb{Z}[t]$, we can consider the polynomial $\iota g(t) = g(-1-t)$. The coefficient of t^k in $g(t)$ contributes $(-1)^k\binom{k}{j}$ times itself to the coefficient of t^j in $\iota g(t)$: this contribution is exactly the number of $(j-1)$-dimensional faces of the $(k-1)$-dimensional simplex. Thus, we can interpret ι as a map on formal sums of simplices, as follows.

Let S be the set of simplices of a given simplicial complex X with vertex set VX. We will think of every simplex $\sigma \in S$ as a subset of VX. Now we can write

$$\iota : \mathbb{Z}[S] \to \mathbb{Z}[S], \quad \sigma \mapsto (-1)^{\dim\sigma+1} \sum_{\tau\subseteq\sigma} \tau.$$

We will identify a subdivision of X by the triple $(X, \widetilde{X}, \phi)$, where $\widetilde{X}$ is the simplicial complex subdividing X (the 'result' of the subdivision) and $\phi : \widetilde{S} \to S$ is the function associating to each simplex $\widetilde{\sigma} \in \widetilde{S}$ its support in X. Now, a subdivision $(X, \widetilde{X}, \phi)$ induces a linear map

$$b_\phi : \mathbb{Z}[S] \to \mathbb{Z}[\widetilde{S}], \quad \sigma \mapsto \sum_{\phi(\widetilde{\sigma})=\sigma} \widetilde{\sigma}.$$

A *subdivision method* Φ is a collection of subdivisions $\Phi := \{(\sigma_n, \widetilde{\sigma}_n, \phi_n)\}_{n\geq 0}$ such that for every k-face $i_k : \sigma_k \to \sigma_m$ of the standard m-simplex, the map ϕ_k is the restriction of ϕ_m to $i_k\sigma_k$. This ensures that, given any simplicial complex X, the complex $\Phi(X)$, called *subdivision of X according to the rule* Φ is uniquely defined by requiring that every n-simplex of X is subdivided as $(\sigma_n, \widetilde{\sigma}_n, \phi_n) \in \Phi$. A subdivision method is *nontrivial in dimension* n if ϕ_k is not the identity map for some $k \leq n$. Clearly if a subdivision is nontrivial in dimension n, then ϕ_n is not the identity map.

Given a subdivision method Φ, in view of the linearity of b_ϕ for each subdivision, it makes sense to write

$$b_\Phi(\sum_{\sigma\in X} \sigma) = \sum_{\sigma\in X} b_\Phi \sigma.$$

As with the map b given by barycentric subdivision, for any subdivision method the induced map b_Φ always commutes with the map ι:

Lemma 3.1 *For any subdivision method Φ, $\iota b_\Phi = b_\Phi \iota$.*

This commutativity was the key step in proving the symmetry for the barycentric subdivision, together with the fact that F_d had a dominating eigenvalue with geometric multiplicity 1. The latter property holds for the matrix realizing any subdivision method that is nontrivial in the top dimension. We thus have the desired result.

Theorem D *For any dimension n and any subdivision method Φ which is nontrivial in dimension n, there exists a unique 'limit polynomial' $p_{n,\Phi}(t)$, such that, for any d-dimensional simplicial complex X, the roots of $f^{\Phi^k(X)}(t)$ converge to the roots of $p_{n,\Phi}(t)$ as k increases. The roots of $p_{n,\Phi}(t)$ are invariant under the Möbius transformation $x \mapsto \frac{-x}{x+1}$.*

Remark 3.2 *Since the above interpretation is on the level of formal sums of simplices, the most natural context in which to study it seems to be the Stanley-Reisner ring $\mathbb{K}[X]$, defined by any simplicial complex X and any field $\mathbb{K}$. A good introduction to these rings can be found in Stanley (1996), where some properties of the Stanley-Reisner ring of a subdivision of a simplicial complex are explored. This brings us to ask the following question.*

Question 3.3 *Is there a (multi-)complex in each dimension whose f-polynomial is related to the limit polynomials $p^X_\infty(t)$ or $q^X_\infty(t)$? More generally, is there a geometric interpretation of the coefficients or the roots of $p^X_\infty(t)$ (equivalently, $q^X_\infty(t)$)?*

Brenti and Welker raise the question of defining a general concept of "barycentric subdivision" for a standard graded algebra. We can broaden the question to involve all subdivision methods, and ask whether the formulas developed in (Delucchi et al., 2009, Section 5) can be taken as a starting point to answer this question.

3.1 Computations

Our method allows us to explicitly compute the coefficients of $p_d(t)$, of $q_d(t)$, and thus also the limit roots. We carry out these computations in our full paper. As a sample, we give the values of the roots of $q_d(t)$ for $d \leq 10$ (computations which take less than 1 second of processor time using the formulae we derive in Delucchi et al. (2009)). The roots of $q_d(t)$ are, for $d \leq 10$, approximated by:

$d=2:$	-1	0								
$d=3:$	-1	$-.5$	0							
$d=4:$	-1	$-.76112$	$-.23888$	0						
$d=5:$	-1	$-.88044$	$-.5$	$-.11956$	0					
$d=6:$	-1	$-.93787$	$-.68002$	$-.31998$	$-.06213$	0				
$d=7:$	-1	$-.9668$	$-.79492$	$-.5$	$-.20508$	$-.0332$	0			
$d=8:$	-1	$-.98189$	$-.86737$	$-.63852$	$-.36148$	$-.13263$	$-.01811$	0		
$d=9:$	-1	$-.98996$	$-.91332$	$-.73961$	$-.5$	$-.26039$	$-.08668$	$-.01004$	0	
$d=10:$	-1	$-.99437$	$-.94277$	$-.81205$	$-.61285$	$-.38715$	$-.18795$	$-.05723$	$-.00563$	0

Acknowledgements

The authors would like to thank Laura Anderson for bringing Theorem 1.1 to their attention, Tom Zaslavsky for interesting discussions, and Dennis Pixton for introducing the authors.

References

F. Brenti and V. Welker. f-vectors of barycentric subdivisions. *Math. Z.*, 259(4):849–865, 2008. ISSN 0025-5874. doi: 10.1007/s00209-007-0251-z. URL http://dx.doi.org/10.1007/s00209-007-0251-z.

E. Delucchi, A. Pixton, and L. Sabalka. f-vectors of subdivided simplicial complexes. arXiv:1002.3201, 13 pp., 2009.

R. P. Stanley. *Combinatorics and commutative algebra*, volume 41 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, second edition, 1996. ISBN 0-8176-3836-9.

E. E. Tyrtyshnikov. *A brief introduction to numerical analysis*. Birkhäuser Boston Inc., Boston, MA, 1997. ISBN 0-8176-3916-0.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 573–583

A Combinatorial Formula for Orthogonal Idempotents in the 0-Hecke Algebra of S_N

Tom Denton[1†]

[1]*Department of Mathematics, University of California Davis, One Shields Ave, Davis, California 95616*

Abstract. Building on the work of P.N. Norton, we give combinatorial formulae for two maximal decompositions of the identity into orthogonal idempotents in the 0-Hecke algebra of the symmetric group, $\mathbb{C}H_0(S_N)$. This construction is compatible with the branching from $H_0(S_{N-1})$ to $H_0(S_N)$.

Résumé. En s'appuyant sur le travail de P.N. Norton, nous donnons des formules combinatoires pour deux décompositions maximales de l'identité en idempotents orthogonaux dans l'algèbre de Hecke $H_0(S_N)$ du groupe symétrique à $q = 0$. Ces constructions sont compatibles avec le branchement de $H_0(S_{N-1})$ à $H_0(S_N)$.

Keywords: Iwahori-Hecke algebra, idempotents, semigroups, combinatorics, representation theory

1 Introduction

The 0-Hecke algebra $\mathbb{C}H_0(S_N)$ for the symmetric group S_N can be obtained as the Iwahori-Hecke algebra of the symmetric group $H_q(S_N)$ at $q = 0$. It can also be constructed as the algebra of the monoid generated by anti-sorting operators on permutations of N.

P.N. Norton described the full representation theory of $\mathbb{C}H_0(S_N)$ in Norton (1979): In brief, there is a collection of 2^{N-1} simple representations indexed by subsets of the usual generating set for the symmetric group, and an additional collection of 2^{N-1} projective indecomposable modules. Norton gave a construction for some elements generating these projective modules, but these elements were neither orthogonal nor idempotent. While it was known that an orthogonal collection of idempotents to generate the indecomposable modules exists, there was no known formula for these elements.

Herein, we describe an explicit construction for two different families of orthogonal idempotents in $\mathbb{C}H_0(S_N)$, one for each of the two orientations of the Dynkin diagram for S_N. The construction proceeds by creating a collection of 2^{N-1} *demipotent* elements, which we call *diagram demipotents*, each indexed by a copy of the Dynkin diagram with signs attached to each node. These elements are demipotent in the sense that for each element X, there exists some number $k \leq N-1$ such that X^j is idempotent for all $j \geq k$. The collection of idempotents thus obtained provides a maximal orthogonal decomposition of the identity.

An important feature of the 0-Hecke algebra is that it is the monoid algebra of a $\mathcal{J}$-trivial monoid. As a result, its representation theory is highly combinatorial. This paper is part of an ongoing effort with

[†]Supported in part by the VIGRE grant NSF–DMS0636297 and NSF grants DMS–0652641 and DMS–0652652.

Florent Hivert, Anne Schilling, and Nicolas Thiéry to characterize the representation theory of general J-trivial monoids, continuing the work of Norton (1979), Hivert and Thiéry (2009), Hivert et al. (2009). The fundamentals of the representation theory of semigroups can be found in Ganyushkin et al. (2009). All proofs of statements in this paper will appear in Denton et al. (2010).

The diagram demipotents obey a branching rule which compares well to the situation in Okounkov and Vershik (1996) in their 'New Approach to the Representation Theory of the Symmetric Group.' In their construction, the branching rule for S_N is given primary importance, and yields a canonical basis for the irreducible modules for S_N which pull back to bases for irreducible modules for S_{N-M}.

Okounkov and Vershik further make extensive use of a maximal commutative algebra generated by the Jucys-Murphy elements. In the 0-Hecke algebra, their construction does not directly apply, because the deformation of Jucys-Murphy elements (which span a maximal commutative subalgebra of $\mathbb{C}S_N$) to the 0-Hecke algebra no longer commute. Instead, the idempotents obtained from the diagram demipotents play the role of the Jucys-Murphy elements, generating a commutative subalgebra of $\mathbb{C}H_0(S_N)$ and giving a natural decomposition into indecomposable modules, while the branching diagram describes the multiplicities of the irreducible modules.

The Okounkov-Vershik construction is well-known to extend to group algebras of general finite Coxeter groups (Ram (1997)). It remains to be seen whether our construction for orthogonal idempotents generalizes beyond type A. However, the existence of a process for type A gives hope that the Okounkov-Vershik process might extend to more general 0-Hecke algebras of Coxeter groups.

Section 2 establishes notation and describes the relevant background necessary for the rest of the paper. For further background information on the properties of the symmetric group, one can refer to the books of Humphreys (1990) and Stanley (1997). Section 3 gives the construction of the diagram demipotents. Section 4 describes the branching rule the diagram demipotents obey, and also establishes the Sibling Rivalry Lemma, which is useful in proving the main results, in Theorem 4.8. Section 5 establishes bounds on the power to which the diagram demipotents must be raised to obtain an idempotent. Finally, remaining questions are discussed in Section 6.

2 Background and Notation

Let S_N be the symmetric group defined by the generators s_i for $i \in I = \{1, \dots, N-1\}$ with the usual relations:

- Reflection: $s_i^2 = 1$,
- Commutation: $s_i s_j = s_j s_i$ for $|i - j| > 1$,
- Braid relation: $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$.

The relations between distinct generators are encoded in the *Dynkin diagram* for S_N, which is a graph with one node for each generator s_i, and an edge between the pairs of nodes corresponding to generators s_i and s_{i+1} for each i. Here, an edge encodes the braid relation, and generators whose nodes are not connected by an edge commute. (See figure 3.)

Definition 2.1 *The* 0**-Hecke monoid** $H_0(S_N)$ *is generated by the collection* π_i *for* i *in the set* $I = \{1, \dots, N-1\}$ *with relations:*

- *Idempotence:* $\pi_i^2 = \pi_i$,

- *Commutation:* $\pi_i\pi_j = \pi_j\pi_i$ *for* $|i-j| > 1$*,*
- *Braid Relation:* $\pi_i\pi_{i+1}\pi_i = \pi_{i+1}\pi_i\pi_{i+1}$.

The 0-Hecke monoid can be realized combinatorially as the collection of anti-sorting operators on permutations of N. For any permutation σ, $\pi_i\sigma = \sigma$ if $i+1$ comes before i in the one-line notation for σ, and $\pi_i\sigma = s_i\sigma$ otherwise.

Additionally, $\sigma\pi_i = \sigma s_i$ if the ith entry of σ is less than the $i+1$th entry, and $\sigma\pi_i = \sigma$ otherwise. (The left action of π_i is on *values*, and the right action is on *positions*.)

Definition 2.2 *The* **0-Hecke algebra** $\mathbb{C}H_0(S_N)$ *is the monoid algebra of the* 0*-Hecke monoid.*

Words for S_N and $H_0(S_N)$ Elements. The set $I = \{1, \dots, N-1\}$ is called the *index set* for the Dynkin diagram. A *word* is a sequence $(i_1, \dots, i_k)$ of elements of the index set. To any word w we can associate a permutation $s_w = s_{i_1} \dots s_{i_k}$ and an element of the 0-Hecke monoid $\pi_w = \pi_{i_1} \cdots \pi_{i_k}$. A word w is *reduced* if its length is minimal amongst words with permutation s_w. The *length* of a permutation σ is equal to the length of a reduced word for σ.

Elements of the 0-Hecke monoid are indexed by permutations: Any reduced word $s = s_{i_1} \dots s_{i_k}$ for a permutation σ gives a reduced word in the 0-Hecke monoid, $\pi_{i_1} \cdots \pi_{i_k}$. Furthermore, given two reduced words w and v for a permutation σ, then w is related to v by a sequence of braid and commutation relations. These relations still hold in the 0-Hecke monoid, so $\pi_w = \pi_v$.

From this, we can see that the 0-Hecke monoid has $N!$ elements, and that the 0-Hecke algebra has dimension $N!$ as a vector space. Additionally, the length of a permutation is the same as the length of the associated $H_0(S_N)$ element.

We can obtain a *parabolic sub-object* (group, monoid, algebra) by considering the object whose generators are indexed by a subset $J \subset I$, retaining the relations of the original object. The Dynkin diagram of the corresponding object is obtained by deleting the relevant nodes and connecting edges from the original Dynkin diagram. Every parabolic subgroup of S_N contains a unique longest element, being an element whose length is maximal amongst all elements of the subgroup. We will denote the longest element in the parabolic sub-monoid of $H_0(S_N)$ with generators indexed by $J \subset I$ by w_J^+, and use $\hat{J}$ to denote the complement of J in I. For example, in $H_0(S_8)$ with $J = \{1, 2, 6\}$, then $w_J^+ = \pi_{1216}$, and $w_{\hat{J}}^+ = \pi_{3453437}$.

Definition 2.3 *An element* x *of a semigroup or algebra is* demipotent *if there exists some* k *such that* $x^\omega := x^k = x^{k+1}$*. A semigroup is* aperiodic *if every element is demipotent.*

The 0-Hecke monoid is aperiodic. Namely, for any element $x \in H_0(S_N)$, let:

$$J(x) = \{i \in I \mid \text{ s.t. } i \text{ appears in some reduced word for } x\}.$$

This set is well defined because if i appears in some reduced word for x, then it appears in every reduced word for x. Then $x^\omega = w_{J(x)}^+$.

The Algebra Automorphism Ψ of $\mathbb{C}H_0(S_N)$. $\mathbb{C}H_0(S_N)$ is alternatively generated as an algebra by elements $\pi_i^- := (1 - \pi_i)$, which satisfy the same relations as the π_i generators. There is a unique automorphism Ψ of $\mathbb{C}H_0(S_N)$ defined by sending $\pi_i \to (1 - \pi_i)$.

For any longest element w_J^+, the image $\Psi(w_J^+)$ is a longest element in the $(1 - \pi_i)$ generators; this element is denoted w_J^-.

The Dynkin diagram Automorphism of $\mathbb{C}H_0(S_N)$. A Dynkin diagram automorphism is a graph automorphism of the underlying graph. For the Dynkin diagram of S_N, there is exactly one non-trivial automorphism, sending the node i to $N-i$.

This diagram automorphism induces an automorphism of the symmetric group, sending the generator $s_i \to s_{N-i}$ and extending multiplicatively. Similarly, there is an automorphism of the 0-Hecke monoid sending the generator $\pi_i \to \pi_{N-i}$ and extending multiplicatively.

Bruhat Order. The *(left) weak order* on the set of permutations is defined by the relation $\sigma \leq \tau$ if there exist reduced words v, w such that $\sigma = s_v, \tau = s_w$, and v is a prefix of w in the sense that $w = v_1, v_2, \ldots, v_j, w_j + 1, \ldots, w_k$. The right weak order is defined analogously, where v must appear as a suffix of w.

The left weak order also exists on the set of 0-Hecke monoid elements, with exactly the same definition. Indeed, $s_v \leq s_w$ if and only if $\pi_v \leq \pi_w$.

For a permutation σ, we say that i is a *(left) descent* of σ if $s_i\sigma \leq \sigma$. We can define a descent in the same way for any element π_w of the 0-Hecke monoid. We write $D_L(\sigma)$ and $D_L(\pi_w)$ for the set of all descents of σ and m respectively. Right descents are defined analogously, and are denoted $D_R(\sigma)$ and $D_R(\pi_w)$, respectively.

It is well known that i is a left descent of σ if and only if there exists a reduced word w for σ with $w_1 = i$. As a consequence, if $D_L(\pi_w) = J$, then $w_J^+\pi_w = \pi_w$. Likewise, i is a right descent if and only if there exists a reduced word for σ ending in i, and if $D_R(\pi_w) = J$, then $\pi_w w_J^+ = \pi_w$.

Bruhat order is defined by the relation $\sigma \leq \tau$ if there exist reduced words v and w such that $s_v = \sigma$ and $s_w = \tau$ and v appears as a subword of w. For example, 13 appears as a subword of 123, so $s_{12} \leq s_{123}$ in strong Bruhat order.

Representation Theory The representation theory of $\mathbb{C}H_0(S_N)$ was described in Norton (1979) and expanded to generic finite Coxeter groups in Carter (1986). A more general approach to the representation theory can be taken by approaching the 0-Hecke algebra as a semigroup algebra, as per Ganyushkin et al. (2009). The principal results are reproduced here for ease of reference.

For any subset $J \subset I$, let λ_J denote the one-dimensional representation of H defined by the action of the generators:

$$\lambda_J(\pi_i) = \begin{cases} 0 & \text{if } i \in J, \\ -1 & \text{if } i \notin J. \end{cases}$$

The λ_J are 2^{N-1} non-isomorphic representations, all one-dimensional and thus simple. In fact, these are all of the simple representations of $\mathbb{C}H_0(S_N)$.

The nilpotent radical $\mathcal{N}$ in $\mathbb{C}H_0(S_N)$ is spanned by elements of the form $x - w^+_{J(x)}$, where x is an element of the monoid $H_0(S_N)$, and $w^+_{J(x)}$ is the longest element in the parabolic submonoid whose generators are exactly the generators in any given reduced word for x. This element $w^+_{J(x)}$ is idempotent. If y is already idempotent, then $y = w^+_{J(y)}$, and so $y - w^+_{J(y)} = 0$ contributes nothing to $\mathcal{N}$. However, all other elements $x - w^+_{J(x)}$ for x not idempotent are linearly independent, and thus give a basis of $\mathcal{N}$.

Norton further showed that

$$\mathbb{C}H_0(S_N) = \bigoplus_{J \subset I} H_0(S_N) w_J^- w_{\hat{J}}^+$$

is a direct sum decomposition of $\mathbb{C}H_0(S_N)$ into indecomposable left ideals.

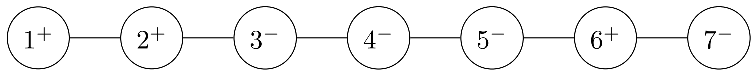


Fig. 1: A signed Dynkin diagram for S_8.

Theorem 2.4 (Norton, 1979) *Let $\{p_J | J \subset I\}$ be a set of mutually orthogonal primitive idempotents with $p_J \in \mathbb{C}H_0(S_N)w_J^- w_{\hat{J}}^+$ for all $J \subset I$ such that $\sum_{J \subset I} p_J = 1$.*

Then $\mathbb{C}H_0(S_N)w_J^- w_{\hat{J}}^+ = \mathbb{C}H_0(S_N)p_J$, and if $\mathcal{N}$ is the nilpotent radical of $\mathbb{C}H_0(S_N)$, $\mathcal{N}w_J^- w_{\hat{J}}^+ = \mathcal{N}p_J$ is the unique maximal left ideal of $\mathbb{C}H_0(S_N)p_J$, and $\mathbb{C}H_0(S_N)p_J/\mathcal{N}p_J$ affords the representation λ_J.

Finally, the commutative algebra $\mathbb{C}H_0(S_N)/\mathcal{N} = \bigoplus_{J \subset I} \mathbb{C}H_0(S_N)p_J/\mathcal{N}p_J = \mathbb{C}^{2^{N-1}}$.

The proof of this theorem is non-constructive, and does not give a formula for the idempotents.

3 Diagram Demipotents

The elements π_i and $(1 - \pi_i)$ are idempotent. There are actually 2^{N-1} idempotents in $H_0(S_N)$, namely the elements w_J^+ for any $J \subset I$. These idempotents are clearly not orthogonal, though.

The goal of this paper is to give a formula for a collection of *orthogonal* idempotents in $\mathbb{C}H_0(S_N)$.

Definition 3.1 *A* signed diagram *is a Dynkin diagram in which each vertex is labeled with a $+$ or $-$.*

Figure 3 depicts a signed diagram for type A_7, corresponding to $H_0(S_8)$. For brevity, a diagram can be written as just a string of signs. For example, the signed diagram in the Figure is written $++---+-$.

We now construct a *diagram demipotent* corresponding to each signed diagram. Let P be a composition of the index set I obtained from a signed diagram D by grouping together sets of adjacent pluses and minuses. For the diagram in Figure 3, we would have $P = \{\{1,2\},\{3,4,5\},\{6,7\}\}$. Let P_k denote the kth subset in P. For each P_k, let $w_{P_k}^{sgn(k)}$ be the longest element of the parabolic sub-monoid associated to the index set P_k, constructed with the generators π_i if $sgn(k) = +$ and constructed with the $(1 - \pi_i)$ generators if $sgn(k) = -$.

Definition 3.2 *Let D be a signed diagram with associated composition $P = P_1 \cup \cdots \cup P_m$. Set:*

$$\begin{aligned} L_D &= w_{P_1}^{sgn(1)} w_{P_2}^{sgn(2)} \cdots w_{P_m}^{sgn(m)}, \text{ and} \\ R_D &= w_{P_m}^{sgn(m)} w_{P_{m-1}}^{sgn(m-1)} \cdots w_{P_1}^{sgn(1)}. \end{aligned}$$

The diagram demipotent *C_D associated to the signed diagram D is then $L_D R_D$. The* opposite diagram demipotent *C'_D is $R_D L_D$.*

Thus, the diagram demipotent for the diagram in Figure 3 is $\pi_{121}^+ \pi_{345343}^- \pi_6^+ \pi_7^- \pi_6^+ \pi_{345343}^- \pi_{121}^+$.

It is not immediately obvious that these elements are demipotent; this is a direct result of Lemma 4.4.

For $N = 1$, there is only the empty diagram, and the diagram demipotent is just the identity.

For $N = 2$, there are two diagrams, $+$ and $-$, and the two diagram demipotents are π_1 and $1 - \pi_1$ respectively. Notice that these form a decomposition of the identity, as $\pi_i + (1 - \pi_i) = 1$.

For $N = 3$, we have the following list of diagram demipotents. The first column gives the diagram, the second gives the element written as a product, and the third expands the element as a sum. For brevity, words in the π_i or π_i^- generators are written as strings in the subscripts. Thus, $\pi_1\pi_2$ is abbreviated to π_{12}.

D	C_D	Expanded Demipotent
$++$	π_{121}	π_{121}
$+-$	$\pi_1\pi_2^-\pi_1$	$\pi_1 - \pi_{121}$
$-+$	$\pi_1^-\pi_2\pi_1^-$	$\pi_2 - \pi_{12} - \pi_{21} + \pi_{121}$
$--$	π_{121}^-	$1 - \pi_1 - \pi_2 + \pi_{12} + \pi_{21} - \pi_{121}$

Observations.

- The idempotent π_{121}^- is an alternating sum over the monoid. This is a general phenomenon: By Norton (1979), w_J^- is the length-alternating signed sum over the elements of the parabolic sub-monoid with generators indexed by J.

- The shortest element in each expanded sum is an idempotent in the monoid with π_i generators; this is also a general phenomenon. The shortest term is just the product of longest elements in nonadjacent parabolic sub-monoids, and is thus idempotent. Then the shortest term of C_D is π_J^+, where J is the set of nodes in D marked with a $+$. Each diagram yields a different leading term, so we can immediately see that the 2^{N-1} idempotents in the monoid appear as a leading term for exactly one of the diagram demipotents, and that they are linearly independent.

- For many purposes, one only needs to explicitly compute half of the list of diagram demipotents; the other half can be obtained via the automorphism Ψ. A given diagram demipotent x is orthogonal to $\Psi(x)$, since one has left and right π_1 descents, and the other has left and right π_1^- descents, and $\pi_1\pi_1^- = 0$.

- The diagram demipotents C^D and C^E for $D \neq E$ do not necessarily commute. Non-commuting demipotents first arise with $N = 6$. However, the idempotents obtained from the demipotents are orthogonal and do commute.

- It should also be noted that these demipotents (and the resulting idempotents) are not in the projective modules constructed by Norton, but generate projective modules isomorphic to Norton's.

- The diagram demipotents C_D listed here are not fixed under the automorphism induced by the Dynkin diagram automorphism. In particular, the 'opposite' diagram demipotents $C'_D = R_D L_D$ really are different elements of the algebra, and yield an equally valid but different set of orthogonal idempotents. For purposes of comparison, the diagram demipotents for the reversed Dynkin diagram are listed below for $N = 3$.

D	C'_D	Expanded Demipotent
$++$	π_{212}	π_{212}
$+-$	$\pi_2\pi_1^-\pi_2$	$\pi_2 - \pi_{212}$
$-+$	$\pi_2^-\pi_1\pi_2^-$	$\pi_1 - \pi_{12} - \pi_{21} + \pi_{212}$
$--$	π_{212}^-	$1 - \pi_1 - \pi_2 + \pi_{12} + \pi_{21} - \pi_{212}$

For $N \leq 4$, the diagram demipotents are actually idempotent and orthogonal. For larger N, raising the diagram demipotent to a sufficiently large power yields an idempotent (see below 4.8); in other words, the diagram demipotents are demipotent. The power that an diagram demipotent must be raised to in order to obtain an actual idempotent is called its *nilpotence degree*.

For $N = 5$, two of the diagram demipotents need to be squared to obtain an idempotent. For $N = 6$, eight elements must be squared. For $N = 7$, there are four elements that must be cubed, and many others must be squared. Some pretty good upper bounds on the nilpotence degree of the diagram demipotents are given in Section 5. As a preview, for $N > 4$ the nilpotence degree is always $\leq N - 3$, and conditions on the diagram can often greatly reduce this bound.

As an alternative to raising the demipotent to some power, we can express the idempotents as a product of diagram demipotents for smaller diagrams. Let D_k be the signed diagram obtained by taking only the first k nodes of D. Then, as we will see, the idempotents can also be expressed as the product $C_{D_1} C_{D_2} C_{D_3} \cdots C_{D_{N-1}=D}$.

The following is an adaptation of a standard lemma for Coxeter groups to the 0-Hecke algebra, which yields triangularity of the diagram demipotents with respect to the weak order.

Lemma 3.3 *Let m be a standard basis element of the 0-Hecke algebra in the π_i basis. Then for any $i \in D_L(m)$, $\pi_i m = m$, and for any $i \notin D_L(m)$ then $\pi_i m$ is lower than m in left weak order.*

Corollary 3.4 (Diagram Demipotent Triangularity) *Let C_D be a diagram demipotent and m an element of the 0-Hecke monoid in the π_i generators. Then $C_D m = \lambda m + x$, where x is an element of H spanned by monoid elements lower in Bruhat order than m, and $\lambda \in \{0, 1\}$. Furthermore, $\lambda = 1$ if and only if $Des(m)$ is exactly the set of nodes in D marked with pluses.*

Theorem 3.5 *Each diagram demipotent is the sum of a non-zero idempotent part and a nilpotent part. That is, all eigenvalues of a diagram demipotent are either 1 or 0.*

4 Branching

There is a very convenient branching of the diagram demipotents for $H_0(S_N)$ into diagram demipotents for $H_0(S_{N+1})$.

Lemma 4.1 *Let $J = \{i, i+1, \ldots, N-1\}$ Then $w_J^+ \pi_N w_J^+$ is the longest element in the generators i through N. Likewise, $w_J^+ \pi_{i-1} w_J^+$ is the longest element in the generators $i-1$ through $N-1$. Similar statements hold for $w_J^- \pi_N^- w_J^-$ and $w_J^- \pi_{i-1}^- w_J^-$.*

The proof of this lemma relies on the formation of the longest words in the symmetric group; one can find an expression for the longest element in the generators $i - 1$ through $N - 1$ as a subword of the product $w_J^+ \pi_{i-1} w_J^+$. The result then follows immediately.

Recall that each diagram demipotent C_D is the product of two elements L_D and R_D. For a signed diagram D, let $D+$ indicate the diagram with an extra $+$ adjoined at the end. Define $D-$ analogously.

Corollary 4.2 *Let $C_D = L_D R_D$ be the diagram demipotent associated to the signed diagram D for S_N. Then $C_{D+} = L_D \pi_N R_D$ and $C_{D-} = L_D \pi_N^- R_D$. In particular, $C_{D+} + C_{D-} = C_D$.*

Corollary 4.3 *The sum of all diagram demipotents for $H_0(S_N)$ is the identity.*

Next we have a key lemma for proving many of the remaining results in this paper:

Lemma 4.4 (Sibling Rivalry) *Sibling diagram demipotents commute and are orthogonal: $C_{D-} C_{D+} = C_{D+} C_{D-} = 0$. Equivalently, $C_D C_{D+} = C_{D+} C_D = C_{D+}^2$ and $C_D C_{D-} = C_{D-} C_D = C_{D-}^2$.*

The proof uses induction on the tree of diagram demipotents, checking four different cases depending on the last two entries of the diagram D. In particular, it is directly checked that $C_{D+++}C_{D++} = C^2_{D+++}$, and $C_{D+-+}C_{D+-} = C^2_{D+-+}$; all other cases and statements follow from symmetry or application of the automorphism Ψ. The first of these calculations, $C_{D+++}C_{D++} = C^2_{D+++}$, is quite instructive.

Corollary 4.5 *The diagram demipotents C_D are demipotent.*

This follows immediately by induction: if $C_D^k = C_D^{k+1}$, then $C_{D+}C_D^k = C_{D+}C_D^{k+1}$, and by sibling rivalry, $C_{D+}^{k+1} = C_{D+}^{k+2}$.

Now we can say a bit more about the structure of the diagram demipotents.

Proposition 4.6 *Let $p = C_D, x = C_{D+}, y = C_{D-}$, so $p = x + y$ and $xy = 0$. Let v be an element of H. Furthermore, let p, x, and y have abstract Jordan decomposition $p = p_i + p_n$, $x = x_i + x_n$, $y = y_i + y_n$, with $p_ip_n = p_np_i$ and $p_i^2 = p_i$, $p_n^k = 0$ for some k, and similar relations for the Jordan decompositions of x and y.*

Then we have the following relations:

1. *If there exists k such that $p^kv = 0$, then $x^{k+1}v = y^{k+1}v = 0$.*
2. *If $pv = v$, then $x(x-1)v = 0$*
3. *If $(x-1)^kv = 0$, then $(x-1)v = 0$*
4. *If $pv = v$ and $x^kv = 0$ for some k, then $yv = v$.*
5. *If $xv = v$, then $yv = 0$ and $pv = v$.*
6. *Let u_i^x be a basis of the 1-space of x, so that $xu_i^x = u_i^x$, $yu_i^x = 0$ and $pu_i^x = v$, and u_j^y a basis of the 1-space of y. Then the collection $\{u_i^x, u_j^y\}$ is a basis for the 1-space of p.*
7. *$p_i = x_i + y_i$, $p_n = x_n + y_n$, $x_iy_i = 0$.*

The proof follows mainly from simple algebraic manipulations.

Corollary 4.7 *There exists a linear basis v_D^j of $\mathbb{C}H_0(S_N)$, indexed by a signed diagram D and some numbers j, such that the idempotent I_D obtained from the abstract Jordan decomposition of C_D fixes every v_D^j. For every signed diagram $E \neq D$, the idempotent I_E kills v_D^j.*

The proof of the corollary further shows that this basis respects the branching from $H_0(S_{N-1})$ to $H_0(S_N)$. In particular, finding this linear basis for $H_0(S_N)$ allows the easy recovery of the bases for the indecomposable modules for any $M < N$.

We now state the main result. For D a signed diagram, let D_i be the signed sub-diagram consisting of the first i entries of D.

Theorem 4.8 *Each diagram demipotent C_D (see Definition 3.2) for $H_0(S_N)$ is demipotent, and yields an idempotent $I_D = C_{D_1}C_{D_2}\cdots C_D = C_D^N$. The collection of these idempotents $\{I_D\}$ form an orthogonal set of primitive idempotents that sum to 1.*

This follows from the previous result and the construction of the diagram demipotents.

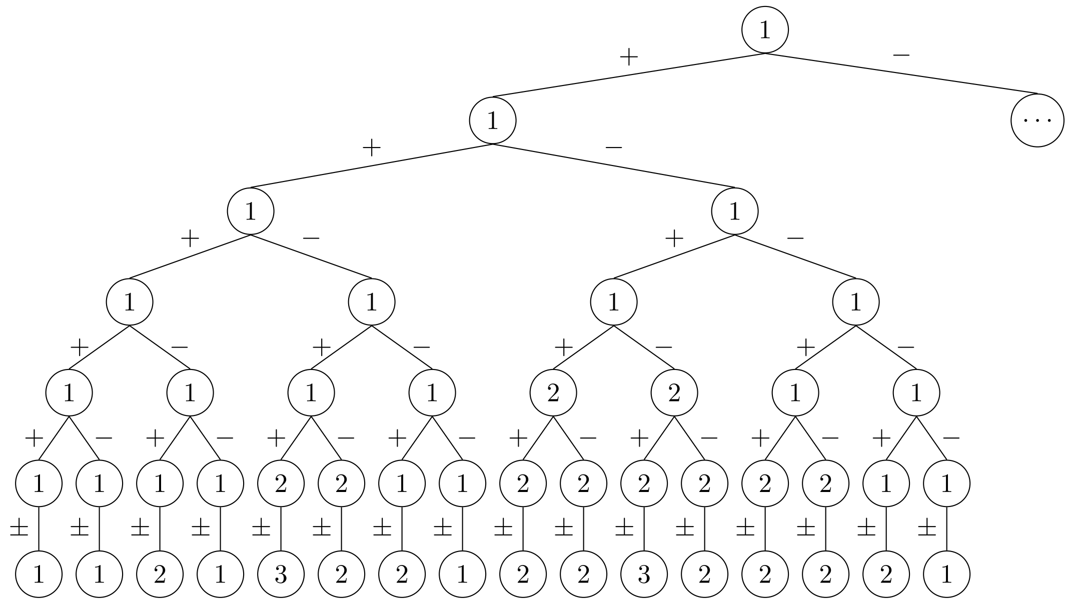


Fig. 2: Nilpotence degree of diagram demipotents. The root node denotes the diagram demipotent with empty diagram (the identity). Since sibling diagram demipotents have the same nilpotence degree, the lowest row has been abbreviated for readability.

5 Nilpotence Degree of Diagram Demipotents

Take any m in the 0-Hecke monoid whose descent set is exactly the set of positive nodes in the signed diagram D. Then $C_D m = m +$ (lower order terms), by a previous lemma, and $I_D m = (C_D)^k(m) = m +$ (lower order terms). The set $\{I_D m | Des(m) = \{\text{positive nodes in D}\}\}$ is thus linearly independent in $H_0(S_N)$, and gives a basis for the projective module corresponding to the idempotent I_D.

We have shown that for any diagram demipotent C_D, there exists a minimal integer k such that $(C_D)^k$ is idempotent. Call k the *nilpotence degree* of C_D. The nilpotence degree of all diagram demipotents for $N \leq 7$ is summarized in Figure 5.

The diagram demipotent $C_{+\cdots+}$ with all nodes positive is given by the longest word in the 0-Hecke monoid, and is thus already idempotent. The same is true of the diagram demipotent $C_{-\cdots-}$ with all nodes negative. As such, both of these elements have nilpotence degree 1.

The following Lemma is easily proved.

Lemma 5.1 *The nilpotence degree of sibling diagram demipotents C_{D+} and C_{D-} is at most one more than the nilpotence degree of the parent C_D. If the nilpotence degree of one sibling is less than or equal to the nilpotence degree of the parent, then the nilpotence degree of the other sibling is equal to the nilpotence degree of the parent.*

Lemma 5.2 *Let D be a signed diagram with a single sign change, or the sibling of such a diagram. Then C_D is idempotent (and thus has nilpotence degree* 1*).*

In particular, this lemma is enough to see why there is no nilpotence before $N = 5$; every signed Dynkin diagrams with three or fewer nodes has no sign change, one sign change, or is the sibling of a diagram with one sign change.

Theorem 5.3 *Let D be any signed diagram with n nodes, and let E be the largest prefix diagram such that E has a single sign change, or is the sibling of a diagram with a single sign change. Then if E has k nodes, the nilpotence degree of D is at most $n - k$.*

This result follows directly from the previous lemma and the fact that the nilpotence degree can increase by at most one with each branching.

This bound is not quite sharp for $H_0(S_N)$ with $N \leq 7$: The diagrams $+ - ++$, $+ - + + +$, and $+ - + + ++$ all have nilpotence degree 2. However, at $N = 7$, the highest expected nilpotence degree is 3 (since every diagram demipotent with three or fewer nodes is idempotent), and this degree is attained by 4 of the demipotents. These diagram demipotents are $+ + - + ++$, $+ - + - ++$, and their siblings.

An open problem is to find a formula for the nilpotence degree directly in terms of the diagram of a demipotent.

6 Remaining Questions

A number of questions still remain.

1. We conjecture that the diagram demipotents C_D have ± 1 coefficients when expanded over $\mathbb{C}$, as this holds for all of the diagram demipotents for $N \leq 8$.

2. Problem: Express the nilpotence degree of C_D as a function of the signed diagram D.

3. Problem: Extend the construction for the idempotents to a more general construction applicable to the 0-Hecke algebra of a general Coxeter group, or, even better, general $\mathcal{J}$-Trivial monoids. The key properties of the idempotents constructed in this paper are construction via a branching rule and invariance of the set of idempotents under the automorphism Ψ; one hopes that a more general construction would retain these properties. One of the impediments to extending to other Coxeter groups is that Lemma 4.1 does not hold for any families of finite Coxeter groups other than S_N, suggesting that other methods of branching must be found.

Acknowledgements

This work was the result of an exploration suggested by Prof. Nicolas M. Thiéry; the notion of branching idempotents was suggested by Alain Lascoux. Additionally, Florent Hivert gave useful insights into working with demipotents elements in an aperiodic monoid. Thanks are also due to my advisor, Anne Schilling, as well as Chris Berg, Andrew Berget, Brant Jones, Steve Pon, and Qiang Wang for their helpful feedback. This research was driven by computer exploration using the open-source mathematical software Sage, developed by Stein et al. (2009) and its algebraic combinatorics features developed by the Sage-Combinat community (2009), and in particular Daniel Bump and Mike Hansen who implemented the Iwahori-Hecke algebras, and the Semigroupe package developed by Jean-Eric Pin.

References

R. W. Carter. Representation theory of the 0-Hecke algebra. *Journal of Algebra*, 104(1):89 – 103, 1986. doi: DOI:10.1016/0021-8693(86)90238-3.

T. Denton, F. Hivert, A. Schilling, and N. M. Thiéry. The representation theory of J-trivial monoids. In preparation, 2010.

O. Ganyushkin, V. Mazorchuk, and B. Steinberg. On the irreducible representations of a finite semigroup. *Proc. Amer. Math. Soc.*, 137(11):3585–3592, 2009. doi: 10.1090/S0002-9939-09-09857-8.

F. Hivert and N. M. Thiéry. The Hecke group algebra of a Coxeter group and its representation theory. *J. Algebra*, 321(8):2230–2258, 2009. doi: 10.1016/j.jalgebra.2008.09.039.

F. Hivert, A. Schilling, and N. M. Thiéry. Hecke group algebras as quotients of affine Hecke algebras at level 0. *J. Combin. Theory Ser. A*, 116(4):844–863, 2009. doi: 10.1016/j.jcta.2008.11.010.

J. E. Humphreys. *Reflection groups and Coxeter groups*, volume 29 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990. ISBN 0-521-37510-X.

P. N. Norton. 0-Hecke algebras. *J. Austral. Math. Soc. Ser. A*, 27(3):337–357, 1979.

A. Okounkov and A. Vershik. A new approach to representation theory of symmetric groups. *Selecta Math. (N.S.)*, 2(4):581–605, 1996. doi: 10.1007/PL00001384.

A. Ram. Seminormal representations of Weyl groups and Iwahori-Hecke algebras. *Proc. London Math. Soc*, 3:7–5, 1997.

T. Sage-Combinat community. Sage-Combinat: enhancing sage as a toolbox for computer exploration in algebraic combinatorics, 2009. `http://combinat.sagemath.org`.

R. P. Stanley. *Enumerative combinatorics. Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997. ISBN 0-521-55309-1; 0-521-66351-2. With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original.

W. Stein et al. *Sage Mathematics Software (Version 3.3)*. The Sage Development Team, 2009. `http://www.sagemath.org`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 584–595

Linear coefficients of Kerov's polynomials: bijective proof and refinement of Zagier's result

Valentin Féray[1] and Ekaterina A. Vassilieva[2]

[1] *LaBRI, Université Bordeaux 1, 351 cours de la libération, 33 400 Talence, France, feray@labri.fr*
[2] *LIX, Ecole Polytechnique, 91128, Palaiseau, France, ekaterina.vassilieva@lix.polytechnique.fr*

Abstract. We look at the number of permutations β of $[N]$ with m cycles such that $(1\ 2\ \dots\ N)\beta^{-1}$ is a long cycle. These numbers appear as coefficients of linear monomials in Kerov's and Stanley's character polynomials. D. Zagier, using algebraic methods, found an unexpected connection with Stirling numbers of size $N+1$. We present the first combinatorial proof of his result, introducing a new bijection between partitioned maps and thorn trees. Moreover, we obtain a finer result, which takes the type of the permutations into account.

Résumé. Nous étudions le nombre de permutations β de $[N]$ avec m cycles telles que $(1\ 2\ \dots\ N)\beta^{-1}$ a un seul cycle. Ces nombres apparaissent en tant que coefficients des monômes linéaires des polynômes de Kerov et de Stanley. À l'aide de méthodes algébriques, D. Zagier a trouvé une connexion inattendue avec les nombres de Stirling de taille N+1. Nous présentons ici la première preuve combinatoire de son résultat, en introduisant une nouvelle bijection entre des cartes partitionnées et des arbres épineux. De plus, nous obtenons un résultat plus fin, prenant en compte le type des permutations.

Keywords: Kerov's Character Polynomials, Bicolored Maps, Long Cycle Factorization

---

1 Introduction

The question of the number of factorizations of the long cycle $(1\ 2\ \dots\ N)$ into two permutations with given number of cycles has already been studied via algebraic or combinatorial[(i)] methods (see [Adr98, SV08]). In these papers, the authors obtain nice generating series for these numbers. Note that the combinatorial approach has been refined to state a result on the number of factorizations of the long cycle $(1\ 2\ \dots\ N)$ in two permutations with given type (see [MV09]).
Unfortunately, in all these results, extracting one coefficient of the generating series gives complicated formulas. The case where one of the two factors has to be also a long cycle is particularly interesting. Indeed, the number $B'(N,m)$ of permutations β of $[N]$ with m cycles, such that $(1\ 2\ \dots\ N)\beta^{-1}$ is a long cycle, is known to be the coefficient of some linear monomial in Kerov's and Stanley's character polynomials (see [Bia03, Theorem 6.1] and [Sta03, Fér10]). These polynomials express the character

---

(i) It can be reformulated in terms of unicellular bicolored maps with given number of vertices, see paragraph 2.1.

value of the irreducible representation of the symmetric group indexed by a Young diagram λ on a cycle of fixed length in terms of some coordinates of λ.
A very simple formula for these numbers was found by D. Zagier [Zag95, Application 3] (see also [Sta09, Corollary 3.3]):

Theorem 1.1 (Zagier, 1995) *Let $m \leq N$ be two positive integers such that $m \equiv N[2]$. Then*

$$\frac{N(N+1)}{2} B'(N,m) = s(N+1,m), \tag{1}$$

where $s(N+1,m)$ is the unsigned Stirling number of the first kind.

It is well-known that $s(N+1,m)$ is the number of permutations of $[N+1]$ with m cycles. As former proofs of this result are purely algebraic, R. Stanley [Sta09] asked for a combinatorial proof of Theorem 1.1. This paper presents the first bijective approach proving this formula. We even prove a finer result, which takes the type[(ii)] of the permutations into consideration and not only their number of cycles. To state it, we need to introduce a few notations. First, we refine the numbers $s(N+1,m)$ and $B'(N,m)$: if $\lambda \vdash n$ (*i.e.* λ is a partition of n), let $A(\lambda)$ (resp. $B(\lambda)$) be the number of permutations $\beta \in S_n$ of type λ (resp. with the additional condition that $(1\ 2\ \dots\ N)\beta^{-1}$ is a long cycle). Then, as Stanley's result deals with permutations of $[N]$ and $[N+1]$, we need operators on partitions which modify their size, but not their length. If μ (resp. λ) has at least one part $i+1$ (resp. i), let $\mu^{\downarrow(i+1)}$ (resp. $\lambda^{\uparrow(i)}$) be the partition obtained from μ (resp. λ) by erasing a part $i+1$ (resp. i) and adding a part i (resp. $i+1$). For instance, using exponential notations (see [Mac95, chapter 1, section 1]), $(1^2 3^1 4^2)^{\downarrow(4)} = 1^2 3^2 4^1$ and $(2^2 3^2 4)^{\uparrow(2)} = 2^1 3^3 4^1$.
Now we can state our main theorem, which implies immediately Theorem 1.1:

Theorem 1.2 (Main result) *For each partition $\mu \vdash N+1$ of length p with $p \equiv N[2]$, one has:*

$$\frac{N+1}{2} \sum_{\lambda=\mu^{\downarrow(i+1)}, i>0} i\, m_i(\lambda)\, B(\lambda) = A(\mu) = \frac{(N+1)!}{z_\mu}, \tag{2}$$

where $m_i(\lambda)$ is the number of parts i in λ and $z_\mu = \prod_i i^{m_i(\mu)} m_i(\mu)!$.

To be comprehensive on the subject, we mention that G. Boccara found an integral formula for $B(\lambda)$ (see [Boc80]), but there does not seem to be any direct link with our result.
As in paper [MV09], the first step (section 2) of our proof of Theorem 1.2 consists in a change of basis in the ring of symmetric functions in order to show the equivalence with the following statement:

Theorem 1.3 *Let λ be a partition of N of length p. Choose randomly (with uniform probability) a set-partition π of $\{1,\dots,N\}$ of type λ and then (again with uniform probability) a permutation β in S_π (that means that each cycle of β is contained in a block of π). Then the probability for $(1\ 2\ \dots\ n)\beta^{-1}$ to be a long cycle is exactly $1/(N-p+1)$.*

Once again, such a simple formula is surprising. We give a bijective proof in sections 3, 4 and 5.

Remark 1 *Theorem 1.2, written for all $\mu \vdash N+1$, gives the collection of numbers $B(\lambda)$ as solution of a sparse triangular system. Indeed, if we endow the set of partitions of N with the lexicographic order, Theorem 1.2, written for $\mu = \lambda^{\uparrow(\lambda_1)}$, gives $B(\lambda)$ in terms of the quantities $A(\mu)$ and $B(\nu)$ with $\nu > \lambda$.*

(ii) The type of a permutation is the sequence of the lengths of its cycles, sorted in increasing order.

2 Link between Theorems 1.2 and 1.3

2.1 Black-partitioned maps

By definition, a map is a graph drawn on a two-dimensional oriented surface (up to deformation), i.e. a graph with a cyclic order on the incident edges to each vertex.

As usual, a couple of permutations (α, β) in S_N can be represented as a bicolored map with N edges labeled with integers from 1 to N. In this identification, $\alpha(i)$ (resp. $\beta(i)$) is the edge following i when turning around its white (resp. black) extremity. White (resp. black) vertices correspond to cycles of α (resp. β). The condition $\alpha \cdot \beta = (1\ 2\ \dots\ N)$ (which we will assume from now on) means that the map is unicellular (i.e. if we remove the edges of the maps from the oriented surface, the resulting surface is homeomorphic to an open disc) and that the positions of the labels are determined by the choice of the edge labeled by 1 (which can be seen as a *root*). In this case, the couple of permutations is entirely determined by β.

Therefore, if $\lambda \vdash N$, the quantity $A(\lambda)$ (resp. $B(\lambda)$) is the number of rooted unicellular maps (resp. *star* maps, that means that we make the additional assumption that the map has only one white vertex) with black vertices' degree distribution λ.
As in the papers [SV08] and [MV09], our combinatorial construction deals with maps with additional information:

Definition 2.1 *A black-partitioned (rooted unicellular) map is a rooted unicellular map with a set partition π of its black vertices. We call degree of a block of π the sum of the degrees of the vertices in π. The* type *of a black-partitioned map is its blocks' degree distribution.*

In terms of permutations, a black-partitioned map consists in a couple (α, β) in S_N with the condition $\alpha\,\beta = (1\ 2\ \dots\ N)$ and a set partition π of $\{1, \dots, N\}$ coarser than the set partition in orbits under the action of β (in other words, if i and j lie in the same cycle of β, they must be in the same part of π).

Example 1 *Let $\beta = (1)(25)(37)(4)(6)$, $\alpha = (1234567)\beta^{-1} = (1267453)$, and π be the partition $\{\{1, 3, 6, 7\}; \{2, 5\}; \{4\}\}$. Associating the triangle, circle and square shape to the blocks, (β, π) is the black-partitioned star map pictured on figure 1.*

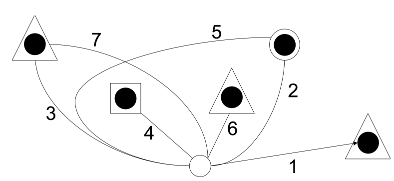


Fig. 1: The black-partitioned map defined in example 1

We denote by $C(\lambda)$ (resp. $D(\lambda)$) the number of black-partitioned maps (resp. black-partitioned star maps) of partition type λ. Equivalently, $C(\lambda)$ (resp. $D(\lambda)$) is the number of couples (β, π) as above such that π is a partition of type λ (resp. and $(1\ 2\ \dots\ N)\beta^{-1}$ is a long cycle). Quantities A and C (resp. B and

D) are linked by the following equations (whose proofs are identical to the one of [MV09, Proposition 1], see also [Mac95, Chapter 1, equation (6.9)])

$$\sum_{\mu \vdash N+1} C(\mu) \operatorname{Aut}(\mu) m_\mu = \sum_{\nu \vdash N+1} A(\nu) p_\nu; \tag{3}$$

$$\sum_{\lambda \vdash N} D(\lambda) \operatorname{Aut}(\lambda) m_\lambda = \sum_{\pi \vdash N} B(\pi) p_\pi, \tag{4}$$

where $m_\bullet$ and $p_\bullet$ denote the monomial and power sum basis of the ring of symmetric functions.

2.2 Permuted star thorn trees and Morales'-Vassilieva's bijection

The main tool of this article is to encode black-partitioned maps into star thorn trees, which have a very simple combinatorial structure. Note that they are a particular case of the notion of thorn trees, introduced by A. Morales and the second author in [MV09].

Definition 2.2 (star thorn tree) *An* (ordered rooted bicolored) star thorn tree *of size N is a tree with a white root vertex, p black vertices and $N-p$ thorns connected to the white vertex (the order in which they are connected matters) and $N-p$ thorns connected to the black vertices. A thorn is an edge connected to only one vertex. We call* type *of a star thorn tree its black vertices' degree distribution (taking the thorns into account). If μ is an integer partition, we denote by $\widetilde{ST}(\mu)$ the number of star thorn trees of type μ.*

An example is given on Figure 2 (for the moment, please do not pay attention to the labels). The interest of this object lies in the following theorem, which corresponds to the case $\lambda = (N)$ of [MV09, Theorem 2] (note that the proof is entirely bijective).

Theorem 2.1 *Let $\mu \vdash N$ be a partition of length p. One has:*

$$C(\mu) = (N-p)! \cdot \widetilde{ST}(\mu). \tag{5}$$

The right-hand side of (5) is the number of couples (τ, σ) where:

- τ is a star thorn tree of type μ.
- σ is a permutation of $[N-p]$, which happens to be exactly the number of thorns with a white (resp. black) extremity in τ. So σ may be seen as a bijection between the thorns with a white extremity and thorns with a black extremity.

We call such a couple a permuted (star) thorn tree. By definition, the type of (τ, σ) is the type of τ. Examples of graphical representations are given on Figure 2: we put symbols on edges and thorns with the following rule. Two thorns have the same label if they are associated by σ and except from that rule, all labels are different (the chosen symbols and their order do not matter, we call that a symbolic labeling).

It is easy to transform a permuted thorn tree (τ, σ) where τ has type $\lambda \vdash N$ to a permuted thorn tree (τ', σ') where τ' has type $\mu = \lambda^{\uparrow(i)}$. We just add a thorn anywhere on the white vertex ($N+1$ possible places) and a thorn anywhere on a black vertex of degree i (there are i possible places on each of the $m_i(\lambda)$ black vertices of degree i). Then we choose σ' to be the extension of σ associating the two new thorns. This procedure is invertible if we remember which thorn of black extremity is the new one (it

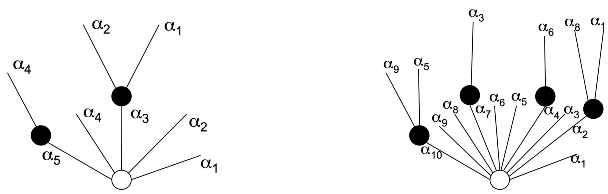


Fig. 2: Example of two permuted star thorn trees $(\tau^1_{\text{ex}}, \sigma^1_{\text{ex}})$ of type $1^1 2^1$ and $(\tau^2_{\text{ex}}, \sigma^2_{\text{ex}})$ of type $2^2 3^2$

must be on a black vertex of degree $i+1$, so there are $i \cdot m_{i+1}(\mu)$ choices). This leads immediately to the following relation:

$$\widetilde{ST}(\mu) \cdot (N+1-p)! \cdot i \cdot m_{i+1}(\mu) = (N+1) \cdot i \cdot m_i(\lambda) \cdot \widetilde{ST}(\lambda) \cdot (N-p)!. \tag{6}$$

2.3 Reduction of the main theorem

Proposition 2.2 *For any partition $\lambda \vdash N$ of length p, one has:*

$$D(\lambda) = \frac{1}{N-p+1}(N-p)!\widetilde{ST}(\lambda).$$

Sections 3, 4 and 5 are devoted to the proof. It is easy to see, with the definition of subsection 2.1 and the bijection of subsection 2.2, that this proposition is a reformulation of Theorem 1.3.

Lemma 2.3 *Proposition 2.2 implies Theorem 1.2.*

Proof: We fix a partition $\mu \vdash N+1$ of length $p < N+1$ and sum equation (6) on $\lambda = \mu^{\downarrow(i+1)}$:

$$\widetilde{ST}(\mu) \cdot (N+1-p)! \cdot (N+1-p) = (N+1) \sum_{\lambda=\mu^{\downarrow(i+1)}, i>0} i \cdot m_i(\lambda) \cdot \widetilde{ST}(\lambda) \cdot (N-p)!.$$

Using Morales'-Vassilieva's bijection and Proposition 2.2, this equality becomes:

$$C(\mu) \cdot (N+1-p) = (N+1) \sum_{\lambda=\mu^{\downarrow(i+1)}, i>0} i \cdot m_i(\lambda) \cdot D(\lambda) \cdot (N+1-p).$$

$$\text{Hence,} \quad \sum_{\substack{\mu \vdash N+1 \\ \mu \neq 1^{(N+1)}}} C(\mu)\,\mathrm{Aut}(\mu) m_\mu = (N+1) \sum_{\substack{\mu \vdash N+1 \\ \mu \neq 1^{(N+1)}}} \sum_{\substack{i>0 \\ \lambda=\mu^{\downarrow(i+1)}}} i \cdot m_i(\lambda)\,\mathrm{Aut}(\mu) D(\lambda) m_\mu;$$

$$= (N+1) \sum_{\lambda \vdash N} \mathrm{Aut}(\lambda) D(\lambda) \left(\sum_{\substack{i>0 \\ \mu=\lambda^{\uparrow(i)}}} i \cdot m_{i+1}(\mu) m_\mu \right).$$

The last equality has been obtained by changing the order of summation and using the trivial fact that, if $\mu = \lambda^{\uparrow(i)}$, one has $\mathrm{Aut}(\mu) \cdot m_i(\lambda) = \mathrm{Aut}(\lambda) \cdot m_{i+1}(\mu)$. Now, observing that the expression in the bracket can be written $\Delta(m_\lambda)$, where Δ is the differential operator $\sum_i x_i^2 \partial / \partial x_i$, one has:

$$\sum_{\mu \vdash N+1} C(\mu)\,\mathrm{Aut}(\mu) m_\mu - (N+1)! m_{1^{N+1}} = (N+1) \cdot \Delta \left(\sum_{\lambda \vdash N} \mathrm{Aut}(\lambda) D(\lambda) m_\lambda \right).$$

Let us rewrite this equality in the power sum basis. The expansion of the two summations in this basis are given by equations (3) and (4). Furthermore, one has: $\Delta(p_\pi) = \sum_i i \cdot m_i(\pi) p_{\pi^{\uparrow(i)}}$. Indeed, the one-part case ($\Delta(p_k) = k \cdot p_{k+1}$) is trivial and the general case follows because Δ is a derivation. We also need the power-sum expansion of $(N+1)! m_{1^{N+1}}$, which is (see [Mac95, Chapter I, equation (2.14')]):

$$(N+1)! m_{1^{N+1}} = (N+1)! e_{N+1} = (N+1)! \sum_{\nu \vdash N+1} \frac{(-1)^{N+1-\ell(\nu)}}{z_\nu} p_\nu = \sum_{\nu \vdash N+1} A(\nu)(-1)^{N+1-\ell(\nu)} p_\nu,$$

where e_{N+1} is the $N+1$-th elementary function. Putting everything together, we have:

$$\sum_{\nu \vdash N+1} A(\nu) p_\nu + \sum_{\nu \vdash N+1} A(\nu)(-1)^{N-\ell(\nu)} p_\nu = (N+1) \sum_\pi B(\pi) \sum_{\rho = \pi^{\uparrow(i)}, i>0} i \cdot m_i(\pi) p_\rho.$$

If we identify the coefficients of p_μ in both sides, we obtain exactly Theorem 1.2 □

Remark 2 *Using Remark 1, the converse statement of Lemma 2.3 can be proved the same way.*

3 Mapping black-partitioned star maps to permuted thorn trees

The following sections provide a combinatorial proof of Proposition 2.2. We proceed in a three step fashion. First we define a mapping Ψ from the set of black-partitioned star maps of type λ (counted by $D(\lambda)$) to a set of permuted star thorn trees of the same type and show it is injective. As a final step, we compute the cardinality of the image set of Ψ and show it is exactly $(1/(N-p+1))\,(N-p)!\widetilde{ST}(\lambda)$. Although there are some related ideas, Ψ is not the restriction of the bijection of paper [MV09].

3.1 Labeled thorn tree

Let (β, π) be a black-partitioned star map. We first construct a labeled star thorn tree $\overline{\tau}$:

(i) Let $(\alpha_k)_{(1 \le k \le N)}$ be integers such that $\alpha_1 = 1$ and the long cycle $\alpha = (1\ 2\ \dots\ N)\beta^{-1}$ is equal to $(\alpha_1 \alpha_2 \alpha_3 \dots \alpha_N)$. The root of $\overline{\tau}$ is a white vertex with N descending edges labeled from right to left with $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N$ (α_1 is the rightmost descending edge and α_N the leftmost).

(ii) Let m_i be the maximum element of block π_i. For $k = 1 \dots N$, if $\alpha_k = \beta(m_i)$ for some i, we draw a black vertex at the other end of the descending edge labeled with α_k. Otherwise the descending edge is a thorn.

Remark 3 *As $\alpha_N = \alpha^{-1}(1) = \beta(N)$ the leftmost descending edge isn't a thorn and is labeled with $\beta(N)$ (N is necessarily the maximum element of the block containing it).*

(iii) For $i = 1 \dots p$, let $(\beta_1^u \dots \beta_{l_u}^u)_{1 \le u \le c}$ be the c cycles included in block π_i such that $\beta_{l_u}^u$ is the maximum element of cycle u. (We have $\Sigma_u l_u = \mid \pi_i \mid$). We also order these cycles according to their maximum, *i.e.* we assume that $\beta_{l_c}^c < \beta_{l_{c-1}}^{c-1} < \dots < \beta_{l_1}^1 = m_i$. As a direct consequence, $\beta_1^1 = \beta(m_i)$.
We connect $\mid \pi_i \mid -1$ thorns to the black vertex linked to the root by the edge $\beta(m_i)$. Moving around the vertex counter-clockwise and starting right after edge $\beta(m_i)$, we label the thorn with $\beta_2^1, \beta_3^1, \dots, \beta_{l_1}^1, \beta_1^2, \dots, \beta_{l_2}^2, \beta_{l_3}^3, \dots \beta_{l_c}^c$. Then $\overline{\tau}$ is the resulting thorn tree.

Remark 4 *Moving around a black vertex clockwise starting with the thorn right after the edge (in clockwise order), a new cycle of β begins whenever we meet a left-to-right maximum of the labels.*

Remark 5 *As the long cycle α and the repartition of the cycles of β in the various blocks appear explicitly in $\overline{\tau}$, one can recover the black-partitioned star map from it.*

The idea behind this construction is to add a root to the map (α, β), select one edge per block, cut all other edges into two thorns and merge the vertices corresponding to the same black block together. Step (i) tells us where to place the root, step (ii) which edges we select and step (iii) how to merge vertices (in maps unlike in graphs, one has to do some choices to merge vertices).

Example 2 *Let us take the black-partitioned star map of example 1. Following construction rules (i) and (ii), one has $m_\triangle = 7$, $m_\bigcirc = 5$, $m_\square = 4$ and the descending edges indexed by $\beta(m_\triangle) = 3$, $\beta(m_\bigcirc) = 2$ and $\beta(m_\square) = 4$ connect a black vertex to the white root. Other descending edges from the root are thorns. Using (iii), we add labeled thorns to the black vertices to get the labeled thorn tree depicted on Figure 3. Focusing on the one connected to the root through the edge 3, we have $(\beta_1^1 \beta_{l_1}^1)(\beta_{l_2}^2)(\beta_{l_3}^3) = (37)(6)(1)$. Reading the labels clockwise around this vertex, we get $1, 6, 7, 3$. The three cycles can be simply recovered looking at the left-to-right maxima 1, 6 and 7.*

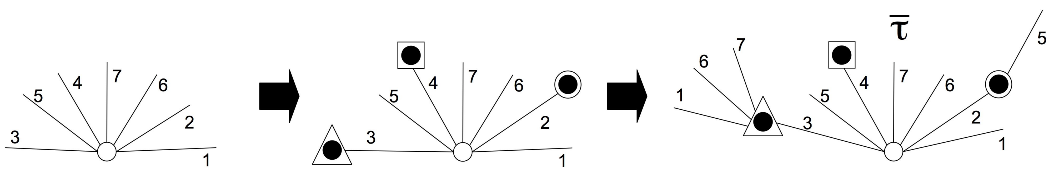


Fig. 3: Labeled thorn tree associated to the black-partitioned star map of Figure 1

3.2 Permuted thorn tree

Using $\overline{\tau}$, we call τ the star thorn tree obtained by removing labels and σ the permutation that associates to a white thorn in τ the black thorn with the same label in $\overline{\tau}$.

Finally, we define: $\Psi(\beta, \pi) = (\tau, \sigma)$.

Example 3 *Following up with example 1, we get the permuted thorn tree $(\tau_{ex}^3, \sigma_{ex}^3)$ drawn on Figure 4. Graphically we use the same convention as in section 2 to represent σ.*

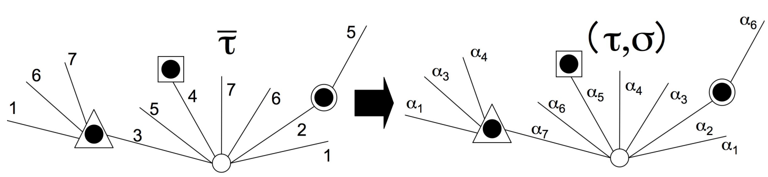


Fig. 4: Permuted thorn tree $(\tau_{\text{ex}}^3, \sigma_{\text{ex}}^3)$ associated to the black-partitioned star map of Figure 1

4 Injectivity and reverse mapping

Assume $(\tau, \sigma) = \Psi(\beta, \pi)$ for some black partitioned star map (β, π). We show that (β, π) is actually uniquely determined by (τ, σ).
As a first step, we recover the labeled thorn tree $\overline{\tau}$. Let us draw the permuted thorn tree (τ, σ) as explained in subsection 2.2. We show by induction that there is at most one possible integer value for each symbolic label.

(i) By construction, the label α_1 of the right-most edge or thorn descending from the root is necessarily 1.

(ii) Assume that for $i \in [N-1]$, we have identified the symbols of values $1, 2, \ldots$, and i. We look at the edge or thorn with label i connected to a black vertex b. In this step, we determine which symbol corresponds to $\beta(i)$.
Recall that, when we move around b clockwise finishing with the edge (in this step, we will always turn in this sense), a new cycle begins whenever we meet a left-to-right maximum. So, to find $\beta(i)$, one has to know whether i is a left-to-right maximum or not.
If all values of labels of thorns before i haven't already been retrieved, then i is not a left-to-right maximum. Indeed, the remaining label values are $i+1, \ldots, N$ and at least one thorn's label on the left of i lies in this interval. Following our construction, necessarily $\beta(i)$ corresponds to the symbolic label of the thorn right at the left of i.
If all the thorns' label values on the left of i have already been retrieved (or there are no thorns at all), then i is a left-to-right maximum. According to the construction of $\overline{\tau}$, $\beta(i)$ corresponds necessarily to the symbolic label of the thorn preceding the next left-to-right maximum. But one can determine which thorn (or edge) corresponds to the next left-to-right maximum: it is the first thorn (or edge) e without a label value retrieved so far (again moving around the black vertex from left to right). Indeed, all the value retrieved so far are less than i and those not retrieved greater than i. Therefore $\beta(i)$ is the thorn right at the left of e. If all the values of the labels of the thorns connected to b have already been retrieved then i is the maximum element of the corresponding block and $\beta(i)$ corresponds to the symbolic label of the edge connecting this black vertex to the root.

(iii) Consider the element (thorn of edge) of white extremity with the symbolic label corresponding to $\beta(i)$. The next element (turning around the root in counter-clockwise order) has necessarily label $\alpha(\beta(i)) = i+1$.
As a result, the knowledge of the thorn or edge with label i uniquely determines the edge or thorn with label $i+1$.

Applying the previous procedure up to $i = N-1$ we see that $\overline{\tau}$ is uniquely determined by (τ, σ) and so is (β, π) (see Remark 5).

Example 4 *Take as an example the permuted thorn tree $(\tau_{ex}^1, \sigma_{ex}^1)$ drawn on the left-hand side of Figure 2, the procedure goes as described on figure 5. First, we identify $\alpha_1 = 1$. Then, as there is a non (value) labeled thorn (α_2) on the left of the thorn connected to a black vertex with label value 1, necessarily 1 is not a left-to-right maximum and α_2 is the label of the thorn right on the left of 1, that is α_2. Then as α_3 follows $\alpha_2 = \beta(1)$ around the white root, we have $\alpha_3 = \alpha(\beta(1)) = 2$.*
We apply the procedure up to the full retrieval of the edges' and thorns' labels. We find $\alpha_2 = 3$, $\alpha_4 = 4$, $\alpha_5 = 5$. Finally, we have $\alpha = (13245)$, $\beta = (213)(4)(5)$, $\pi = \{\{1,2,3\};\{4,5\}\}$ as shown on figure 5.

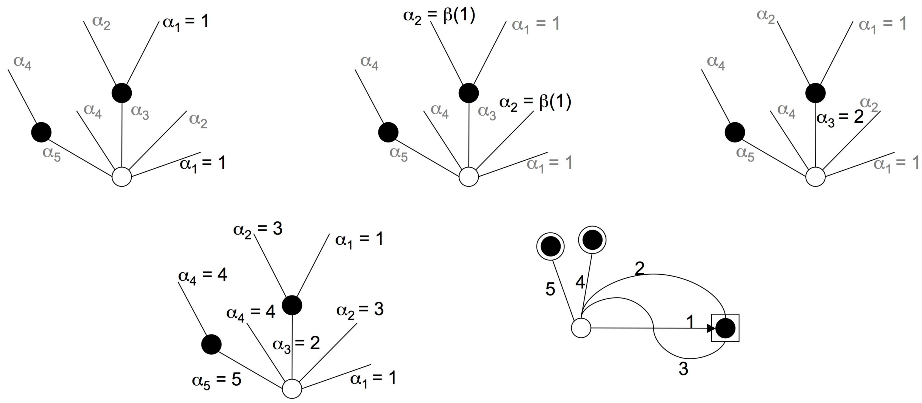


Fig. 5: Reconstruction of the map

5 Characterization and size of the image set $\Im(\Psi)$

5.1 A necessary and sufficient condition to belong to $\Im(\Psi)$

Why Ψ is not surjective? Let us fix a permuted star thorn tree (τ, σ). We can try to apply to it the procedure of section 4 and we distinguish two cases:

- it can happen, for some $i < N$, when one wants to give the label $i+1$ to the edge following $\beta(i)$ (step (iii)), that this edge has already a label j. If so, the procedure fails and (τ, σ) is not in $\Im(\Psi)$.
- if this never happens, the procedure ends with a labeled thorn tree $\overline{\tau}$. In this case, one can find the unique black-partitioned star map M corresponding to $\overline{\tau}$ and by construction $\Psi(M) = (\tau, \sigma)$.

For instance, if we take as (τ, σ) the couple $(\tau_{\text{ex}}^2, \sigma_{\text{ex}}^2)$ on the right of Figure 2, the procedure gives successively : $\alpha_1 = 1,\ \alpha_9 = 2,\ \alpha_{10} = 3,\ \alpha_6 = 4,\ \alpha_5 = 5$ and then we should choose $\alpha_1 = 6$, but this is impossible because we already have $\alpha_1 = 1$.

Lemma 5.1 *If the procedure fails, the label j of the edge that should get a second label $i+1$ is always* 1.

Proof: If $j > 1$, this means that $\beta(j-1) = \beta(i)$. Let us distinguish two cases.
If $j-1$ is a left-to-right maximum, the label i must be at the right of $\beta(j-1)$ and not a left-to-right maximum. But this is impossible because all thorns at the left of $\beta(j-1)$ (including $\beta(j-1)$) have labels smaller than j.
If $j-1$ is not a left-to-right maximum, the label $j-1$ must be at the right of $\beta(j-1) = \beta(i)$ and i is a left-to-right maximum. Then $\beta(i)$ is before the next left-to-right maximum. So the edge at the right of $\beta(i)$ has a label greater than i and can not be $j-1$. □

An auxiliary oriented graph Remark 3 gives a necessary condition for (τ, σ) to be in $\Im(\Psi)$: its leftmost edge leaving the root must be a real edge e_0, and not a thorn. From now on, we call this property $(P1)$: note that, among all permuted thorn tree of a given type $\lambda \vdash N$ of length p, exactly p over N have this property. When $(P1)$ is satisfied, we denote π_0 the black extremity of e_0. The lemma above shows

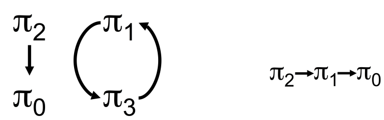


Fig. 6: Two examples of auxiliary graphs.

that the procedure fails if and only if e_0 is chosen as $\beta(i)$ for some $i < N$. But this can not happen at any time. Indeed, the following lemma is a direct consequence from step (ii) of the inverting procedure:

Lemma 5.2 *A real edge (i.e. which is not a thorn) e can be chosen as $\beta(i)$ only if the edge and all thorns leaving the corresponding black vertex have labels smaller or equal to i. If this happen, we say that the black vertex is completed at step i.*

Corollary 5.3 *Let e be a real edge of black extremity $\pi \neq \pi_0$. Let us denote e' the element (edge or thorn) right at the left of e on the white vertex. Let π' be the black extremity of the element e'' associated to e' (i.e. e' itself if it is an edge, its image by σ else). Then π' can not be completed before π.*

Proof: If π' is completed at step i, by Lemma 5.2, element e'' has a label $j \leq i$. As e' has the same label, this implies that e has label $\beta(j-1)$ or in other words, that π is completed at step $j-1 < i$. $\square$

When applied for every black vertex $\pi \neq \pi_0$, this corollary gives some partial information on the order in which the black vertices can be completed. We will summarize this in an oriented graph $G(\tau, \sigma)$: its vertices are the black vertices of τ and its edges are $\pi \to \pi'$, where π and π' are in the situation of the corollary above. This graph has one edge leaving each of its vertex, except for π_0. As examples, the graphs corresponding to $(\tau^2_{\text{ex}}, \sigma^2_{\text{ex}})$ and to $(\tau^3_{\text{ex}}, \sigma^3_{\text{ex}})$ (see Figures 2 and 4) are drawn on Figure 6.

The graph $G(\tau, \sigma)$ gives all the information we need! Can we decide, using only $G(\tau, \sigma)$, whether (τ, σ) belongs to $\Im(\Psi)$ or not? There are two cases, in which the answer is obviously yes:

1. Let us suppose that $G(\tau, \sigma)$ is an oriented tree of root π_0 (all edges are oriented towards the root). In this case, we say that (τ, σ) has property $(P2)$. Then, the vertex π_0 can be completed only when all other vertices have been completed, *i.e.* when all edges and thorns have already a label. That means that e_0 can be chosen as $\beta(i)$ only for $i = N$. Therefore, in this case, the procedure always succeeds and (τ, σ) belongs to $\Im(\Psi)$. This is the case of $(\tau^3_{\text{ex}}, \sigma^3_{\text{ex}})$.

2. Let us suppose that $G(\tau, \sigma)$ contains an oriented cycle (eventually a loop). Then all the vertices of this cycle can never be completed. Therefore, the procedure always fails in this case and (τ, σ) does not belong to $\Im(\Psi)$. This is the case of $(\tau^2_{\text{ex}}, \sigma^2_{\text{ex}})$.

In fact, we are always in one of these two cases (the proof of the following lemma is left to the reader):

Lemma 5.4 *Let G be an oriented graph whose vertices have out-degree 1, except for one vertex v_0 which has out-degree 0. Then G is either an oriented tree of root v_0 or contains an oriented cycle.*

Finally, one has the following result:

Proposition 5.5 *The mapping Ψ defines a bijection:*

$$\left\{ \begin{array}{c} \textit{black-partitioned star maps} \\ \textit{of type } \lambda \end{array} \right\} \simeq \left\{ \begin{array}{c} \textit{permuted star thorn trees of type } \lambda \\ \textit{with properties } (P1) \textit{ and } (P2) \end{array} \right\}. \qquad (7)$$

5.2 Proportion of permuted thorn trees (τ, σ) in $\Im(\Psi)$

To finish the proof of Proposition 2.2, one just has to compute the size of the right-hand side of (7):

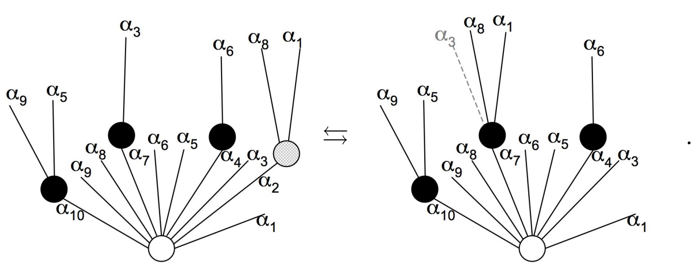
Proposition 5.6 *Let λ be a partition of N of length p. Denote by $P(\lambda)$ (resp. $P'(\lambda)$) the proportion of couples (τ, σ) with properties $(P1)$ and $(P2)$ among all the permuted thorn trees of type λ (resp. among permuted thorn trees of type λ with property $(P1)$) of type λ. Then, one has:*

$$P'(\lambda) = \frac{N}{p(N-p+1)} \text{ and, hence, } P(\lambda) = \frac{1}{N-p+1}.$$

Proof (by induction on p): The case $p = 1$ is easy: as $G(\tau, \sigma)$ has only one vertex and no edges, it is always a tree. Therefore, for any $N \geq 1$, one has $P'((N)) = 1$.
Suppose that the result is true for any λ of length $p-1$ and fix a partition $\mu \vdash N$ of length $p > 1$. Consider the permuted thorn trees (τ, σ) of type μ, verifying $(P1)$, with a marked black vertex $\overline{\pi} \neq \pi_0$: as there are always $p-1$ choices for the marked vertex, the proportion of these objects verifying $(P2)$ is still $P'(\mu)$. Let us now split this set, depending on the degrees (in τ) of the marked vertex and of the end of the edge leaving $\overline{\pi}$ in the graph $G(\tau, \sigma)$. The proportion of marked star thorn trees of type μ (with property $(P1)$) whose marked vertex has degree k_0 is $m_{k_0}(\mu)/p$. We denote $k = \deg_\tau(\overline{\pi})$ and $\mu' = \mu \backslash k$ (*i.e.* the partition obtained from μ by deleting one part k).

- In $k-1$ cases over $N-1$, this second extremity is also $\overline{\pi}$. So $G(\tau, \sigma)$ contains a loop and (τ, σ) does not fulfill $(P2)$.
- For every j, in $j \cdot m_j(\mu')$ cases over $N-1$, this second extremity is a vertex $\pi' \neq \overline{\pi}$ of degree j (in τ). But one has an easy bijection φ:

$$\left\{ \begin{array}{c} (\tau, \sigma) \text{ of type } \mu \text{ verifying } (P1) \\ \text{with a marked black vertex } \overline{\pi} \neq \pi_0 \\ \text{of size } k \text{ such that } \overline{\pi} \to_{G(\tau,\sigma)} \pi' \\ \text{with } \pi' \neq \overline{\pi} \text{ of size } j \end{array} \right\} \simeq \left\{ \begin{array}{c} (\tau', \sigma') \text{ of type } \mu^{\downarrow(j,k)} := \mu \backslash (j,k) \cup (j+k-1) \\ \text{verifying } (P1) \text{ with the edge or one of the first } j-1 \\ \text{thorns of a black vertex of size } j+k-1 \text{ marked} \\ \text{(always } j \cdot m_{j+k-1}(\mu^{\downarrow(j,k)}) \text{ choices)} \end{array} \right\}$$



From left-to-right: erase the marked black vertex $\overline{\pi}$ with its edge $e_{\overline{\pi}}$ and move its thorns to the black vertex π' (at the right of its own thorns). Choose as marked the element (edge of thorn) with a black extremity with the same symbolic label as the element right at the left of $e_{\overline{\pi}}$.
From right to left: look at the white thorn corresponding to the marked thorn $\overline{e}$ (if the marked element is an edge, just take the edge itself). Then add a new edge with a black vertex just at the right

of this thorn (or edge). Finally, move the $k-1$ right-most thorns of the black extremity of $\overline{e}$ to this new black vertex. The marked black vertex is the new one.
This bijection keeps property (P2). Indeed, if $\varphi(\tau,\sigma,\overline{\pi}) = (\tau',\sigma',\overline{e})$, the graph $G(\tau',\sigma')$ is obtained from $G(\tau,\sigma)$ by contracting the edge of origin $\overline{\pi}$. Therefore, the proportion of couples having property $(P2)$ on the left-hand side is the same as on the right-hand side. But, as $\mu^{\downarrow(j,k)}$ has length $p-1$ and size $N-1$, by induction hypothesis, this proportion is:

$$\frac{N-1}{(p-1)\big((N-1)-(p-1)+1\big)}.$$

We can now put the different cases together to compute $P'(\mu)$:

$$P'(\mu) = \sum_k \frac{m_k(\mu)}{p}\left(\sum_j \frac{j\cdot m_j(\mu')}{N-1}\cdot\frac{N-1}{(p-1)\big(N-p+1\big)}\right) = \frac{N}{p\big(N-p+1\big)}.$$

The last equality is obtained by a straight-forward computation and ends the proof of Proposition 5.6 and, therefore, of Proposition 2.2. □

References

[Adr98] NM Adrianov. An analogue of the Harer-Zagier formula for unicellular two-color maps. *Funct. Anal. Appl*, 31(3):149–155, 1998.

[Bia03] P. Biane. Characters of symmetric groups and free cumulants. In *Asymptotic combinatorics with applications to mathematical physics (St. Petersburg, 2001)*, volume 1815 of *Lecture Notes in Math.*, pages 185–200. Springer, Berlin, 2003.

[Boc80] G. Boccara. Nombre de reprśentations d'une permutation comme produit de deux cycles de longueurs donnés. *Discrete Math.*, 29:105–134, 1980.

[Fér10] V. Féray. Stanley's formula for characters of the symmetric group. *Annals of Combinatorics*, 13(4):453–461, 2010.

[Mac95] I.G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Univ P., 2nd edition, 1995.

[MV09] A. Morales and E. Vassilieva. Bijective enumeration of bicolored maps of given vertex degree distribution. *DMTCS Proceedings (FPSAC)*, AK:661–672, 2009.

[Sta03] R. P. Stanley. Irreducible symmetric group characters of rectangular shape. *Sém. Lothar. Combin.*, 50:Art. B50d, 11 pp. (electronic), 2003.

[Sta09] R.P. Stanley. Two enumerative results on cycles of permutations. to appear in European J. Combinatorics, 2009.

[SV08] G. Schaeffer and E. Vassilieva. A bijective proof of Jackson's formula for the number of factorizations of a cycle. *J. Comb. Theory, Ser. A*, 115(6):903–924, 2008.

[Zag95] D. Zagier. On the distribution of the number of cycles of elements in symmetric groups. *Nieuw Arch. Wisk.*, 13(3):489–495, 1995.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 596–607

Balanced binary trees in the Tamari lattice

Samuele Giraudo[1]

[1]*Institut Gaspard Monge, Université Paris-Est Marne-la-Vallée, 5 Boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, France*

Abstract. We show that the set of balanced binary trees is closed by interval in the Tamari lattice. We establish that the intervals $[T_0, T_1]$ where T_0 and T_1 are balanced trees are isomorphic as posets to a hypercube. We introduce tree patterns and synchronous grammars to get a functional equation of the generating series enumerating balanced tree intervals.

Résumé. Nous montrons que l'ensemble des arbres équilibrés est clos par intervalle dans le treillis de Tamari. Nous caractérisons la forme des intervalles du type $[T_0, T_1]$ où T_0 et T_1 sont équilibrés en montrant qu'en tant qu'ensembles partiellement ordonnés, ils sont isomorphes à un hypercube. Nous introduisons la notion de motif d'arbre et de grammaire synchrone dans le but d'établir une équation fonctionnelle de la série génératrice qui dénombre les intervalles d'arbres équilibrés.

Keywords: balanced trees, Tamari lattice, posets, grammars, generating series, combinatorics

1 Introduction

Binary search trees are used as data structures to represent dynamic totally ordered sets [7, 6, 3]. The algorithms solving classical related problems such as the insertion, the deletion or the search of a given element can be performed in a time logarithmic in the cardinality of the represented set, provided that the encoding binary tree is balanced. Recall that a binary tree is balanced if for each node x, the height of the left subtree of x and the height of the right subtree of x differ by at most one.

The algorithmic of balanced trees relies fundamentally on the so-called rotation operation. An insertion or a deletion of an element in a dynamic ordered set modifies the tree encoding it and can imbalance it. The efficiency of these algorithms comes from the fact that binary search trees can be rebalanced very quickly after the insertion or the deletion, using no more than two rotations [2].

Surprisingly, this operation appears in a different context since it defines a partial order on the set of binary trees of a given size. A tree T_0 is smaller than a tree T_1 if it is possible to transform the tree T_0 into the tree T_1 by performing a succession of right rotations. This partial order, known as the Tamari order [8, 10], defines a lattice structure on the set of binary trees of a given size.

Since binary trees are naturally equipped with this order structure induced by rotations, and the balance of balanced trees is maintained doing rotations, we would like to investigate if balanced trees play a particular role in the Tamari lattice. Our goal, in this is paper, is to combine the two points of view of the rotation operation. A first simple computer observation is that the intervals $[T_0,\ T_1]$ where T_0 and T_1 are balanced trees are only made up of balanced trees. The main goal of this paper is to prove this property. As a consequence, we give a characterization on the shape of these intervals and, using grammars allowing to generate trees, enumerate them.

This article is organized as follows. In Section 2, we set the essential notions about binary trees and balanced trees, and we give the definition of the Tamari lattice in our setting. Section 3 is devoted to establish the main result: the set of balanced trees is closed by interval in the Tamari lattice. In Section 4, we define tree patterns and synchronous grammars. These grammars allow us to generate trees avoiding a given set of tree patterns. We define a subset of balanced trees where elements hold a peculiar position in the Tamari lattice and we give, using the synchronous grammar generating these, a functional equation of the generating series enumerating these. Finally, in Section 5, we look at balanced tree intervals and show that they are, as posets, isomorphic to hypercubes. Encoding balanced tree intervals by particular trees, and establishing the synchronous grammar generating these trees, we give a functional equation satisfied by the generating series enumerating balanced tree intervals.

Acknowledgments

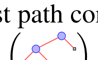
The author would like to thank Florent Hivert for introducing him to the problem addressed in this paper, and Jean-Christophe Novelli and Florent Hivert for their invaluable advice and their improvement suggestions. The computations of this work have been done with the open-source mathematical software Sage [1].

2 Preliminaries

2.1 Complete rooted planar binary trees

In this article, we consider complete rooted planar binary trees. Nodes are denoted by circles like  and leaves by squares like . The empty tree is also denoted by . Assuming L and R are complete rooted planar binary trees, let $L \wedge R$ be the (unique) complete rooted planar binary tree which has L as left subtree and R as right subtree. Let also $\mathcal{T}_n$ be the set of complete rooted planar binary trees with n nodes and $\mathcal{T}$ be the set of all complete rooted planar binary trees. We use in the sequel the standard terminology (ie. *child*, *ancestor*, *edge*, *path*, ...) about complete rooted planar binary trees [3].

Recall that the nodes of a complete rooted planar binary tree T can be visited in the infix order: it consists in visiting recursively the left subtree of T, then the root, and finally the right subtree. We say that a node y is *on the right compared to* a node x in T if the node x appears strictly before the node y in the infix order and we denote that by $x \rightsquigarrow_T y$. We extend this notation to subtrees saying that a subtree S of root y of T is on the right compared to a node x in T if for all nodes y' of S we have $x \rightsquigarrow_T y'$. We say that a node x of T is the *leftmost node* of T if x is the first visited node in the infix order.

If T is a complete rooted planar binary tree, we shall denote by $\mathrm{ht}(T)$ the *height* of T, that is the length of the longest path connecting the root of T to one of its leaves. For example, we have $\mathrm{ht}() = 0$, $\mathrm{ht}\left(\right.$$\left.\right) = 1$, and $\mathrm{ht}\left(\right.$$\left.\right) = 2$.

In the sequel, we shall mainly talk about complete rooted planar binary trees so we shall call them simply *trees*.

2.2 Balanced trees

Let us define, for each tree T, the mapping γ_T called the *imbalance mapping* which associates an element of $\mathbb{Z}$ with a node x of T, namely the *imbalance value* of x. It is defined for a node x by:

$$\gamma_T(x) = \mathrm{ht}(R) - \mathrm{ht}(L) \tag{2.1}$$

where L (resp. R) is the left (resp. right) subtree of x.

Balanced trees form a subset of $\mathcal{T}$ composed of trees which have the property of being balanced:

Definition 2.1 *A tree T is* balanced *if for all node x of T, we have*

$$\gamma_T(x) \in \{-1, 0, 1\}. \tag{2.2}$$

Let us denote by $\mathcal{B}_n$ the set of balanced trees with n nodes (see Figure 1 for the first sets) and $\mathcal{B}$ the set of all balanced trees.

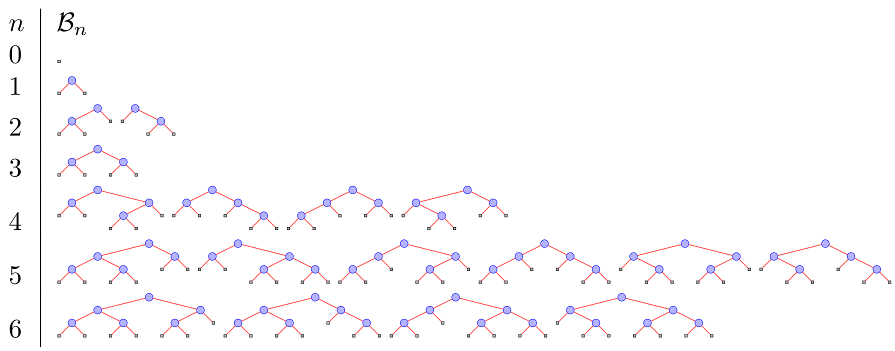


Fig. 1: The first balanced trees.

2.3 The Tamari lattice

The Tamari lattice can be defined in several ways [10, 5] depending on which kind of catalan object (ie. in bijection with trees) the order relation is defined. We give here the most convenient definition for our use. First, let us recall the right rotation operation:

Definition 2.2 *Let T_0 be a tree and $S_0 = (A \wedge B) \wedge C$ be the subtree of root y of T_0. If T_1 is the tree obtained by replacing the tree S_0 by the tree $A \wedge (B \wedge C)$ in T_0 (see Figure 2), we say that T_1 is obtained from T_0 by a* right rotation *of* root y.

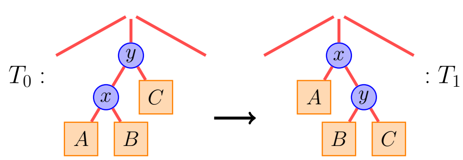


Fig. 2: The right rotation of root y.

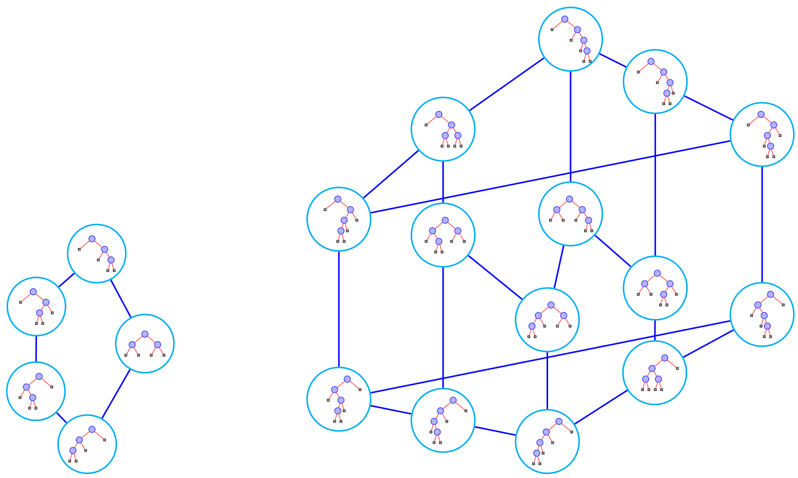
We write $T_0 \lessdot T_1$ if T_1 can be obtained by a right rotation from T_0. We call the relation $\lessdot$ the *partial Tamari relation*.

Remark 2.3 *Applying a right rotation to a tree does not change the infix order of its nodes.*

In the sequel, we only talk about right rotations, so we call these simply *rotations*. We are now in a position to give our definition of the Tamari relation:

Definition 2.4 *The* Tamari relation*, written $\preccurlyeq$, is the reflexive and transitive closure of the partial Tamari relation $\lessdot$.*

The Tamari relation is an order relation. For $n \geq 0$, the set $\mathcal{T}_n$ with the $\preccurlyeq$ order relation defines a lattice: the Tamari lattice. We denote by $\mathbb{T}_n = (\mathcal{T}_n, \preccurlyeq)$ the Tamari lattice of order n.



(a) $\mathbb{T}_3$ (b) $\mathbb{T}_4$

Fig. 3: The Tamari lattices $\mathbb{T}_3$ and $\mathbb{T}_4$.

3 Closure by interval of the set of balanced trees

3.1 Rotations and balance

Let us first consider the modifications of the imbalance values of the nodes of a tree $T_0 = (A \wedge B) \wedge C$ when a rotation at its root is applied. Let T_1 be the tree obtained by this rotation, y the root of T_0 and x the left child of y in T_0. Note first that the imbalance values of the nodes of the trees A, B and C are not modified by the rotation. Indeed, only the imbalance values of the nodes x and y are changed. Since T_0 is balanced, we have $\gamma_{T_0}(x) \in \{-1, 0, 1\}$ and $\gamma_{T_0}(y) \in \{-1, 0, 1\}$. Thus, the pair $(\gamma_{T_0}(x), \gamma_{T_0}(y))$ can take nine different values. Here follows the list of the imbalance values of the nodes x and y in the trees T_0 and T_1:

	(B1)	(U1)	(U2)	(B2)	(U3)	(U4)	(U5)	(U6)	(U7)
$(\gamma_{T_0}(x), \gamma_{T_0}(y))$	**(-1, -1)**	(-1, 0)	(-1, 1)	**(0, -1)**	(0, 0)	(0, 1)	(1, -1)	(1, 0)	(1, 1)
$(\gamma_{T_1}(x), \gamma_{T_1}(y))$	**(1, 1)**	(2, 2)	(3, 3)	**(1, 0)**	(2, 1)	(3, 2)	(2, 0)	(3, 1)	(4, 2)

Tab. 1: Imbalance values of the nodes x and y in T_0 and T_1.

Notice that only in (B1) and (B2) the tree T_1 is balanced. We have the following lemma:

Lemma 3.1 *Let T_0 and T_1 be two balanced trees such that $T_0 \lessdot T_1$. Then, the trees T_0 and T_1 have the same height.*

Proof: Since T_0 and T_1 are both balanced, the rotation modifies a subtree S_0 of T_0 such that the imbalance values of the root of S_0, namely y, and the left child of y, namely x, satisfy (B1) or (B2). Let S_1 be the tree obtained by the rotation of root y from S_0. Computing the height of the trees S_0 and S_1, we have $\mathrm{ht}(S_0) = \mathrm{ht}(S_1)$. Thus, as a rotation modifies a tree locally, we have $\mathrm{ht}(T_0) = \mathrm{ht}(T_1)$. $\square$

A rotation transforming a tree T_0 into a tree T_1 is a *conservative balancing rotation* if both T_0 and T_1 are balanced. Considering y the root of this rotation and x the left child of y, we see, by the previous computations and Lemma 3.1, that T_0 and T_1 are both balanced if and only if T_0 is balanced and

$$(\gamma_{T_0}(x), \gamma_{T_0}(y)) \in \{(-1,-1),(0,-1)\}. \tag{3.1}$$

Similarly, a rotation is an *unbalancing rotation* if T_0 is balanced but T_1 not.

Lemma 3.2 *Let T_0 be a balanced tree and T_1 be an unbalanced tree such that $T_0 \lessdot T_1$. Then, there exists a node z in T_1 such that $\gamma_{T_1}(z) \geq 2$ and the left subtree and the right subtree of z are both balanced.*

Proof: Immediate, looking at (U1), (U2), (U3), (U4), (U5), (U6) and (U7). □

3.2 Admissible words

Definition 3.3 *A word $z \in \mathbb{N}^*$ is* admissible *if either $|z| \leq 1$ or we have $z_1 - 1 \leq z_2$, and the word obtained by applying the substitution*

$$z_1.z_2 \longrightarrow \begin{cases} \max\{z_1, z_2\} + 1 & \textit{if } z_1 - 1 \leq z_2 \leq z_1 + 1, \\ z_2 & \textit{otherwise} \end{cases} \tag{3.2}$$

to z is admissible. Let us denote by $\mathcal{A}$ the set of admissible words.

For example, we can check that the word $z = 00122$ is admissible. Indeed, applying the substitution (3.2), we have $00122 \to 1122 \to 222 \to 32 \to 4$ and at each step, the condition $z_1 - 1 \leq z_2$ holds. The word $z' = 1234488$ is also admissible: $1234488 \to 334488 \to 44488 \to 5488 \to 688 \to 88 \to 9$. The word $z'' = 3444$ is not admissible because we have $3444 \to 544 \to 64$ and since that $6 - 1 \nleq 4$, we have $z'' \notin \mathcal{A}$.

Remark 3.4 *If z is an admissible word, then, for all $1 \leq i \leq |z| - 1$ the inequality $z_i - 1 \leq z_{i+1}$ holds.*

Remark 3.5 *The prefixes and suffixes of an admissible word are still admissible.*

Remark 3.6 *If $z = u.v$ where $z, u, v \in \mathbb{N}^*$ are admissible words, after applying the substitution (3.2) to v to obtain the word v', the word $z' = u.v'$ is still admissible.*

Let the *potential* $\mathrm{P}(z)$ of an admissible word z be the outcome of the application of the substitution (3.2). In the previous examples, we have $\mathrm{P}(z) = 4$ and $\mathrm{P}(z') = 9$.

Let T be a tree, x be a node of T, $(x = x_1, x_2, \ldots, x_\ell)$ be the sequence of all ancestors of x whose right sons are not themselves ancestors of x, ordered from bottom to top and $(S_{x_i})_{1 \leq i \leq \ell}$ be the sequence of the right subtrees of the nodes x_i (see Figure 4). The word z on the alphabet $\mathbb{N}$ defined by $z_i = \mathrm{ht}(S_{x_i})$ is called the *characteristic word* of the node x in the tree T and denoted by $\mathrm{c}_T(x)$.

Lemma 3.7 *Let T be a balanced tree, x a node of T, and z the characteristic word of x. Then, z is admissible and $\mathrm{P}(z) \leq \mathrm{ht}(T)$.*

Proof: By structural induction on balanced trees. The lemma is obviously true for the trees of the set $\mathcal{B}_0 \cup \mathcal{B}_1$. Let L and R be two balanced trees such that $T = L \wedge R$ is balanced too and assume that the lemma is true for both L and R. Let x be a node of T. Distinguishing the cases where x is a node of L, a node of R, or the root of T, we have, by induction, the statement of the lemma. □

Lemma 3.8 *Let T be a tree and y a node of T such that $\mathrm{c}_T(y)$ is admissible and all subtrees of the sequence $(S_{y_i})_{1 \leq i \leq \ell}$ are balanced. Then, for all node x of T such that $y \rightsquigarrow_T x$, the word $\mathrm{c}_T(x)$ is admissible.*

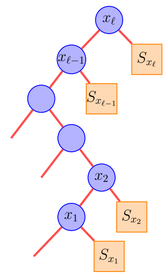


Fig. 4: The sequence $(S_{x_i})_{1\le i\le \ell}$ associated to the node $x = x_1$.

Proof: If x is an ancestor of y, the word $\mathrm{c}_T(x)$ is a suffix of $\mathrm{c}_T(y)$, thus we have, by Remark 3.5, $\mathrm{c}_T(x) \in \mathcal{A}$. Otherwise, let S be the subtree of T such that x is a node of S and the parent of S in T is an ancestor of y. We have $\mathrm{c}_T(y) = u.\,\mathrm{ht}(S).v$ where $u, v \in \mathcal{A}$. As $y \rightsquigarrow_T S$, we have $S \in \mathcal{B}$ and by Lemma 3.7, we have $\mathrm{c}_S(x) \in \mathcal{A}$ and $\mathrm{P}(\mathrm{c}_S(x)) \le \mathrm{ht}(S)$. Thus, thanks to Remark 3.5, $\mathrm{ht}(S).v \in \mathcal{A}$, so that $\mathrm{c}_T(x) = \mathrm{c}_S(x).v \in \mathcal{A}$. $\square$

3.3 The main result

Theorem 3.9 *Let T_0 and T_1 be two balanced trees such that $T_0 \preccurlyeq T_1$. Then, the interval $[T_0, T_1]$ only contains balanced trees. In other words, all successors of a tree obtained doing an unbalancing rotation into a balanced tree are unbalanced.*

Proof: To prove the theorem, we shall show that for all balanced tree T_0 and an unbalanced tree T_1 such that $T_0 \lessdot T_1$, all trees T_2 such that $T_1 \preccurlyeq T_2$ are unbalanced. Indeed, T_1 has a property guaranteeing it is unbalanced that can be kept for all its successors.

Let $\mathrm{Imb}_T(x)$ be the property: the node x of T and the node y which is the leftmost node of the left subtree of x satisfy: (see Figure 5):

(1) $\gamma_T(x) \ge 2$;

(2) the left subtree of x is balanced;

(3) all the subtrees S such that $y \rightsquigarrow_T S$ are balanced;

(4) $\mathrm{c}_T(y) \in \mathcal{A}$.

Point (2) guarantees that each tree having the previous property is unbalanced.

First, let us show that there exists a node x such that $\mathrm{Imb}_{T_1}(x)$ is true. The tree T_1 is obtained by an unbalancing rotation from T_0. By Lemma 3.2, there exists a node x in T_1 satisfying points (1) and (2). As the left and right subtrees of x are balanced and as all the trees on the right compared to x are balanced in T_0, they remain balanced in T_1, so that point (3) checks out. To establish (4), denoting by y the leftmost node of the left subtree of x in T_1, we have, by Remark 3.6 and Lemmas 3.7 and 3.8, $\mathrm{c}_{T_1}(y) \in \mathcal{A}$.

Now, let us show that given a tree T_1 such that $\mathrm{Imb}_{T_1}(x)$ is satisfied for a node x of T_1, for all tree T_2 such that $T_1 \lessdot T_2$, there exists a node x' of T_2 such that $\mathrm{Imb}_{T_2}(x')$ is satisfied. Let y be the leftmost node of the left subtree of x in T_1 and r be the root of the rotation that transforms T_1 into T_2. We will treat all cases depending on the position of r compared to y.

If the node r belongs to a subtree of T_1 which is on the left compared to y, the rotation does not modify any of the subtrees on the right compared to y. Thus we have $\mathrm{Imb}_{T_2}(x)$.

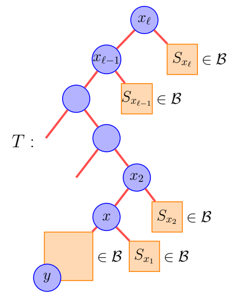


Fig. 5: The imbalance property $\mathrm{Imb}_T(x)$. The node y is the leftmost node of the left subtree of the node x.

If the subtree S_1 of root r satisfies $y \rightsquigarrow_{T_1} S_1$, let S_2 be the subtree of T_2 obtained by the rotation of S_1 which transforms T_1 into T_2. If S_2 is balanced, by Lemma 3.1, $\mathrm{ht}(S_1) = \mathrm{ht}(S_2)$ and we have $\mathrm{Imb}_{T_2}(x)$. If S_2 is not balanced, by the study of the initial case, we have $\mathrm{Imb}_{S_2}(x')$ for a node x' of S_2. Besides, by Remark 3.6 and Lemma 3.8, denoting by y' the leftmost node of the left subtree of x' in T_2, we have $\mathrm{c}_{T_2}(y') \in \mathcal{A}$ and thus, $\mathrm{Imb}_{T_2}(x')$.

If the node r is an ancestor of y and the left child of r is still an ancestor of y, let B be the right subtree of r and A the right subtree of the left child of r in T_1. The rotation replaces the trees A and B by the tree $A \wedge B$. As $\mathrm{c}_{T_1}(y) \in \mathcal{A}$, we have, by Remark 3.4, $\mathrm{ht}(A) - 1 \leq \mathrm{ht}(B)$. Thus, if $A \wedge B$ is balanced, we have $\mathrm{Imb}_{T_2}(x)$. Indeed, points (1), (2) and (3) are clearly satisfied and, by Remark 3.6, we have (4). If $A \wedge B$ is unbalanced, calling x' the root of this tree in T_2, we have $\gamma_{T_2}(x') \geq 2$, and, calling y' the leftmost node of A, we have, by Lemma 3.8, $\mathrm{c}_{T_2}(y') \in \mathcal{A}$. Thus we have $\mathrm{Imb}_{T_2}(x')$.

If the node r is an ancestor of y and the right child of r is still an ancestor of y, the rotation does not modify any of the subtrees on the right compared to y. Thus, we have $\mathrm{Imb}_{T_2}(x)$. □

4 Tree patterns and synchronous grammars

Word patterns are usually used to describe languages by considering the set of words avoiding them. We use the same idea to describe sets of trees. We show first that we can describe two interesting subsets of the set of balanced trees only by two-nodes patterns.

Next, we follow the methods of [7, 4] to characterize, in our setting, a way to obtain a functional equation admitting as fixed point the generating series enumerating balanced trees. In this purpose, we introduce *synchronous grammars*, allowing to generate trees iteratively. This method gives us a way to enumerate trees avoiding a set of tree patterns because, as we shall see, functional equations of generating series can be extracted from synchronous grammars.

4.1 Tree patterns

Definition 4.1 *A* tree pattern *is a nonempty non complete rooted planar binary tree with labels in* $\mathbb{Z}$*.*

Let T be a tree and T_γ be the labeled tree of shape T where each node of T_γ is labeled by its imbalance value. The tree T admits an *occurrence* of a tree pattern p if a connected component of T_γ has the same shape and same labels as p.

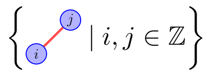
Now, given a set P of tree patterns, we can define the set composed of the trees that do not admit any occurrence of the elements of P. For example, the set

$$\{ i \mid i \notin \{-1, 0, 1\}\} \tag{4.1}$$

describes the set of balanced trees; the set

$$\{ i \mid i \neq 0\} \tag{4.2}$$

describes the set of perfect trees and



(4.3)

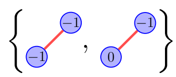
describes the set of right comb trees.

4.2 Two particular subsets of balanced trees

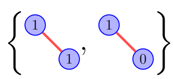
Let us describe a subset of the balanced trees and its counterpart such that its elements are, roughly speaking, at the end of the balanced trees subset in the Tamari lattice:

Definition 4.2 *A balanced tree T_0 (resp. T_1) is* maximal *(resp.* minimal*) if for all balanced tree T_1 (resp. T_0) such that $T_0 \lessdot T_1$ we have T_1 (resp. T_0) unbalanced.*

Proposition 4.3 *A balanced tree T is maximal if and only if it avoids the set of tree patterns*

$P_{\max} :=$  (4.4)

Similarly, a balanced tree T is minimal if and only if it avoids the set of tree patterns

$P_{\min} :=$  (4.5)

Proof: Assume that T is maximal. For all tree T_1 such that $T \lessdot T_1$ we have T_1 unbalanced. Thus, it is impossible to do a conservative balancing rotation from T and it avoids the set $P_{\max}$.

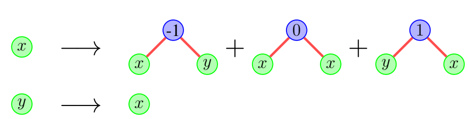
Assume that T avoids the two tree patterns of $P_{\max}$, then, for every tree T_1 such that $T \lessdot T_1$, the tree T_1 is unbalanced because we can do only unbalancing rotations in T. Thus, the tree T is maximal.

The proof of the second part of the proposition is done in an analogous way. □

4.3 Synchronous grammars and enumeration of balanced trees

Let us first describe a way to obtain the functional equation admitting as fixed point the generating series which enumerates balanced trees [7, 4].

The idea is to generate trees by allowing them to grow from the root to the leaves step by step. For that, we generate *bud trees*, that are non complete rooted planar binary trees with the particularity that the set of external nodes (the nodes without descendant) are *buds*. A bud tree grows by *simultaneously* substituting all of its buds by new bud trees. Trees are finally obtained replacing buds by leaves. The rules of substitution allowing to generate bud trees form a *synchronous grammar*. The link between tree patterns and synchronous grammars is that synchronous grammars generate trees controlling the imbalance value of the nodes. The rules generating balanced trees are



(4.6)

(4.7)

The role of the bud x is to generate a node which has -1, 0 or 1 as imbalance value, the only values that a balanced tree can have. The role of the bud y is to delay the growth of the bud tree to enable the creation of the imbalance values -1 and 1. We have the following theorem:

Theorem 4.4 *Let B be a bud tree generated from the bud* ⓧ *by the previous synchronous grammar. If B does not contain any bud* ⓨ*, replacing all buds* ⓧ *by leaves, we obtain a tree T where each node z of T is labeled by $\gamma_T(z)$. In this way, the previous synchronous grammar generates exactly the set of balanced trees.*

Figure 6 shows an example of such a generation.

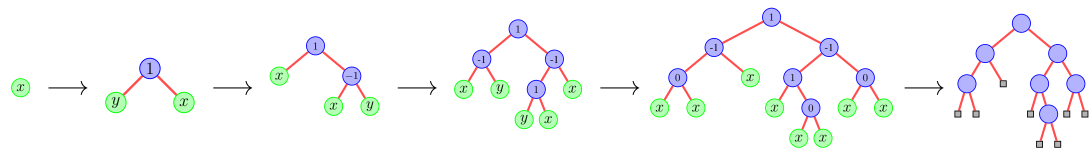


Fig. 6: Generation of a balanced tree.

The main purpose of synchronous grammars is to obtain a way to enumerate the trees generated. We can translate the set of rules to obtain a functional equation of the generating series enumerating them. For balanced trees, we have [7, 4, 9]:

Theorem 4.5 *The generating series enumerating balanced trees according to the number of leaves of trees is $G_{\text{bal}}(x) := A(x, 0)$ where*

$$A(x, y) := x + A(x^2 + 2xy, x). \tag{4.8}$$

The resolution, or, in other words, the coefficient extraction for this kind of functional equation, is made by iteration. We proceed by computing the sequence of polynomials $(A_i)_{i\geq 0}$ defined by:

$$A_i(x, y) = \begin{cases} x & \text{if } i = 0, \\ x + A_{i-1}(x^2 + 2xy, x) & \text{otherwise.} \end{cases} \tag{4.9}$$

The first iterations give

$$A_0 = x, \tag{4.10}$$
$$A_1 = x + 2xy + x^2, \tag{4.11}$$
$$A_2 = x + 2xy + x^2 + 4x^2y + 2x^3 + 4x^2y^2 + 4x^3y + x^4. \tag{4.12}$$

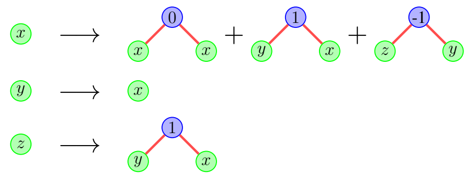
The fixed point of the sequence $(A_i)_{i\geq 0}$, after substituting 0 to the parameter y in order to ignore bud trees with some buds ⓨ, is the generating series of balanced trees counted according to the number of leaves.

We can refine this idea to enumerate maximal balanced trees:

Proposition 4.6 *The generating series enumerating maximal balanced trees according to the number of leaves of the trees is $G_{\max}(x) := A(x, 0, 0)$ where*

$$A(x, y, z) := x + A(x^2 + xy + yz, x, xy). \tag{4.13}$$

Proof: To obtain this functional equation, let us use the following synchronous grammar which generates maximal balanced trees:



(4.14)
(4.15)
(4.16)

This grammar must generate only maximal balanced trees. By Proposition 4.3, the generated trees must avoid the two tree patterns of P_{max}. To do that, we have to control the growth of the bud x when it generates a tree S such that its root has an imbalance value of -1. Indeed, if the root of the left subtree of S grows with an imbalance value of -1 or 0, one of the two tree patterns is not avoided. The idea is to force the imbalance value of the root of left subtree of S to be 1, role played by the bud z. $\square$

The solution of this functional equation give us the following first values for the number of maximal trees in the Tamari lattice: 1, 1, 1, 1, 2, 2, 2, 4, 6, 9, 11, 13, 22, 38, 60, 89, 128, 183, 256, 353, 512, 805, 1336, 2221, 3594, 5665, 8774, 13433, 20359.

5 The shape of the balanced tree intervals

5.1 Isomorphism between balanced tree intervals and hypercubes

A hypercube of dimension k can be seen as a poset whose elements are subsets of a set $\{e_1, \dots, e_k\}$ ordered by the relation of inclusion. Let us denote by $\mathbb{H}_k$ the hypercube poset of dimension k.

We have the following characterization of the shape of balanced tree intervals:

Theorem 5.1 *Let T_0 and T_1 be two balanced trees such that $T_0 \preccurlyeq T_1$. Then there exists $k \geq 0$ such that the posets $([T_0, T_1], \preccurlyeq)$ and $\mathbb{H}_k$ are isomorphic.*

Proof: First, note by Theorem 3.9, that $I = [T_0, T_1] \subseteq \mathcal{B}$. Thus, every covering relation of the interval I is a conservative balancing rotation.

Then, note that the rotations needed to transform T_0 into T_1 are disjoint in the sense that if y is a node of $T_2 \in I$ and x its left child, if we apply a conservative balancing rotation of root y in T_2 to obtain $T_3 \in I$, all the rotations in the successors of T_3 of root y and of root x are unbalancing rotations. Indeed, by Lemma 3.1, each conservative balancing rotation modifies only the imbalance values of the root of the rotation and its left child, and, according to the values obtained, these two nodes cannot thereafter be roots of conservative balancing rotations.

Besides, by the nature of the conservative balancing rotations and by Theorem 3.9, we can see that all the ways to transform T_0 into T_1 solicit the same rotations, possibly in a different order.

Now, we can associate to a tree $T \in I$ a subset of $\mathbb{N}$ containing the positions in the infix order of the nodes y such that, to obtain T from T_0, we have done, among other, a rotation of root y. The interval I is isomorphic to the poset $\mathbb{H}_k$ where k is the number of rotations needed to transform T_0 into T_1. $\square$

5.2 Enumeration of balanced tree intervals

Let us make use again of the synchronous grammars:

Proposition 5.2 *The generating series enumerating balanced tree intervals in the Tamari lattice according to the number of leaves of the trees is $G_{\text{inter}}(x) := A(x, 0, 0)$ where*

$$A(x, y, z) := x + A(x^2 + 2xy + z, x, x^3 + x^2 y). \tag{5.1}$$

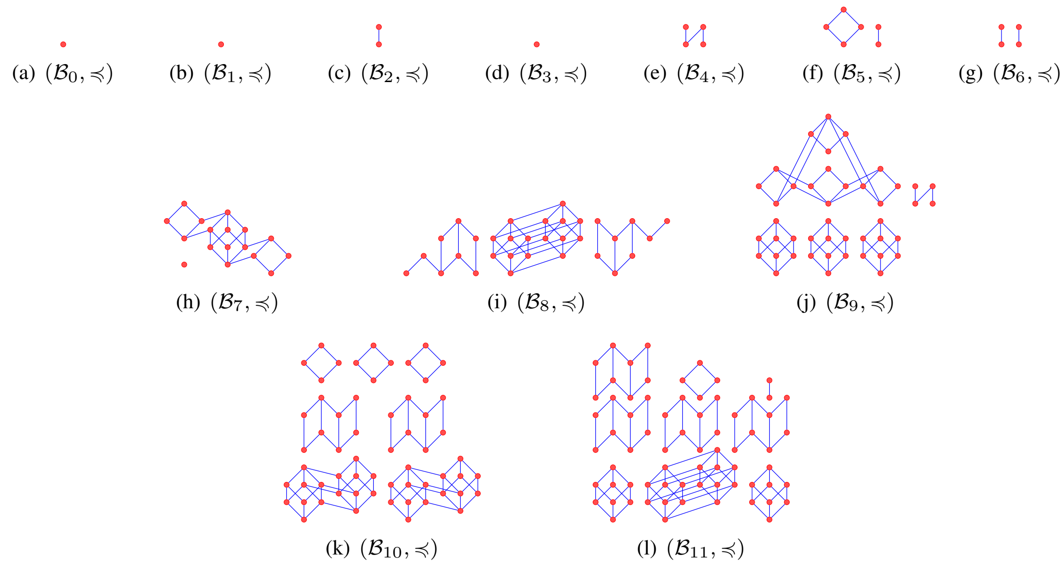
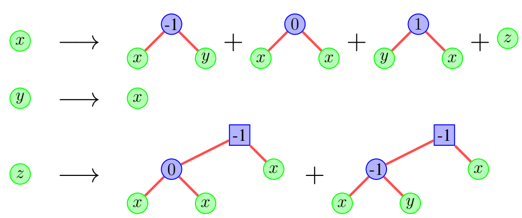


Fig. 7: Hasse diagrams of the first $(\mathcal{B}_n, \preccurlyeq)$ posets.

Proof: Let $I = [T_0, T_1]$ be a balanced tree interval. This interval can be encoded by the tree T_0 in which we mark the nodes which are roots of the conservative balancing rotations needed to transform T_0 into T_1. If a node y of T_0 is marked, then its left child cannot be marked too because the rotations of the interval I are disjoint (see the proof of Theorem 5.1). To generate these objects, we use the following synchronous grammar that generates marked trees (the marked nodes are represented by a rectangle instead of a circle):



(5.2)

(5.3)

(5.4)

□

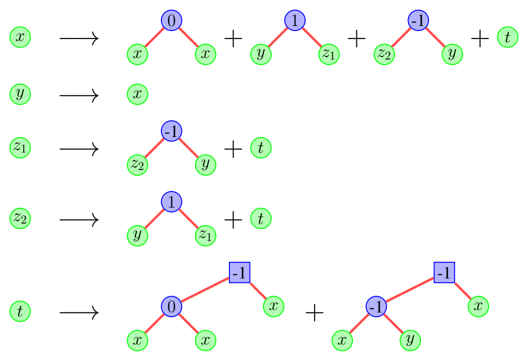
The solution of this functional equation gives us the following first values for the number of balanced tree intervals in the Tamari lattice: 1, 1, 3, 1, 7, 12, 6, 52, 119, 137, 195, 231, 1019, 3503, 6593, 12616, 26178, 43500, 64157, 94688, 232560, 817757, 2233757, 5179734.

The interval $[T_0, T_1]$ is a *maximal balanced tree interval* if T_0 (resp. T_1) is a minimal (resp. maximal) balanced tree.

Proposition 5.3 *The generating series enumerating maximal balanced tree intervals in the Tamari lattice according to the number of leaves of the trees is* $G_{\text{intermax}}(x) := A(x, 0, 0, 0)$ *where*

$$A(x, y, z, t) := x + A(x^2 + 2yz + t, x, yz + t, x^3 + x^2 y). \tag{5.5}$$

Proof: Let $I = [T_0, T_1]$ be a maximal balanced tree interval. This interval can be encoded by the minimal tree T_0 in which we mark the nodes which are roots of the conservative balancing rotations needed to transform T_0 into T_1. Since T_1 is a maximal balanced tree, by Proposition 4.3, it avoids the tree patterns of $P_{\max}$, thus, the object which encodes I must not have a node which is root of a conservative balancing rotation not marked if its parent or its left child is not marked. To generate these objects, we use the following synchronous grammar:



(5.6)
(5.7)
(5.8)
(5.9)
(5.10)

Note that the buds z_1 and z_2 play the same role so that the functional equation is simplified. □

The solution of this functional equation gives us the following first values for the number of maximal balanced tree intervals in the Tamari lattice: 1, 1, 1, 1, 3, 2, 2, 6, 9, 15, 15, 17, 41, 77, 125, 178, 252, 376, 531, 740, 1192, 2179, 4273, 7738, 13012, 20776, 32389, 49841, 75457, 113011.

References

[1] Sage mathematics software, version 4.2, 2009. `http://www.sagemath.org/`.

[2] G.M. Adelson-Velsky and E. M. Landis. An algorithm for the organization of information. *Soviet Mathematics Doklady, 3:1259-1263*, 1962.

[3] A. Aho and J. Ullman. *Foundations of Computer Science*. W. H. Freeman, 1994.

[4] F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial Species and Tree-like Structures*. Cambridge University Press, 1994.

[5] O. Bernardi and N. Bonichon. Catalan's intervals and realizers of triangulations. *FPSAC*, 2007.

[6] T.H. Cormen, C. E. Leiserson, R.L. Rivest, and C. Stein. *Introduction to algorithms*. McGraw-Hill, 2003.

[7] D. Knuth. *The Art of Computer Programming. Volume 3. Sorting and searching*. Addison Wesley Longman, 1998.

[8] D. Knuth. *The Art of Computer Programming. Volume 4, Fascicle 4. Generating all trees - History of combinatorial generation*. Addison Wesley Longman, 2004.

[9] N. J. A. Sloane. The on-line encyclopedia of integer sequences. `http://www.research.att.com/~njas/sequences/`.

[10] R. P. Stanley. *Enumerative Combinatorics II*. Cambridge University Press, 1990.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 608–619

Enumeration of inscribed polyominos

Alain Goupil[1†], Hugo Cloutier[1] and Fathallah Nouboud[1]

[1]*Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières, 3351 boul des Forges, c.p. 500, Trois-Rivières (QC) Canada*
alain.goupil@uqtr.ca, hugo854@yahoo.ca, Fathallah.Nouboud@uqtr.ca

Abstract. We introduce a new family of polyominos that are inscribed in a rectangle of given size for which we establish a number of exact formulas and generating functions. In particular, we study polyominos inscribed in a rectangle with minimum area and minimum area plus one. These results are then used for the enumeration of lattice trees inscribed in a rectangle with minimum area plus one.

Résumé. Nous introduisons une nouvelle famille de polyominos inscrits dans un rectangle de format donné pour lesquels des formules exactes et des séries génératrices sont présentées. Nousétudions en particulier les polyominos inscrits d'aire minimale et ceux d'aire minimale plus un. Ces résultats sont ensuite utilisés pour l'énumération de polyominos arbres inscrits dans un rectangle d'aire minimum plus un.

Keywords: inscribed polyomino, enumeration, rectangle, generating function, lattice tree, minimal area.

1 Introduction

S. Golomb introduced polyominos in 1952 [6]. Various families of polyomominos have been defined and investigated since then (see [1], [2], [4], [5] and ref. therein). Algorithms have also been developed for their enumeration (see [7]). But the problem of their enumeration in the general case remains unsolved. In this work, we have developed formulas that, to our knowledge, are counting polyominos of a family not described in the existing literature so we could not connect our work with it.

A polyomino, sometimes called an animal, is a set of unit square cells in the discrete plane $\mathbb{N} \times \mathbb{N}$ connected by their edges up to translation. We are interested in the number $p(n)$ of polyominos with n cells where n is called the *area* of these polyominos. A polyomino is inscribed in a rectangle $b \times k$ when it is included in the rectangle and each of the four edges of the rectangle is touched by a cell of the polyomino. The minimum number of cells in a polyomino inscribed in a $b \times k$ rectangle has $b + k - 1$ cells and we will denote respectively by $p_{min}(b, k)$ and $p_{min+1}(b, k)$ the number of polyominos that are inscribed in a $b \times k$ rectangle and have minimum area and minimum area plus one. A lattice tree is a polyomino that contains no cycle and we will also be interested with lattice trees inscribed in a rectangle. The main results of this work are the following.

[†]Work supported in part by NSERC

Theorem 1 *For integers $b \geq 2, k \geq 2$, the number $p_{min}(b,k)$ of polyominos inscribed in a rectangle $b \times k$ with minimal area $n = b + k - 1$ is given by the fomula*

$$p_{min}(b,k) = 2k + 2b - 3bk - 8 + 8\binom{k+b-2}{b-1}$$

Corollary 1 *For all integers $n \geq 1$ the number $p_{min}(n)$ of polyominos with n cells inscribed in a rectangle of perimeter $2(n+1)$ is given by the fomula*

$$p_{min}(n) = 2^{n+2} - \frac{1}{2}(n^3 - n^2 + 10n + 4)$$

The polyominos in the previous corollary can also be seen as animals occupying a rectangular region of maximal perimeter with respect to their area.

Theorem 2 *The two variables generating function for the number $p_{min}(b,k)$ of polyominos of minimal area inscribed in a rectangle $b \times k$ has the following rational form :*

$$\begin{aligned}\sum_{b,k\geq 1} f_{min}(b,k)x^b y^k \quad &= \quad 2\left(1 + \frac{xy}{(1-x)(1-y)}\right)^2 \frac{xy}{(1-x-y)} - \\ &\left(\frac{xy}{(1-x)^2(1-y)^2} - \frac{xy^2}{(1-y)^2} - \frac{x^2y}{(1-x)^2}\right)\end{aligned}$$

Theorem 3 *For all integers $b, k \geq 1$, the number $p_{min+1}(b,k)$ of polyominos inscribed in a rectangle $b \times k$ that have minimum area plus one is*

$$p_{min+1}(b,k) = \begin{cases} 0 & \text{if } b = 1 \text{ or } k = 1 \\ 1 & \text{if } b = k = 2 \\ 4b^2 - 16b + 18 & \text{if } k = 2 \text{ and } b > 2 \\ 8(b+k-22)\binom{b+k-4}{b-2} + \frac{8(2k^2+2kb+b-13k+13)}{(k-2)}\binom{b+k-4}{b-1} & \\ +\frac{8(2b^2+2kb+k-13b+13)}{(b-2)}\binom{b+k-4}{k-1} + 48\binom{b+k-2}{b-1} & \\ -\frac{4}{3}(b^3+k^3) - 12(b^2k + bk^2) + 16(b^2+k^2) & \\ +72bk - \frac{266}{3}(b+k) + 120 & \text{if } b \geq 3 \text{ and } k \geq 3 \end{cases}$$

Corollary 2 *For all integers $n \geq 5$, the number $p_{min+1}(n)$ of polyominos with n cells inscribed in a rectangle of perimeter $2n$ is*

$$p_{min+1}(n) = 2^{n-1}(5n-6) - \frac{2}{3}(4n^4 - 44n^3 + 215n^2 - 451n + 318)$$

A consequence of theorem 3 and corollary 2 is to obtain exact formulas for corresponding sets of lattice trees inscribed in a rectangle.

It is clear that any general polyomino is always inscribed in a rectangle so that the set $Po(n)$ of polyominos with area n can be partitionned into classes given by the the dimensions $b \times k$ of the circonscribed rectangles. Our approach in counting inscribed polyominos thus constitute a fair strategy to attack the well known problem of counting the total number $po(n) = card(Po(n))$ of polyominos of area n.

Notations. As a general rule we will use capital letters for sets and corresponding small letters for their cardinalities. We will introduce specific notations as they are needed.

2 Proofs of the formulas

Proof of theorem 1. We begin with two geometric observations on inscribed polyominos with minimal area. 1- All inscribed polyominos with minimal area are oriented along one of the two diagonals of the $b \times k$ rectangle noting that polyominos with a cross shape (figure (1 d)) are the only polyominos with minimal area that can be seen as oriented along both diagonals. 2- Minimal area polyominos all have a structure in three parts: one hook, possibly reduced to a unique cell, on each end of the diagonal connected on their corner by a stair polyomino in the direction of the diagonal as shown in figure 1 c). A stair polyomino (figure (1 b)) along one diagonal, going say from north-west to south-east $\backslash$ is a path allowed only two directions for adjacent cells: east $\rightarrow$ and south $\downarrow$.

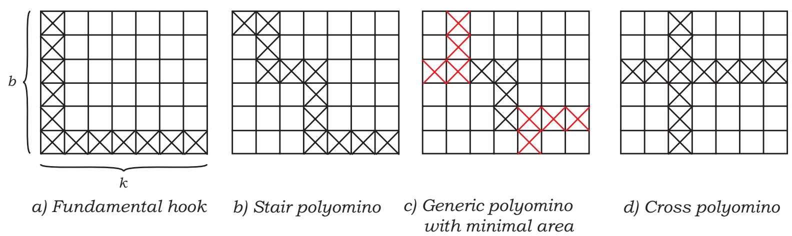


Fig. 1: Inscribed minimal polyominos

The geometric triple-structure of polyominos with minimal area appearing in figure 1 c) can also be given a biological interpretation. Animals with n cells that need to touch the edges of a rectangle of maximal perimeter must have this geometric triple-structure and shape.

We have produced two proofs of theorem 1. Each proof consists in a case study of the set of polyominos of minimal area. The first proof uses the triple-structure hook-stair-hook of minimal polyominos and the second proof is a dynamic construction of the polyominos beginning with the fundamental hook (figure 1 a)) and moving the square cells horizontally or vertically to form a new inscribed polyomino. We present here only the first proof.

Let $P_{min,\backslash}(b,k)$ be the set of polyominos of minimal area inscribed in a rectangle $b \times k$ along the diagonal from north-west to south-east. Denote by $p_{min,\backslash}(b,k)$ the cardinality of $P_{min,\backslash}(b,k)$. Let $P_{min,/}(b,k)$ and $p_{min,/}(b,k)$ be similarly defined for the other diagonal. Since there is clearly a bijection between the two sets $P_{min,\backslash}(b,k)$ and $P_{min,/}(b,k)$, we need only consider one of the diagonals of the rectangle. The set $P_+(b,k)$ of Cross polyominos satisfies $P_+(b,k) = P_{min,\backslash}(b,k) \cap P_{min,/}(b,k)$, so that we have

$$p_{min}(b,k) = 2p_{min,\backslash}(b,k) - p_+(b,k) \tag{1}$$

Let $P_{min,(i,j)}(b,k)$ be the set of polyominos in $P_{min,\backslash}(b,k)$ having the corner cell of their upper left hook in position (i,j) in matrix notation. Thus $p_{min,(1,1)}(b,k)$ is the number of polyominos in $P_{min,\backslash}(b,k)$ that have a cell in the upper left corner of the rectangle $b \times k$. Let us count these polyominos. First observe that the cell in position $(1,1)$ must be in one of three situations: a) it is connected to a cell on its right with no cell below. b) it is connected to a cell below with no cell on the right. c) It has a cell on the right and a cell below. In the first two cases, if we remove the cell $(1,1)$ we obtain polyominos with minimal areas inscribed in a smaller rectangle. There is only one polyomino with minimal area in the third case. Thus we obtain the recurrence

$$p_{min,(1,1)}(b,k) = p_{min,(1,1)}(b,k-1) + p_{min,(1,1)}(b-1,k) + 1 \quad \forall\, b,h \geq 1 \tag{2}$$

with the initial conditions $p_{min,(1,1)}(b,0) = p_{min,(1,1)}(0,k) = 0$. There is also an exact expression for $p_{min,(1,1)}(b,k)$. The key observation is the well known fact that the number of stair polyominos inscribed in a rectangle with cells in each corner of a diagonal is given by a binomial coefficient. Let $P_{stair}(b,k)$ be the set of stair polyominos in $P_{min,\backslash}(b,k)$ with end cells in each end of the main diagonal of the rectangle $b \times k$. Then

$$p_{stair}(b,k) = \binom{b+k-2}{b-1} \tag{3}$$

Polyominos in $P_{min,(1,1)}(b,k)$ are in bijective correspondance with polyominos in $(\cup_{i<b,j<k} P_{stair}(i,j)) \cup P_{stair}(b.k)$ so that using basic binomial identities we obtain

$$p_{min,(1,1)}(b,k) = \sum_{i=1}^{b-1}\sum_{j=1}^{k-1}\binom{i+j-2}{i-1} + \binom{b+k-2}{b-1} = 2\binom{b+k-2}{b-1} - 1 \tag{4}$$

Moreover it is also immediate that

$$p_{min,(i,j)}(b,k) = p_{min,(1,1)}(b-i+1, k-j+1) \tag{5}$$

so that by equation (4) we have

$$p_{min,(i,j)}(b,k) = 2\binom{b+k-i-j}{b-i} - 1. \tag{6}$$

Since $p_+(b,k) = bk$ and

$$p_{min,\backslash}(b,k) = p_{min,(1,1)}(b,k) + \sum_{i=2}^{b}\sum_{j=2}^{k} p_{min,(i,j)}(b,k), \tag{7}$$

using (1) we obtain

$$p_{min}(b,k) = 2\left(p_{min,(1,1)}(b,k) + \sum_{i=2}^{b}\sum_{j=2}^{k} p_{min,(i,j)}(b,k)\right) - bk, \tag{8}$$

Now using (8) and (6) we obtain an exact expression for $p_{min}(b,k)$:

$$\begin{aligned} p_{min}(b,k) &= 2\left(2\binom{b+k-2}{b-1} - 1 + \sum_{i=2}^{b}\sum_{j=2}^{k}\left(2\binom{b+k-i-j}{b-i} - 1\right)\right) - bk \\ &= 8\binom{b+k-2}{b-1} - 6 - 2(b-1)(k-1) - bk \quad \forall b,k \geq 1. \end{aligned} \tag{9}$$

which proves theorem 1. $\diamond$

Proof of corollary 1. Polyominos with minimal area inscribed in a rectangle can also be seen as polyominos that are maximally stretched. These polyominos occupy a rectangle of maximal perimeter $2n+2$ when their area is n. If we sum maximally stretched polyominos over all rectangles of perimeter $2n+2$, we obtain the number $p_{min}(n)$ of maximally stretched polyominos:

$$\begin{aligned} p_{min}(n) &= \sum_{b=1}^{n} p_{min}(b,k) = 2 + \sum_{b=2}^{n-1} \left(8\binom{b+k-2}{b-1} - 6 - 2(b-1)(k-1) - bk \right) \\ &= 2^{n+2} - \frac{1}{2}(n^3 - n^2 + 10n + 4) \end{aligned} \tag{10}$$

which proves corollary 1. Observe that we have computed separately the cases $b = 1$ and $b = n$ in equation (10) $\diamond$

n	1	2	3	4	5	6	7	8	9	10
$p_{min}(n)$	1	2	6	18	51	134	328	758	1677	3594

Tab. 1: Numbers $p_{min}(n)$ of maximally stretched polyominos of area n

Proof of theorem 2. We construct the rational form of the generating function $\sum_{b,k\geq 1} f_{min,\backslash}(b,k)x^b y^k$ from its triple-structure $hook - stair - hook$ described before and the multiplication principle.
Since there is at most one hook in the upper left corner of the rectangle having its corner in position (i,j) and because we choose not to count the corner cell, the generating function for hooks with corner in position (i,j) is

$$1 + \sum_{i,j\geq 2} x^{i-1}y^{j-1} = 1 + \frac{xy}{(1-x)(1-y)} \tag{11}$$

Recall that the number of stairs from the upper left corner of a rectangle to the cell (i,j) is $\binom{i+j-2}{i-1}$ so that the generating function for stair polyominos is

$$\begin{aligned} \sum_{i,j\geq 1} \binom{i+j-2}{i-1} x^i y^j &= xy \sum_{i,j\geq 1} \binom{i+j-2}{i-1} x^{i-1}y^{j-1} = xy \sum_{k\geq 0} (x+y)^k \\ &= \frac{xy}{1-x-y} \end{aligned} \tag{12}$$

Now applying the multiplication principle and equations (11) and (12) we obtain

$$\sum_{b,k\geq 1} f_{min,\backslash}(b,k)x^b y^k = \left(1 + \frac{xy}{(1-x)(1-y)}\right)^2 \frac{xy}{(1-x-y)} \tag{13}$$

Finally recalling equation 1 we deduce theorem 2 when we agree that the generating function for crosses is

$$\sum_{b,k\geq 1} f_+(b,k)x^b y^k = \sum_{k\geq 1} xy^k + \sum_{b\geq 2} x^b y + \sum_{b,k\geq 2} bkx^b y^k = \frac{xy}{(1-x)^2(1-y)^2} - \frac{xy^2}{(1-y)^2} - \frac{x^2 y}{(1-x)^2}$$

Proof of theorem 3. The proof of theorem 3 is also a case study, with more cases though than in the proof of theorem 1. We will need $2 \times t$ *bench polyominos* that we define as polyominos of area $t + 2$ inscribed in a $2 \times t$ rectangle, $t \geq 2$, with one full row, or column, of cells plus the two cells at each end of the other row as shown in figure 2 a). We will use the following triple-structure description of polyominos

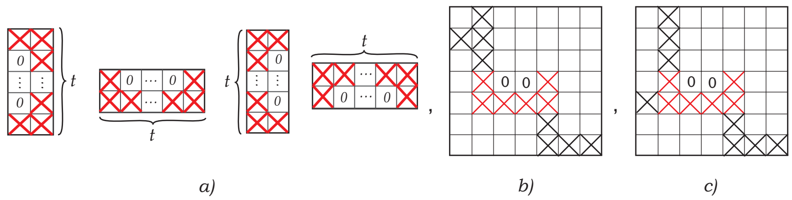


Fig. 2: Bench polyominos

of area $min + 1$ inscribed in a rectangle.

Facts. a) A polyomino of area $min + 1$ inscribed in a rectangle contains exactly one bench polyomino in one of the four possibles positions of figure 2 a). b) Moreover there is exactly two ways to complete a fixed bench into a polyomino of area $min+1$ along one diagonal of the $b \times k$ rectangle. First, a polyomino of minimal area is attached to a corner of the bench (figure 2 b)) and if it is a hook, it may have its corner cell on any cell of the $2 \times t$ circonscribed rectangle (figure 2 c)), provided the connectivity condition is satisfied. Second, a polyomino of minimal area is attached on the opposite corner and if it is a hook, it may have its corner on any cell of the $2 \times t$ circonscribed rectangle up to connectivity.
c) Starting on the north-west corner and moving clockwise, let c_1, c_2, c_3 and c_4 be the four corner cells of a bench polyomino B included in a $b \times k$ rectangle. Let f_1, f_2, f_3, f_4 be the number of polyominos inscribed in the respective rectangles determined by the diagonals from the northwest corner of the $b \times k$ rectangle to the northwest corner $c_1 \in B$ and so on for the other three rectangles as in figure 3.

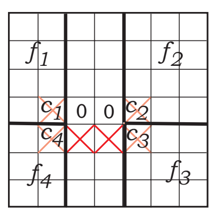


Fig. 3: A polyomino of area $min + 1$ constructed from a bench polyomino

The number $p(B)$ of polyominos of area $min + 1$ inscribed in a $b \times k$ rectangle and containing the bench polyomino B is given by the formula

$$p(B) = f_1 \cdot f_3 + f_2 \cdot f_4 - 8t \tag{14}$$

Case 1. The bench is in a corner. Let us start by considering the case where a bench is in one corner of the rectangle.

Proposition 1 *For integers $t \geq 2$ let $p_{t,1}(b,k)$, resp. $p_{t,2}(b,k)$, be the number of polyominos in $P_{min+1}(b,k)$ containing a $2 \times t$ bench in the northwest corner of the rectangle with the seating part (figure 4 a), resp. the leg part (figure 4 b), upward. Then we have*

$$p_{t,1}(b,k) = 2\binom{b+k-t-2}{b-2} + 2 \tag{15}$$

$$p_{t,2}(b,k) = 2\binom{b+k-t-2}{b-2} + 2(t-1) \tag{16}$$

Proof: We have to observe that once a bench is placed in a corner of the rectangle, we may complete it into a polyomino of area $min + 1$ either by adding a polyomino of minimal area inscribed in the sub-rectangle with corners given by the southeast corner of the bench and the southeast corner of the rectangle or by adding a hook as shown in figure 4. In the case where the legs of the bench are upwards, the corner of the hook, sometimes absent, is any of the $2t$ cells of the rectangle containing the bench. In the case where the legs of the bench are downwards, there are 4 possible hooks, one of which is already counted. Formulas (15) and (16) then follow from equation (4). □

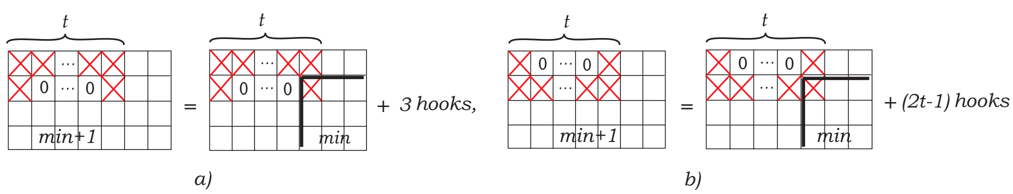


Fig. 4: Case 1. A bench in a corner

Corollary 3 *For integers $b, k \geq 3$, the number $g_1(b,k)$ of polyominos of area $min + 1$ inscribed in a $b \times k$ rectangle with a bench polyomino $2 \times t$ in any corner of the rectangle is given by the formula*

$$\begin{aligned} g_1(b,k) &= (4\sum_{t=3}^{k-1} p_{t,1}(b,k) + 4) + (4\sum_{t=3}^{k-1} p_{t,2}(b,k) + 2k) \qquad (17)\\ &\quad +(4\sum_{t=3}^{b-1} p_{t,1}(k,b) + 4) + (4\sum_{t=3}^{b-1} p_{t,2}(k,b) + 2b), \\ &= 16\left(\binom{b+k-4}{b-1} + \binom{b+k-4}{k-1}\right) + 2k(2k-1) + 2b(2b-1) - 72 \qquad (18) \end{aligned}$$

Proof: This is a consequence of proposition 1 and of a careful study of the particular cases involved. In each corner of the rectangle there are up to four benches to consider; the sums in formula (17) cannot be taken up to $t = k$ because there are less cases to consider. Also symmetry in b, k have been integrated to shorten the expressions. □

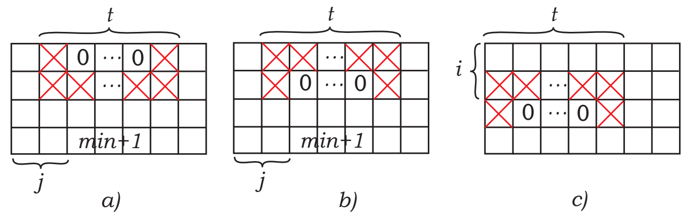


Fig. 5: The three dispositions of a horizontal bench along one side

Case 2. The bench touches one side of the rectangle and is not in a corner. There are three ways to put a horizontal bench on one side of a $b \times k$ rectangle as shown in figure 5.

Proposition 2 *Let $g_{2-horiz}(b,k)$ be the number of polyominos of area $min+1$ inscribed in a $b \times k$ rectangle with a horizontal bench polyomino of length $t \geq 3$ along one of the sides without being in a corner of the $b \times k$ rectangle. We have*

$$\begin{aligned}
g_{2-horiz}(b,k) &= 2\left[\sum_{t=3}^{k-2}\sum_{j=2}^{k-t} 2p_{t,1}(b,k-j+1) + 2p_{t,1}(b,j+t-1) - 8\right] + \\
&\quad 2\left[\sum_{t=3}^{k-2}\sum_{j=2}^{k-t} 2p_{t,2}(b,k-j+1) + 2p_{t,2}(b,j+t-1) - 4t\right] + \\
&\quad 4\left[\sum_{t=3}^{k-1}\sum_{i=2}^{b-2} tp_{t,1}(b-i+1,k) + 2p_{t,2}(i+1,k) - 4t\right] + 4k(b-3) \\
&= 8\left[-2k+6+2\binom{b+k-4}{b}\right] + 2\left[8\binom{b+k-4}{b} + \frac{2}{3}(k-3)(k^2-6k-4)\right] \\
&\quad +2\left[\frac{2(5b+k-7)}{(k-2)}\binom{b+k-4}{b-1} - 7bk + 2b + bk^2 - 4k^2 + 14k + 2\right] + 4k(b-3)
\end{aligned}$$

Proof: Omitted. □

Corollary 4 *Let $g_2(b,k)$ be the number of polyominos inscribed in a $b \times k$ rectangle and area $min+1$ containing a bench polyomino $2 \times t,\ t \geq 3$ touching one of the sides without being in a corner of the $b \times k$ rectangle. We have*

$$\begin{aligned}
g_2(b,k) &= g_{2-horiz}(b,k) + g_{2-horiz}(k,b) \\
&= 32\left(\binom{b+k-4}{b} + \binom{b+k-4}{k}\right) + \\
&\quad 8\left(\frac{(5k+b-7)}{(b-2)}\binom{b+k-4}{k-1} + \frac{(5b+k-7)}{(k-2)}\binom{b+k-4}{b-1}\right) + \\
&\quad \frac{4}{3}(b^3+k^3) - 28(k^2+b^2) - 48bk + \frac{164}{3}(b+k) + 4(bk^2+b^2k) + 144
\end{aligned}$$

Proof: The first equality partitions polyominos into polyominos containing horizontal and vertical benches and the second equality is obtained from proposition 2. □

Case 3. The bench touches no side of the rectangle. Let $g_{3-horiz}(b,k)$ and $g_{3-vert}(b,k)$ be the number of polyominos inscribed in a $b \times k$ rectangle and area $min+1$ containing a $2 \times t$, $t \geq 3$ horizontal and vertical bench polyomino respectively that touches no side of the rectangle.

Proposition 3 *We have*

$$
\begin{aligned}
g_{3-horiz}(b,k) &= 2\sum_{t=3}^{k-2}\sum_{i=2}^{b-2}\sum_{j=2}^{k-t} p_{t,1}(i+1,k-j+1)p_{t,2}(b-i+1,j+t-1)+\\
&\qquad p_{t,1}(i+1,j+t-1)p_{t,2}(b-i+1,k-j+1)-8t\\
&= 2\sum_{t=3}^{k-2}\sum_{i=2}^{b-2}\sum_{j=2}^{k-t}(2\binom{i+k-j-t}{i-1}+2)(2\binom{b+j-i-2}{j-1}+2(t-1))+\\
&\quad 2\sum_{t=3}^{k-2}\sum_{i=2}^{b-2}\sum_{j=2}^{k-t}(2\binom{i+j-2}{i-1}+2)(2\binom{b+k-j-i-t}{b-i-1}+2(t-1)-8t\\
&= 64\,kb-\frac{352}{3}\,k-\frac{8}{3}k^3+40k^2-32(b-1)-16k^2b+16\,(k-4)\binom{b+k-4}{b-2}+\\
&\quad \frac{16b\left(k^2-5\,k+8\right)}{(b+k-3)}\binom{b+k-3}{k-2}-32\binom{b+k-4}{k-4}
\end{aligned}
$$

Proof: As before, we surround the bench with rectangles that reduce our enumeration to case 1 using inclusion-exclusion for two sets. The four surrounding rectangles are arranged in pairs that allow the completion of polyominos along one of the diagonals of the rectangle as shown in figure 6. This gives the first equality of proposition 3. Then we use equations (15) and (16) to obtain the second binomial expression which we reduce to the third expression using standard binomial identities. □

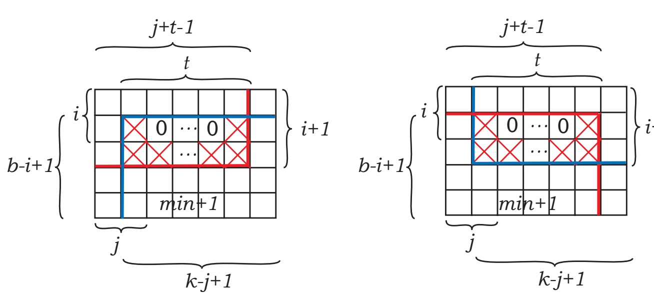


Fig. 6: Decompositions of a polyomino with an inner horizontal bench

Now to complete our count for polyominos in case 3, observe that $g_{3-vert}(b,k) = g_{3-horiz}(k,b)$.

Corollary 5 *The number $g_3(b,k)$ of polyominos inscribed in a rectangle $b\times k$ of area $min+1$ containing a bench polyomino of length $t \geq 3$ touching no side of the rectangle is given by*

$$\begin{aligned}
g_3(b,k) &= g_{3-horiz}(b,k) + g_{3-horiz}(k,b) \\
&= \frac{8}{3}\Big[24 - 6(b^2k+bk^2) + 48bk - 56(b+k) + 15(b^2+k^2) - (b^3+k^3) \\
&\quad -12\left(\binom{b+k-4}{b} + \binom{b+k-4}{k}\right) + 6(b+k-6)\left(\binom{b+k-4}{b-1} + \binom{b+k-4}{k-1}\right) \\
&\quad -60\binom{b+k-4}{b-2} + 18\binom{b+k-2}{b-1}\Big]
\end{aligned}$$

Proof: This is immediate from proposition 3. □

One ingredient is missing to obtain a formula for the number $f_{min+1}(b,k)$. We have to analyse separately the case where the bench has format 2×2 because it contains more symmetries than the other benches and the formulas are not special cases of the formulas for $2\times t$ benches.

Case 4. 2×2 **benches.** The cases are similar to the cases for $2\times t$ benches with $t\geq 3$.

Proposition 4 *a) The number $p_{2\times 2-corner}(b,k)$ of polyominos inscribed in a $b\times k$ rectangle and of area $min+1$ containing a 2×2 bench in the upper left corner satisfies the formulas*

$$\begin{aligned}
p_{2\times 2-corner}(b,k) &= p_{min,(1,1)}(b-1,k-1) + 3 \\
&= \left(2\binom{b+k-4}{b-2} - 1\right) + 3
\end{aligned} \tag{19}$$

b) The number $p_{2\times 2-side}(b,k)$ of polyominos inscribed in a $b\times k$ rectangle of area $min+1$ containing a 2×2 bench along one side and not in a corner of the rectangle satisfies the formula

$$p_{2\times 2-side}(b,k) = \begin{cases} 4(b-3) & \text{if } k=2 \text{ and } b\geq 3 \\ 4(k-3) & \text{if } b=2 \text{ and } k\geq 3 \\ 16\left[\binom{b+k-4}{k-1} + \binom{b+k-4}{b-1} - 2\right] & \text{if } k\geq 3 \text{ and } b\geq 3 \end{cases} \tag{20}$$

c) For integers $b\geq 3$ and $k\geq 3$, the number $p_{2\times 2-center}(b,k)$ of polyominos inscribed in a rectangle $b\times k$, of area $min+1$ and containing a 2×2 bench polyomino that touches no side of the rectangle is given by

$$p_{2\times 2-center}(b,k) = 8\left[\binom{b+k-4}{b-3}(k-3) + \binom{b+k-4}{k-3}(b-3) + \binom{b+k-4}{b-2} + b + k - bk + 1\right] \tag{21}$$

Proof: The proof is similar to the proof for benches of length $t\geq 3$ and we will not repeat the arguments. □

Corollary 6 *For all positive integers b, k, the number $p_{2\times 2}(b,k)$ of polyominos inscribed in a rectangle with area $min+1$ and containing a 2×2 bench is given by the formula*

$$p_{2\times 2}(b,k) = \begin{cases} 0 & \text{if } k=1 \text{ or } b=1 \\ 1 & \text{if } k=2 \text{ and } b=2 \\ 4(b+k-4) & \text{if } (k=2 \text{ and } b>2) \text{ or } (k>2 \text{ and } b>2) \\ 8\left[\binom{b+k-4}{b-2}+2\binom{b+k-4}{b-1}+2\binom{b+k-4}{k-1}-3\right] & \text{if } (k=3 \text{ and } b\geq 3) \text{ or } (k\geq 3 \text{ and } b=3) \\ 8\left[\left(\binom{b+k-4}{b-2}+1\right)(b+k-2)-bk\right] & \text{if } k\geq 4 \text{ and } b\geq 4 \end{cases}$$

Proof: The first three cases are immediate and the last two cases are consequences of proposition 4. □

We are now ready to complete the proof of theorem 3 which is an immediate consequence of the identity

$$p_{min+1}(b,k) = g_1(b,k)+g_2(b,k)+g_3(b,k)+p_{2\times 2}(b,k)$$

and of corollaries 3, 4, 5 and 6.

Proof of corollary 2. This is a consequence of theorem 3 and the identity

$$p_{min+1}(n) = \sum_{b=2}^{n-2} p_{min+1}(b, n-b)$$

n	4	5	6	7	8	9	10	11	12	13
$p_{min+1}(n)$	1	12	80	384	1468	4756	13656	35982	88740	209420

Tab. 2: Number $p_{min+1}(n)$ of polyominos of area n inscribed in a rectangle of perimeter $2n$

3 Applications

We observe two consequences of the formulas developped in the previous section. First, it is possible to count the number $\ell_{min+1}(b,k)$ of lattice animals inscribed in a $b\times k$ rectangle with area $min+1$ because

$$\ell_{min+1}(b,k) = p_{min+1}(b,k) - p_{2\times 2}(b,k). \tag{22}$$

Proposition 5 *For positive integers b, k, the number $\ell_{min+1}(b,k)$ of lattice trees inscribed in a rectangle $b\times k$ and of area $min+1$ is given by the formula*

$$\ell_{min+1}(b,k) = \begin{cases} 0 & \text{if } k\leq 2 \text{ and } b\leq 2 \\ 4b^2-20b+26 & \text{if } (k=2 \text{ and } b>2) \\ & \text{ or } (k>2 \text{ and } b>2) \\ 8(b+k-23)\binom{b+k-4}{b-2}+\frac{8(2k^2+2kb+b-15k+17)}{k-2}\binom{b+k-4}{b-1} & \\ +\frac{8(2b^2+2kb+k-15b+17)}{b-2}\binom{b+k-4}{k-1}+48\binom{b+k-2}{b-1}-\frac{4}{3}(b^3+k^3) & \text{if } (k=3 \text{ and } b\geq 3) \\ -12(b^2k+bk^2)+16(b^2+k^2)+72bk-\frac{266}{3}(b+k)+144, & \text{or } (k\geq 3 \text{ and } b=3) \\ -160\binom{b+k-4}{b-2}+\frac{8(2k^2+2kb+b-13k+13)}{k-2}\binom{b+k-4}{b-1}+136 & \\ +\frac{8(2b^2+2kb+k-13b+13)}{b-2}\binom{b+k-4}{k-1}+48\binom{b+k-2}{b-1}+80bk & \\ -\frac{4}{3}(b^3+k^3)-12(b^2k+bk^2)+16(b^2+k^2)-\frac{290}{3}(b+k), & \text{if } k\geq 4 \text{ and } b\geq 4 \end{cases}$$

Proof: This is an immediate consequence of equation (22), theorem 3 and corollary 6. □

Corollary 7 *For all integers $n \geq 5$ the number $\ell_{min+1}(n)$ of lattice trees of area n inscribed in a rectangle of perimeter $2n$ is given by the formula*

$$\ell_{min+1}(n) = 2^{n+1}(n-1) - \frac{2}{3}(4n^4 - 46n^3 + 227n^2 - 473n + 318) \tag{23}$$

Proof: This is a consequence of proposition 5 and the equation

$$\ell_{min+1}(n) \quad = \quad \sum_{b=2}^{n-2} \ell_{min+1}(b, n-b)$$

□

n	4	5	6	7	8	9	10	11	12	13
$\ell_{min+1}(n)$	0	4	40	232	988	3420	10240	27680	69588	166132

Tab. 3: Number $\ell_{min+1}(n)$ of lattice trees of area n inscribed in a rectangle of perimeter $2n$

Remark. All the formulas described in this paper have been verified numerically with independent computer programs that can construct and count the relevant polyominos.

References

[1] E. Barcucci, A. Frosini, S Rinaldi; *On directed-convex polyominoes in a rectangle*, Discrete Mathematics, 298:62–78, 2005.

[2] M. Bousquet-Mélou; *Codage des polyominos convexes et équations pour l'énumération suivant l'aire*, Discrete Applied Mathematics, 48, 21–43, 1994.

[3] S. Brlek, G. Labelle, A. Lacasse; *Algorithms for polyominoes based on the discrete Green theorem*, Discrete Applied Mathematics, 147, 187–205, 2005.

[4] E. Deutsch; *Enumerating symmetric directed convex polyominoes*, Discrete Mathematics, 280, 225–231, 2004.

[5] J.P. Dubemard, I. Dutour, *Enumération de polyominos convexes dirigés*, Discrete Mathematics, 157,79–90, 1996.

[6] S. Golomb, *Checker Boards and Polyominoes*, Amer. Math. Monthly 61, 675–682, 1954.

[7] I. Jensen, *Enumerations of lattice animals and trees*, Journal of Statistical Physics. 102 no 3-4, 865–881, 2001.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 620–631

Word equations in a uniquely divisible group

Christopher J. Hillar[1†] and Lionel Levine[2‡] and Darren Rhea[3]

[1]*The Mathematical Sciences Research Institute, 17 Gauss Way, Berkeley, CA 94720-5070, USA*
[2]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA USA*
[3]*Department of Mathematics, University of California, Berkeley, CA 94720, USA*

Abstract. We study equations in groups G with unique m-th roots for each positive integer m. A *word equation* in two letters is an expression of the form $w(X, A) = B$, where w is a finite word in the alphabet $\{X, A\}$. We think of $A, B \in G$ as fixed coefficients, and $X \in G$ as the unknown. Certain word equations, such as $XAXAX = B$, have solutions in terms of radicals: $X = A^{-1/2}(A^{1/2}BA^{1/2})^{1/3}A^{-1/2}$, while others such as $X^2AX = B$ do not. We obtain the first known infinite families of word equations not solvable by radicals, and conjecture a complete classification. To a word w we associate a polynomial $P_w \in \mathbb{Z}[x, y]$ in two *commuting* variables, which factors whenever w is a composition of smaller words. We prove that if $P_w(x^2, y^2)$ has an absolutely irreducible factor in $\mathbb{Z}[x, y]$, then the equation $w(X, A) = B$ is not solvable in terms of radicals.

Résumé. Nous étudions des équations dans les groupes G avec les m-th racines uniques pour chaque nombre entier positif m. Une *équation de mot dans deux lettres* est une expression de la forme $w(X, A) = B$, où w est un mot fini dans l'alphabet $\{X, A\}$. Nous pensons $A, B \in G$ en tant que coefficients fixes, et $X \in G$ en tant que inconnu. Certaines équations de mot, telles que $XAXAX = B$, ont des solutions en termes de radicaux: $X = A^{-1/2}(A^{1/2}BA^{1/2})^{1/3}A^{-1/2}$, alors que d'autres tel que $X^2AX = B$ ne font pas. Nous obtenons les familles infinies d'abord connues des équations de mot non solubles par des radicaux, et conjecturons une classification complété. Á un mot w nous associons un polynôme $P_w \in \mathbb{Z}[x, y]$ dans deux variables de permutation, qui factorise toutes les fois que w est une composition de plus petits mots. Nous montrons que si $P_w(x^2, y^2)$ a un facteur absolument irréductible dans $\mathbb{Z}[x, y]$, alors l'équation $w(X, A) = B$ n'est pas soluble en termes de radicaux.

Keywords: absolutely irreducible, polynomials over finite fields, solutions in radicals, uniquely divisible group, word equation

1 Introduction

A group G is called *uniquely divisible* if for every $B \in G$ and each positive integer m, there exists a unique $X \in G$ such that $X^m = B$. We denote the unique such X by $B^{1/m}$, and its inverse by $B^{-1/m}$. In the literature, such groups are also referred to as $\mathbb{Q}$-groups. Note that if it is not the trivial group, then G must be torsion-free, hence infinite. Examples of uniquely divisible groups include the group of positive

[†]Partially supported by an NSA Young Investigators Grant and an NSF All-Institutes Postdoctoral Fellowship administered by the Mathematical Sciences Research Institute through its core grant DMS-0441170
[‡]Supported by an NSF Postdoctoral Fellowship

units of a real closed field, unipotent matrix groups, noncommutative power series with unit constant term, and the group of characters of a connected Hopf algebra over a field of characteristic zero [1].

Inspired by trace conjectures in matrix analysis (see section 2), we study here the natural question of which equations in a uniquely divisible group have solutions in terms of radicals. As a motivating example, consider the Riccati equation $XAX = B$ with $A, B \in G$ given and $X \in G$ unknown. This equation has a unique solution X in any uniquely divisible group; moreover, its solution may be written explicitly (albeit in two distinct ways) as

$$X = A^{-1/2}(A^{1/2}BA^{1/2})^{1/2}A^{-1/2} = B^{1/2}(B^{-1/2}A^{-1}B^{-1/2})^{1/2}B^{1/2}. \tag{1}$$

More generally, let $w(X, A)$ be a finite word in the two-letter alphabet $\{X, A\}$. An expression of the form

$$w(X, A) = B \tag{2}$$

is called a *word equation*. We are interested in classifying those word equations that have a solution in every uniquely divisible group. Clearly, the more general situation in which positive rational exponents on X, A, and B are allowed reduces to this one.

The main tool in our analysis is a new combinatorial object $P_w \in \mathbb{Z}[x, y]$, called the *word polynomial*. If $w = A^{a_0}XA^{a_1}X\cdots A^{a_{n-1}}XA^{a_n}$, we define

$$P_w(x, y) := y^{a_0} + xy^{a_0+a_1} + x^2y^{a_0+a_1+a_2} + \cdots + x^{n-1}y^{a_0+\cdots+a_{n-1}}. \tag{3}$$

For a prime p, let $(\mathbb{Z}/p\mathbb{Z})^*$ denote the set of nonzero elements of the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Theorem 1.1 *There exists a uniquely divisible group G with the following property: For all finite words w in the alphabet $\{X, A\}$, if the equation $P_w(x^2, y^2) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p, then there exist elements $A, B, B' \in G$ for which the word equation $w(X, A) = B$ has no solution $X \in G$, and the word equation $w(X, A) = B'$ has at least two solutions $X \in G$.*

We prove this theorem in section 6. The group G is constructed from an infinite collection of pq-groups whose orders are chosen using Dirichlet's theorem on primes in arithmetic progressions.

Despite appearances, Theorem 1.1 yields a computationally efficient sufficient condition for the equation $w(X, A) = B$ to have no solution in G (see Corollary 1.7). To put Theorem 1.1 in context, we next discuss a family of word equations which *do* have solutions in every uniquely divisible group, along with a conjectured complete classification of such words.

In this paper, "word" will always mean a finite word over the alphabet $\{X, A\}$ (unless another alphabet is specified). The word w is called *universal* if (2) has a solution $X \in G$ for every uniquely divisible group G and each two elements $A, B \in G$; if this solution is always unique, then we say that w is *uniquely universal*. A related class of words is those for which (2) has a solution "in terms of radicals." This notion is defined carefully in section 3. Our explorations give evidence for the surprising conjecture that all three of these classes are in fact the same, and can be characterized as follows.

Definition 1.2 *A word w in the alphabet $\{X, A\}$ is totally decomposable if it is the image of the letter X under a composition of maps of the form*

- $\pi_{m,k} : w \mapsto (wA^k)^m w$, *for $m \geq 1$, $k \geq 0$*

- $r : w \mapsto wA$
- $l : w \mapsto Aw.$

For example, the word $w = XAX^2AXAXAX^2AX$ is totally decomposable, as witnessed by the composition $w = \pi_{1,1} \circ \pi_{1,0} \circ \pi_{1,1}(X)$. According to the following lemma, any totally decomposable word is uniquely universal.

Lemma 1.3 *Let G be a uniquely divisible group, and let w be a totally decomposable word. For any $A, B \in G$, the equation $w(X, A) = B$ has a unique solution $X \in G$, and this solution can be expressed in terms of radicals.*

We conjecture the converse: totally decomposable words are the only universal words.

Conjecture 1.4 *Let w be a finite word in the alphabet $\{X, A\}$. The following are equivalent.*

1. *w is totally decomposable.*
2. *w is uniquely universal.*
3. *w is universal.*
4. *$w(X, A) = B$ has a solution in terms of radicals. (see Definition 3.4)*

The implications $(1) \Rightarrow (2) \Rightarrow (3)$ and $(3) \Leftrightarrow (4)$ are straightforward (see section 3). The remaining implication $(4) \Rightarrow (1)$ is the difficult one. Theorem 1.1 arose out of our attempts to prove this implication. It reduces the noncommutative question about word equations to a commutative question about solutions to polynomial equations mod p. More concretely, we use Theorem 1.1 to prove that several infinite families of word equations are not solvable in terms of radicals (Corollary 1.8). To our knowledge, these are the first such infinite families known.

Together with Theorem 1.1, the following would imply Conjecture 1.4.

Conjecture 1.5 *If w is a word that is not totally decomposable, then the equation $P_w(x^2, y^2) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p.*

The questions outlined above (e.g., asking whether a particular word w is universal, or uniquely universal, or solvable in terms of radicals) are examples of decidability questions in first-order theories of groups. Determining whether a set of equations has a solution in a group is known as the *Diophantine problem*. More generally, given a set of axioms for a class of groups, one would like to provide an algorithm which decides the truth or falsehood of any given sentence in the theory. That such an algorithm exists for free groups follows from pioneering work of Kharlampovich and Myasnikov [16] (see also the independent work of Sela [27]), but it is still open whether one exists for free uniquely divisible groups $F^{\mathbb{Q}}$. The Diophantine problem for $F^{\mathbb{Q}}$ does admit such an algorithm [15] although the time complexity of the algorithm described there is likely at least doubly exponential (the proof uses the decidability of Presburger arithmetic). In contrast, Conjecture 1.4 says that the Diophantine problem of a single equation in one variable reduces to an easily verifiable combinatorial condition (total decomposability).

Example 1 *Consider the word $w = X^2AX$, which is not totally decomposable; we use Theorem 1.1 to show that w is not universal. Its word polynomial is $P_w(x,y) = 1 + x + x^2y$. We need to verify that the equation $1 + x^2 + (x^2y)^2 = 0$ has a nonzero solution modulo p, for all sufficiently large primes p. A standard pigeonhole argument shows that for all primes p, there exist $a, b \in \mathbb{Z}$ with $a \neq 0$ such that $1 + a^2 + b^2 \equiv 0$ (mod p). If $b = 0$ and $p \geq 5$, then $1 + (-1 + 4^{-1})^2 + (a + a4^{-1})^2 \equiv 0$ (mod p) so that we may assume both a, b nonzero. Setting $x = a$ and $y = ba^{-2}$ gives us our solution.*

The word polynomial for the totally decomposable word $v = XAXAX$, on the other hand, is $P_v(x,y) = 1 + xy + x^2y^2$. Let p be a prime greater than 3. If $x, y \in (\mathbb{Z}/p\mathbb{Z})^*$ satisfy $P_v(x^2, y^2) = 0$, then setting $z = x^2y^2$ we have $z^3 = 1$ and $z \neq 1$, which forces $p \equiv 1$ (mod 3). □

Recall that a polynomial over a field K is *absolutely irreducible* if it remains irreducible over every algebraic extension of K. The next result shows that to verify Conjecture 1.5 for a particular word w, it suffices to prove that a factor of $P_w(x^2, y^2)$ is absolutely irreducible.

Proposition 1.6 *Suppose $F \in \mathbb{Z}[x,y]$ satisfies $F(0,0) \neq 0$, and F has a factor $f \in \mathbb{Z}[x,y]$ which is irreducible over $\mathbb{C}[x,y]$. Then the equation $F(x,y) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p.*

Corollary 1.7 *If w is a word in the alphabet $\{X, A\}$ beginning with X, and if $P_w(x^2, y^2)$ has a factor $f \in \mathbb{Z}[x,y]$ such that f is irreducible in $\mathbb{C}[x,y]$, then w is not universal.*

Example 2 *We show the usefulness of Corollary 1.7 by revisiting Example 1. The word $w = X^2AX$ has $P_w(x^2, y^2) = 1 + x^2 + x^4y^2$, which is irreducible over $\mathbb{C}$ (since $1 + x^2$ is not a square in $\mathbb{C}(x)$). It follows that X^2AX is not universal. In contrast, the totally decomposable word $v = XAXAX$ has $P_v(x^2, y^2) = 1 + x^2y^2 + x^4y^4 = (1 + xy + x^2y^2)(1 - xy + x^2y^2)$. Each factor on the right side is irreducible over $\mathbb{Z}$ but factors over $\mathbb{C}$.* □

In Section 7, we use Corollary 1.7 to verify Conjecture 1.5 for the following infinite families of words.

Corollary 1.8 *The following families of words do not have their equations solvable in terms of radicals:*

$$X^nAX^m,\ m,n \geq 1,\ m \neq n;\ \ XA^{m+2n}XA^{m+n}XA^mX,\ m \geq 0, n \geq 1;$$

$$XAX^nAX,\ n \geq 3;\ \ X^2(AX)^nX,\ n \geq 2.$$

Using Corollary 1.7 and the symbolic computation software Maple, we have also verified Conjecture 1.4 for all words of length at most 10. The most difficult part of the computation is to check whether a given bivariate polynomial over $\mathbb{Z}$ is irreducible over $\mathbb{C}$. This can be done in polynomial-time using the algorithm of Gao [8] (and is implemented in Maple).

We do not know if the condition in Corollary 1.7 is sufficient to prove Conjecture 1.4.

Question 1.9 *If the word w is not totally decomposable and begins with X, must $P_w(x^2, y^2)$ have a factor in $\mathbb{Z}[x,y]$ which is irreducible over $\mathbb{C}[x,y]$?*

The remainder of the paper is organized as follows. In section 2, we give additional motivation arising from the BMV trace conjecture in quantum statistical mechanics. In section 3, we review the basic properties of uniquely divisible groups and construct a free uniquely divisible group on two free generators. This construction allows us to define the notion of solvability in terms of radicals, but it is not needed for the proof of Theorem 1.1. Section 4 describes some important examples of uniquely divisible groups.

Most of the standard examples have the property that *every* word equation with nonnegative exponents has a unique solution; the need to construct more exotic groups is part of what makes Conjecture 1.4 so difficult. Section 5 discusses properties of the word polynomial P_w, and section 6 is devoted to the proof of Theorem 1.1. These two sections form the heart of the paper. Finally, Section 7 contains the proof of Corollary 1.8.

2 Background and Motivation

The Lieb-Seiringer formulation [23] of the long-standing Bessis-Moussa-Villani (BMV) trace conjecture [5, 21, 26, 11, 9, 19, 18, 17, 7] in statistical physics says that the trace of $S_{m,k}(A,B)$, the sum of all words of length m in A and B with k Bs, is nonnegative for all $n \times n$ positive semidefinite matrices A and B. In the case of 2×2 matrices, more is true: every word in two positive semidefinite letters has nonnegative trace (in fact, nonnegative eigenvalues). It was unknown whether such a fact held in general until [14] appeared where it was found (with the help of Shaun Fallat) that the word $w = BABAAB$ has negative trace with the positive definite matrices

$$A_1 = \begin{bmatrix} 1 & 20 & 210 \\ 20 & 402 & 4240 \\ 210 & 4240 & 44903 \end{bmatrix} \quad \text{and} \quad B_1 = \begin{bmatrix} 36501 & -3820 & 190 \\ -3820 & 401 & -20 \\ 190 & -20 & 1 \end{bmatrix}.$$

Finding such examples is surprisingly difficult, and randomly generating millions of matrices (from the Wishart distribution) fails to produce them. Nonetheless, it is believed that most words can have negative trace, and it was conjectured [14] that if a word has positive trace for every pair of real positive definite A and B, then it is a palindrome or a product of two palindromes (the converse is well-known). If we replace the words "positive trace" in the previous sentence with "positive eigenvalues," we obtain a weaker conjecture which was also studied in [14]. Further evidence for this conjecture can be found in [12], where it was proved that a generic word has positive definite complex Hermitian matrices A and B giving it a nonpositive eigenvalue.

Positive definite matrices which give a word a negative trace are also potential counterexamples to the BMV conjecture, and it is useful to be able to generate these matrices (see [10, §4.1] and [2, §11] for two such examples). As remarked above, this is difficult since random sampling does not seem to work. This discussion explains some of the subtlety of the BMV conjecture: most words occurring in $S_{m,k}(A,B)$ likely can be made to have negative trace; however, a particular word has a small proportion of matrices which witness this.

Although the set of $n \times n$ positive definite matrices is not a group for $n > 1$, every positive definite matrix has a unique positive definite m-th root for any m. More remarkably, it turns out [13, 2] that every word equation $w(X,A) = B$ with w palindromic (and containing at least one X) has a positive definite solution X for each pair of positive definite A and B (although this solution can be non-unique [2]). Using A_1 and B_1, it follows that any word of the form $wAwAAw$ with $w = w(B,A)$ palindromic (and containing at least one B) can have negative trace. This gives an infinite family verifying the conjecture of [14], and moreover, provides an infinite number of potential counterexamples to the BMV conjecture. The existence proof in [13] uses fixed point methods, although for special cases (e.g. when w contains four or less Xs), one may express solutions X explicitly (and computationally efficiently) in terms of A, B and fractional powers [2, §5]. Computing solutions without using these formal representations "in terms of radicals" is difficult [2, Remark 11.3], and it is believed that most equations do not have solutions expressable in this manner. For instance, there is no known expression for the solution to $XAX^3AX = B$ although there is always a unique positive definite solution [20].

3 Radical words and the free uniquely divisible group

In this section we review some basic properties of uniquely divisible groups, and construct the free uniquely divisible group on two generators. This construction allows us to define precisely the notion of "solvability in terms of radicals" (but we emphasize that the proof of Theorem 1.1 does not rely on this construction). The following lemma shows that rational powers of group elements are well-defined and behave as expected.

Lemma 3.1 *Let G be a uniquely divisible group and $a \in G$. Define $a^{n/m} := (a^n)^{1/m}$ for $n \in \mathbb{Z}$ and $0 \neq m \in \mathbb{N}$, and define $a^0 := 1$. Then if $p, q \in \mathbb{Q}$, we have $(a^p)^q = (a^q)^p = a^{pq}$ and $a^p a^q = a^q a^p = a^{p+q}$.*

A detailed study of uniquely divisible groups can be found in the thesis of Baumslag [3] where they are called *divisible R-groups*. See also [22] for a study of the metabelian case. As remarked in [3], one of the difficulties is that there is no clear normal form for uniquely divisible group elements (for example, see (1) from the introduction). There is, however, the notion of a free uniquely divisible group which comes out of Birkhoff's theory of "varieties of algebras" [6]. Since the construction is simple, we briefly outline the main ideas here. Our perspective is model-theoretic (see [25] for background) although we will use only basic notions from that subject.

Let T be the first-order theory of uniquely divisible groups. The underlying language and axioms of this theory are those of groups, with an additional (countably infinite) set of axioms expressing that every element has a unique m-th root for each positive integer m. Consider the smallest set S of finite, formal expressions containing letters $\{A, B\}$, exponents of the form $^{n/m}$ ($n \in \mathbb{Z}$, $0 \neq m \in \mathbb{N}$), and balanced parentheses that is closed under taking concatenations and powers (and contains the empty expression). For example, S contains the two rightmost expressions in (1).

If G is a uniquely divisible group and $a, b \in G$, then an expression $e = e(A, B) \in S$ defines unambiguously (by Lemma 3.1) an element $e(a, b) \in G$ by replacing letters $\{A, B\}$ with corresponding group elements $\{a, b\}$ and then evaluating the result in G. When two expressions $e, f \in S$ evaluate to the same group element for each pair $a, b \in G$ in every uniquely divisible group G, we write $e \sim f$. For instance, the two rightmost expressions in (1) are equivalent in this way. Although we will not need it here, Gödel's completeness theorem (along with soundness) implies that $e \sim f$ if and only if there is a (finite) formal proof from the axioms of T that they are equal.

Note that $\sim$ is an equivalence relation on S, and we write $[e]$ for the equivalence class containing $e \in S$.

Definition 3.2 *The set $\mathcal{F} := \{[e] : e \in S\}$ with multiplication $[e] \cdot [f] = [ef]$ is called the free uniquely divisible group on letters $L = \{A, B\}$.*

The definition extends in the obvious way to define the free uniquely divisible group on any set L, but (except for a remark at the very end of the paper) we shall only use the case of two generators.

The main facts about $\mathcal{F}$ that we will need are summarized in the following lemma. We remark that any homomorphism $\psi : F \to G$ between uniquely divisible groups is easily seen to satisfy $\psi(a^q) = \psi(a)^q$ for all $a \in F$ and $q \in \mathbb{Q}$.

Lemma 3.3 *$\mathcal{F}$ is a uniquely divisible group. Moreover, $\mathcal{F}$ satisfies a universal property with respect to the map $\theta : L \to \mathcal{F}$ sending $A \mapsto [A]$ and $B \mapsto [B]$: Given any uniquely divisible group G and any map*

$\phi : L \to G$, there exists a unique homomorphism (of uniquely divisible groups) $\psi : \mathcal{F} \to G$ such that $\psi \circ \theta = \phi$.

This discussion allows us to formally define the concept of solution in terms of radicals mentioned in the introduction (specifically, in the statement of Conjecture 1.4).

Definition 3.4 *A word w is called radical (and has equation $w(X, A) = B$ solvable in terms of radicals) if the equation $w(X, [A]) = [B]$ has a solution $X \in \mathcal{F}$.*

We now connect this definition with the idea of word equations having solutions in radicals. Let G be a uniquely divisible group. A subgroup $R \subseteq G$ is *radical* if $x^m \in R$ implies $x \in R$ for all $x \in G$ and all positive integers m. Given a subset H of G, recursively define sets R_n for $n \geq 0$ by setting $R_0 = H$ and

$$R_{n+1} = \{(xy)^q : x, y \in R_n,\ q \in \mathbb{Q}\}.$$

We call the union $\mathcal{R}(H) := \bigcup_{n \in \mathbb{N}} R_n$ the *radical subgroup of G generated by H*. One easily checks that $\mathcal{R}(H)$ is a radical subgroup of G and that it is the intersection of all radical subgroups containing H.

For a uniquely divisible group G and $a, b \in G$, the subgroup $\mathcal{R}(\{a, b\})$ can be thought of as the radical expressions generated by a and b. Given a specific instance of the word equation $w(x, a) = b$, any solution $x \in \mathcal{R}(\{a, b\})$ can be viewed as one "in terms of radicals." Of course, whether a particular word equation in a group has a solution in terms of radicals in this sense depends on the group (and the elements $a, b \in G$). However, as the next lemma shows, a radical word always has such a solution. Note that this verifies the implication $(4) \Rightarrow (3)$ in Conjecture 1.4.

Lemma 3.5 *Let w be a radical word, and let G be a uniquely divisible group. For any $a, b \in G$, the equation $w(x, a) = b$ has a solution x which lies in the radical subgroup $\mathcal{R}(\{a, b\}) \subseteq G$.*

The proof of Lemma 1.3 shows that $(1) \Rightarrow (2)$ in Conjecture 1.4. As the implications $(2) \Rightarrow (3) \Rightarrow (4)$ are trivial, the sole unproved implication is $(4) \Rightarrow (1)$.

4 Examples of Uniquely Divisible Groups

In addition to the free uniquely divisible group encountered in the previous section, there are many interesting examples of uniquely divisible groups. We discuss several of them here, although this list is far from exhaustive.

Recall that a *real closed field* is an ordered field K whose positive elements are squares and such that any polynomial of odd degree with coeffients in K has a zero in K. It follows from the definition that each positive element of a real closed field has a positive m-th root for every positive integer m. Moreover, since the field is ordered, this positive root is unique. The group of positive elements of a real closed field is therefore uniquely divisible.

The rest of our examples are noncommutative. The free group F_2 on the alphabet $\{A, B\}$ may be embedded via the Magnus homomorphism ϕ_M [24] into the algebra $\mathbb{Q}\langle\langle a, b\rangle\rangle$ of noncommutative power series via $A \mapsto 1 + a$ and $B \mapsto 1 + b$. The image of this map is a subgroup of the group D of noncommutative power series with constant term 1. Using the binomial series, it can be shown that D is uniquely divisible (see also Proposition 4.1 below for another proof). In particular, F_2 is a subgroup of a uniquely divisible group.

Our next result shows that every word equation with coefficients in D has a unique solution in that set. However, this solution might not be in the radical subgroup $\mathcal{R}(\phi_M(F_2))$ generated by $\phi_M(F_2)$.

Proposition 4.1 *For any $A_1, \ldots, A_m, B \in D$, the equation $\prod_{i=1}^{m}(A_i X) = B$ has a unique solution $X \in D$.*

Let $\psi : \mathcal{F} \to D$ be the homomorphism of uniquely divisible groups with $\psi([A]) = 1 + a$ and $\psi([B]) = 1 + b$ given by Lemma 3.3. Surprisingly, while the Magnus homomorphism $\phi_M : F_2 \to D$ is an embedding of groups, it is a very old open question whether ψ is also an embedding. As far as we know, Baumslag has the best result on this problem [4], giving injectivity when ψ is restricted to certain one-relator subgroups of $\mathcal{F}$.

Our next example is a matrix group. Let K be a field of characteristic 0 and let UT_n be the group of $n \times n$ *unipotent matrices* over K. These are the upper triangular matrices with coefficients in K with 1's along the diagonal.

Proposition 4.2 *For any $A_1, \ldots, A_m, B \in UT_n$, the equation $\prod_{i=1}^{m}(A_i X) = B$ has a unique solution $X \in UT_n$. (In particular, UT_n is a uniquely divisible group.)*

5 The word polynomial

Given a finite word w over the alphabet $\{X, A\}$, write $w = A^{a_0} X A^{a_1} X \cdots A^{a_{n-1}} X A^{a_n}$ for nonnegative integers $a_0, \ldots, a_n$. The *word polynomial* of w is the polynomial in commuting variables x and y given by (3). For example, the word $w = X^{n-1}AX$ has word polynomial: $P_w(x, y) = 1 + x + x^2 + \cdots + x^{n-2} + x^{n-1}y$. Note that if w ends in X (i.e., $a_n = 0$) then w can be uniquely recovered from P_w.

If u is another word over the same alphabet, the composition $u \circ w$ is the word obtained by replacing each occurrence of the letter X in u by the word w. Although composition of words is not commutative, it can be modeled by multiplication of polynomials according to the following lemma.

Lemma 5.1 *Let u and w be finite words in the alphabet $\{X, A\}$ ending with X, and let m, n be respectively the number of letters in w equal to A, X. Then $P_{u\circ w}(x, y) = P_u(x^n y^m, y) P_w(x, y)$.*

Our next lemma shows another context in which the word polynomial P_w arises: from substituting certain affine transformation matrices for the letters of w.

Lemma 5.2 *Let x, y, z be commuting indeterminates, let $w(X, A)$ be a word, and let m, n be respectively the number of letters in w equal to A, X. Then,* $w\left(\begin{bmatrix} x & z \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} y & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} x^n y^m & P_w(x, y) z \\ 0 & 1 \end{bmatrix}$.

6 Proof of Thoerem 1.1

We begin with the following elementary fact.

Lemma 6.1 *Let G be a group of order n. If m and n are relatively prime, then every element of G has a unique m-th root.*

Let G_i $(i = 1, 2, \ldots)$ be an infinite sequence of finite groups with the following property: For every positive integer m, there exists an N such that

$$m \text{ and } \#G_i \text{ are relatively prime for all } i > N. \tag{4}$$

By Lemma 6.1, these groups have a limiting kind of unique divisibility, which suggests taking the quotient of the direct product of the G_i by their direct sum.

Lemma 6.2 *If $G_1, G_2, \ldots$ is a sequence of finite groups satisfying (4), then $G = \prod_{i=1}^{\infty} G_i \Big/ \bigoplus_{i=1}^{\infty} G_i$ is uniquely divisible.*

This lemma allows us to construct many examples of uniquely divisible groups. Next we describe the sequence of groups G_i that we will use to prove Theorem 1.1.

Let p be an odd prime, and let $q = \frac{p-1}{2}$. Since the group $(\mathbb{Z}/p\mathbb{Z})^*$ of units mod p is cyclic of order $p-1$, we can pick an element $t \in (\mathbb{Z}/p\mathbb{Z})^*$ whose multiplicative order mod p is q (namely, t can be the square of any generator). The powers of t are exactly the nonzero squares, i.e. quadratic residues, in $\mathbb{Z}/p\mathbb{Z}$. We take G_p to be the semidirect product $(\mathbb{Z}/q\mathbb{Z}) \ltimes (\mathbb{Z}/p\mathbb{Z})$, which has the presentation

$$G_p = \left\langle S, T \ : \ S^t T = TS,\, T^q = 1,\, S^p = 1 \right\rangle.$$

The group G_p can be realized concretely as the group of affine transformations of $\mathbb{Z}/p\mathbb{Z}$ of the form $z \mapsto az + b$, where $a \in (\mathbb{Z}/p\mathbb{Z})^*$ is a quadratic residue and $b \in \mathbb{Z}/p\mathbb{Z}$ is arbitrary. Thus we can view G_p as the group of all 2×2 matrices of the form $\begin{bmatrix} t^k & b \\ 0 & 1 \end{bmatrix}$ where $k \in \mathbb{Z}/q\mathbb{Z}$ and $b \in \mathbb{Z}/p\mathbb{Z}$. The generators S and T correspond to the affine transformations $z \mapsto z + 1$ and $z \mapsto tz$, or the matrices: $S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $T = \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}$.

Lemma 6.3 *Let $\alpha, \beta \in \{0, \ldots, q-1\}$ and $\gamma \in \{0, \ldots, p-1\}$. For any word $w = w(X, A)$, the following identity holds in the group G_p: $w\big(S^\gamma T^\beta, T^\alpha\big) = S^{\gamma P_w(t^\beta, t^\alpha)} T^{\alpha m + \beta n}$, where m and n are respectively the number of A's and X's in w.*

Lemma 6.4 *Let w be a finite word in the alphabet $\{X, A\}$, and let n be the number of letters in w equal to X. Let p be a prime such that $q = \frac{p-1}{2}$ is relatively prime to n. If the equation $P_w(x^2, y^2) = 0$ has a solution $(x, y) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$, then there exist $a, b \in G_p$ for which the word equation $w(X, a) = b$ has no solution $X \in G_p$.*

Proof: Suppose $(x, y) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ solves $P_w(x^2, y^2) = 0$. Since any quadratic residue mod p is a power of t, we can find integers α, δ such that $x^2 = t^\delta$ and $y^2 = t^\alpha$. Let $a = T^\alpha$ and $b = ST^{\alpha m + \delta n}$, where m is the number of letters in w equal to A. By Lemma 6.3, an element $X = S^\gamma T^\beta \in G_p$ solves the word equation $w(X, a) = b$ if and only if

$$S^{\gamma P_w(t^\beta, t^\alpha)} T^{\alpha m + \beta n} = b = ST^{\alpha m + \delta n}. \tag{5}$$

Equating powers of T, we obtain $\beta n \equiv \delta n \pmod{q}$. Since n and q are relatively prime, it follows that $\beta \equiv \delta \pmod{q}$, and hence $t^\beta \equiv t^\delta \pmod{p}$. Now equating powers of S in (5) yields

$$1 \equiv \gamma P_w(t^\beta, t^\alpha) \equiv \gamma P_w(t^\delta, t^\alpha) \equiv \gamma P_w(x^2, y^2) \equiv 0 \pmod{p},$$

so there is no solution X to $w(X, a) = b$ in G_p. □

Proof of Theorem 1.1: Let $\pi_0 = 2, \pi_1 = 3, \ldots$ be the primes in increasing order. By the Chinese remainder theorem, for each $i \geq 1$ there is an integer k_i satisfying

$$\begin{aligned} k_i &\equiv 3 \pmod{4} \\ k_i &\equiv 2 \pmod{\pi_j}, \qquad j = 1, \ldots, i. \end{aligned}$$

By Dirichlet's theorem on primes in arithmetic progression, for each i there exists a prime p_i satisfying $p_i \equiv k_i \pmod{4\pi_1 \dots \pi_i}$. By construction, $\frac{p_i-1}{2}$ is not divisible by any of $2, \pi_1, \dots, \pi_i$. Since $\#G_{p_i} = \frac{p_i(p_i-1)}{2}$, the sequence of groups $G_{p_1}, G_{p_2}, \dots$ satisfies condition (4), so by Lemma 6.2 the quotient group $G = \prod_{i\geq 1} G_{p_i} \Big/ \bigoplus_{i\geq 1} G_{p_i}$ is uniquely divisible.

Now let w be a word in the alphabet $\{X, A\}$, and let n be the number of letters in w equal to X. Let π_{i_0} be the largest prime divisor of n. For $i > i_0$ we have $\frac{p_i-1}{2}$ relatively prime to n. By hypothesis, we can choose $i_1 \geq i_0$ sufficiently large so that the equation $P(x^2, y^2) = 0$ has a solution $(x_i, y_i) \in (\mathbb{Z}/p_i\mathbb{Z})^* \times (\mathbb{Z}/p_i\mathbb{Z})^*$ for all $i > i_1$. By Lemma 6.4, for each $i > i_1$, there exist $a_i, b_i \in G_{p_i}$ for which the word equation $w(X, a_i) = b_i$ has no solution $X \in G_{p_i}$. Let $A, B \in G$ be the images of the sequences $(a_i)_{i>i_1}$ and $(b_i)_{i>i_1}$ under the quotient map $\prod_{i>i_1} G_{p_i} \longrightarrow G$. The equation $w(X, A) = B$ has no solution $X \in G$.

Finally, we prove that there exists $B' \in G$ such that $w(X, A) = B'$ has at least two solutions $X \in G$. Using the result just proved, for all $i > i_1$, the map $G_{p_i} \to G_{p_i}$ sending $g \mapsto w(g, a_i)$ is not surjective, hence not injective. Let $g_i, \tilde{g}_i$ be distinct elements of G_{p_i} such that $w(g, a_i) = w(\tilde{g}_i, a_i)$, and let B' be the image in G of the sequence $(w(g_i, a_i))_{i>i_1}$. Then the images in G of the sequences $(g_i)_{i>i_1}$ and $(\tilde{g}_i)_{i>i_1}$ are distinct solutions to $w(X, A) = B'$. ☐

7 Word equations not solvable by radicals

In this section, we use Corollary 1.7 to find several infinite families of words that are not universal, and consequently not solvable in terms of radicals.

Lemma 7.1 *Let m and n be distinct positive integers, and let $w = X^m A X^n$. Then $P_w(x^2, y^2)$ is irreducible in $\mathbb{C}[x, y]$.*

Proof: We view the word polynomial $P_w(x^2, y^2) = \frac{x^{2m}-1}{x^2-1} + y^2 x^{2m} \frac{x^{2n}-1}{x^2-1}$ as a polynomial in y with coefficients in $\mathbb{C}(x)$. If $m < n$, then there exists $\zeta \in \mathbb{C}$ such that $\zeta^{2m} \neq 1$ and $\zeta^{2n} = 1$, in which case ζ is a simple pole of $f(x) = x^{-2m}(x^{2m} - 1)/(x^{2n} - 1)$; likewise, if $m > n$, then f has a simple root. Thus f is not a square in $\mathbb{C}(x)$, which implies that $P_w(x^2, y^2)$ is irreducible in $\mathbb{C}(x)[y]$. ☐

By Corollary 1.7, it follows that for positive integers $m \neq n$, the word equation $X^m A X^n = B$ has no solution in terms of radicals. Our next result shows that for $m \geq 0$ and $n \geq 1$, the word equation $XA^{m+2n}XA^{m+n}XA^mX = B$ has no solution in terms of radicals.

Lemma 7.2 *Let $m \geq 0$ and $n \geq 1$ be integers, and let $w = XA^{m+2n}XA^{m+n}XA^mX$. Then $P_w(x^2, y^2)$ has a factor in $\mathbb{Z}[x, y]$ which is irreducible over $\mathbb{C}[x, y]$.*

Next, we show that the word equation $XAX^nAX = B$ has no solution in terms of radicals if $n \geq 3$.

Lemma 7.3 *Let $n \geq 3$ be an integer, and let $w = XAX^nAX$. Then $P_w(x^2, y^2)$ is irreducible in $\mathbb{C}[x, y]$.*

To further extend these families of words not solvable by radicals, note that under the hypotheses of Theorem 1.1, we can actually derive a slightly stronger conclusion.

Corollary 7.4 *Let u, w be finite words in the alphabet $\{X, A\}$. If $P_w(x^2, y^2) = 0$ has a solution $(x_p, y_p) \in (\mathbb{Z}/p\mathbb{Z})^* \times (\mathbb{Z}/p\mathbb{Z})^*$ for all but finitely many primes p, then the word equations $u \circ w(X, A) = B$ and $w \circ u(X, A) = B$ have no solution in terms of radicals.*

A simple substitution also proves the following.

Corollary 7.5 *Let $n \geq 2$ be an integer. The word equation $X^2(AX)^nX = B$ is not solvable by radicals.*

8 Acknowledgements

The authors thank George Bergman, Pavel Etingof, Kiran Kedlaya, Igor Klep, Alexei Miasnikov, Bjorn Poonen, and Kate Stange for helpful conversations.

References

[1] M. Aguiar, N. Bergeron, and F. Sottile. Combinatorial Hopf algebras and generalized Dehn-Sommerville equations. *Compos. Math.*, 142:1–30, 2006.

[2] S. Armstrong and C. Hillar. Solvability of symmetric word equations in positive definite letters. *J. London Math. Soc.*, 76:777–796, 2007.

[3] G. Baumslag. Some aspects of groups with unique roots. *Acta Math.*, 10:277–303, 1960.

[4] G. Baumslag. On the residual nilpotence of certain one-relator groups. *Comm. Pure Appl. Math.*, 21:491–506, 1968.

[5] D. Bessis, P. Moussa, and M. Villani. Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics. *J. Math. Phys.*, 16:2318–2325, 1975.

[6] G. Birkhoff. On the structure of abstract algebras. *Proc. Cambridge Philos. Soc.*, 31:433–454, 1935.

[7] B. Collins, K. Dykema, and F. Torres-Ayala. Sum-of-squares results for polynomials related to the Bessis-Moussa-Villani conjecture, 2009. `http://arxiv.org/abs/0905.0420`.

[8] S. Gao. Factoring multivariate polynomials via partial differential equations. *Math. Comput.*, 72(242):801–822, 2003.

[9] D. Hagele. Proof of the cases $p \leq 7$ of the Lieb-Seiringer formulation of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.*, 127:1167–1171, 2007.

[10] F. Hansen. Trace functions as Laplace transforms. *J. Math. Phys.*, 47:043504, 2006.

[11] C. Hillar. Advances on the Bessis-Moussa-Villani trace conjecture. *Lin. Alg. Appl.*, 426:130–142, 2007.

[12] C. Hillar and C. R. Johnson. Positive eigenvalues of generalized words in two Hermitian positive definite matrices. In P. Pardalos and H. Wolkowicz, editors, *Novel Approaches to Hard Discrete Optimization*, volume 37 of *Fields Institute Communications*, pages 111–122, 2003.

[13] C. Hillar and C. R. Johnson. Symmetric word equations in two positive definite letters. *Proc. Amer. Math. Soc.*, 132:945–953, 2004.

[14] C. R. Johnson and C. Hillar. Eigenvalues of words in two positive definite letters. *SIAM J. Matrix Anal. Appl.*, 23:916–928, 2002.

[15] O. Kharlampovich and A. Myasnikov. Equations in a free $\mathbf{Q}$-group. *Trans. Amer. Math. Soc.*, 350(3):947–974, 1998.

[16] O. Kharlampovich and A. Myasnikov. Elementary theory of free non-abelian groups. *J. Algebra*, 302(2):451–552, 2006.

[17] I. Klep and M. Schweighofer. Connes' embedding conjecture and sums of Hermitian squares. *Adv. Math.*, 217:1816–1837, 2008.

[18] I. Klep and M. Schweighofer. Sums of Hermitian squares and the BMV conjecture. *J. Stat. Phys.*, 133:739–760, 2008.

[19] P. Landweber and E. Speer. On D. Hagele's approach to the Bessis-Moussa-Villani conjecture. *Lin. Alg. Appl.*, 431:1317–1324, 2009.

[20] J. Lawson and Y. Lim. Solving symmetric matrix word equations via symmetric space machinery. *Lin. Alg. Appl.*, 414:560–569, 2006.

[21] K. J. Le Couteur. Representation of the function Tr$(\exp(A - \lambda B))$ as a Laplace transform with positive weight and some matrix inequalities. *J. Phys. A: Math. Gen.*, 13:3147–3159, 1980.

[22] J. Ledlie. Representations of free metabelian $\mathcal{D}_\pi$-groups. *Trans. Amer. Math. Soc.*, 153:307–346, 1971.

[23] E. H. Lieb and R. Seiringer. Equivalent forms of the Bessis-Moussa-Villani conjecture. *J. Stat. Phys.*, 115:185–190, 2004.

[24] W. Magnus. On a theorem of Marshall Hall. *Ann. of Math.*, pages 764–768, 1939.

[25] D. Marker. *Model theory: an introduction*, volume 217 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2010.

[26] P. Moussa. On the representation of tr $\left(e^{A-\lambda B}\right)$ as a Laplace transform. *Rev. Math. Phys.*, 12:621–655, 2000.

[27] Z. Sela. Diophantine geometry over groups. VI. The elementary theory of a free group. *Geom. Funct. Anal.*, 16(3):707–730, 2006.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 632–642

Constant term evaluation for summation of C-finite sequences

Qing-Hu Hou[1] and Guoce Xin[2]

[1] *Center for Combinatorics, LPMC-TJKLC, Nankai University, Tianjin 300071, PR China*
[2] *Department of Mathematics, Capital Normal University, Beijing 100084, PR China*
[1] *hou@nankai.edu.cn,* [2] *guoce.xin@gmail.com*

Abstract. Based on constant term evaluation, we present a new method to compute a closed form of the summation $\sum_{k=0}^{n-1}\prod_{j=1}^{r}F_j(a_jn+b_jk+c_j)$, where $\{F_j(k)\}$ are C-finite sequences and a_j and a_j+b_j are nonnegative integers. Our algorithm is much faster than that of Greene and Wilf.

Résumé. En s'appuyant sur l'évaluation de termes constants, nous présentons une nouvelle méthode pour calculer une forme close de la somme $\sum_{k=0}^{n-1}\prod_{j=1}^{r}F_j(a_jn+b_jk+c_j)$, où les $\{F_j(k)\}$ sont des suites C-finies, et où les a_j et les a_j+b_j sont des entiers positifs ou nuls. Notre algorithme est beaucoup plus rapide que celui de Greene et Wilf.

Keywords: C-finite sequences, constant term, summation, closed form

1 Introduction

A sequence $\{F(k)\}_{k\geq 0}$ is C-finite (see [Zei90]) if there exist constants $c_1,\ldots,c_d$ such that

$$F(k)=c_1F(k-1)+c_2F(k-2)+\cdots+c_dF(k-d),\quad \forall\, k\geq d.$$

Correspondingly, the integer d is called the order of the recurrence. Greene and Wilf [GW07] provided a method to compute a closed form of the summation

$$\sum_{k=0}^{n-1}\prod_{j=1}^{r}F_j(a_jn+b_jk+c_j),$$

where $\{F_j(k)\}$ are C-finite sequences and a_j,b_j are integers satisfying $a_j\geq 0$ and $a_j+b_j\geq 0$. They proved that the sum must be a linear combination of the terms

$$\prod_{j=1}^{r}F_j((a_j+b_j)n+i_j)\quad\text{and}\quad \phi_{i_1,\ldots,i_j}(n)\prod_{j=1}^{r}F_j(a_jn+i_j),\quad (0\leq i_j<d_j)$$

where d_j is the order of the recurrence of $\{F_j(k)\}$ and $\phi_{i_1,\ldots,i_r}(n)$ is a polynomial in n with given degree bound. Then the explicit formula of the sum can be computed by the method of undetermined coefficients.

In this paper, we provide another approach which is based on MacMahon's partition analysis [Mac16] and the Omega calculations [APR01, Xin04]. We first introduce an extra variable z and consider the summation

$$S(z) = \sum_{k=0}^{n-1} z^k \prod_{j=1}^{r} F_j(a_j n + b_j k + c_j).$$

Then we rewrite $S(z)$ as the constant term (with respect to $x_1, \ldots, x_r$) of the Laurent series

$$f_1(x_1) f_2(x_2) \cdots f_r(x_r) \sum_{k=0}^{n-1} z^k \prod_{j=1}^{r} x_j^{-a_j n - b_j k - c_j},$$

where

$$f_j(x_j) = \sum_{k=0}^{\infty} F_j(k) x_j^k$$

is the generating function. Using partial fraction decomposition, we can derive an explicit formula for $S(z)$ in terms of $\prod_{j=1}^{r} F_j((a_j + b_j)n + i_j)$ and $\prod_{j=1}^{r} F_j(a_j n + i_j)$, where $0 \leq i_j < d_j$. Finally, the substitution of $z = 1$ leads to a closed form of the original summation.

2 Basic tools by partial fraction decomposition

Let K be a field. Fix a polynomial $D(x) \in K[x]$. For any polynomial $P(x) \in K[x]$, we use $\mathrm{rem}(P(x), D(x), x)$ (or $\mathrm{rem}(P(x), D(x))$ for short) to denote the remainder of $P(x)$ when divided by $D(x)$. This notation is extended for rational function $R(x) = P(x)/Q(x)$ when $Q(x)$ is coprime to $D(x)$:

$$\mathrm{rem}(R(x), D(x)) := \mathrm{rem}(P(x)\beta(x), D(x)), \quad \text{if } \alpha(x)D(x) + \beta(x)Q(x) = 1. \tag{1}$$

In algebraic language, the remainder is the standard representative in the quotient ring $K[x]/\langle D(x) \rangle$.

It is convenient for us to use the following notation:

$$\left\{ \frac{P(x)/Q(x)}{D(x)} \right\} = \frac{\mathrm{rem}(P(x)/Q(x), D(x))}{D(x)}. \tag{2}$$

Equivalently, if we have the following partial fraction decomposition:

$$\frac{P(x)}{Q(x)D(x)} = p(x) + \frac{r_1(x)}{D(x)} + \frac{r_2(x)}{Q(x)},$$

where $p(x), r_1(x), r_2(x)$ are polynomials with $\deg r_1(x) < \deg D(x)$, then we claim that $r_1(x) = \mathrm{rem}(P(x)/Q(x), D(x))$ and hence

$$\left\{ \frac{P(x)/Q(x)}{D(x)} \right\} = \frac{r_1(x)}{D(x)}.$$

Note that we do not need $\deg r_2(x) < \deg Q(x)$.

The following properties are transparent:

$$P_1(x) \equiv P_2(x) \pmod{D(x)} \Rightarrow \left\{ \frac{R(x)P_1(x)}{D(x)} \right\} = \left\{ \frac{R(x)P_2(x)}{D(x)} \right\}; \tag{3}$$

$$\alpha(x)D(x) + \beta(x)Q(x) = 1 \Rightarrow \left\{ \frac{P(x)/Q(x)}{D(x)} \right\} = \left\{ \frac{P(x)\beta(x)}{D(x)} \right\}; \tag{4}$$

$$\left\{ \frac{aR_1(x) + bR_2(x)}{D(x)} \right\} = a \left\{ \frac{R_1(x)}{D(x)} \right\} + b \left\{ \frac{R_2(x)}{D(x)} \right\}, \quad \forall a, b \in K. \tag{5}$$

The crucial lemma in our calculation is as follows.

Lemma 1 *Let $R(x), D(x)$ be as above and assume $D(0) \neq 0$. Then for any Laurent polynomial $L(x)$ with $\deg L(x) \leq 0$, we have*

$$\operatorname*{CT}_x L(x) \left\{ \frac{R(x)}{D(x)} \right\} = \operatorname*{CT}_x \left\{ \frac{L(x)R(x)}{D(x)} \right\}, \tag{6}$$

where $\operatorname*{CT}_x g(x)$ means to take constant term of the Laurent series $g(x)$ in x.

Proof: By linearity, we may assume $L(x) = x^{-k}$ for some $k \geq 0$.

Assume $r(x) = \operatorname{rem}(R(x), D(x))$. Since $D(0) \neq 0$, we have the following partial fraction decomposition:

$$\frac{r(x)}{x^k D(x)} = \frac{p(x)}{x^k} + \frac{r_1(x)}{D(x)},$$

where $\deg p(x) < k$ and $\deg r_1(x) < \deg D(x)$. Then taking constant term in x gives

$$\operatorname*{CT}_x \frac{r(x)}{x^k D(x)} = \operatorname*{CT}_x \frac{r_1(x)}{D(x)} = \operatorname*{CT}_x \left\{ \frac{x^{-k} r(x)}{D(x)} \right\} = \operatorname*{CT}_x \left\{ \frac{x^{-k} R(x)}{D(x)} \right\}.$$

This is just (6) when $L(x) = x^{-k}$. □

Let $\mathbb{Z}$ and $\mathbb{N}$ denote the set of integers and nonnegative integers respectively. Suppose that $\{F(k)\}_{k \in \mathbb{N}}$ is a C-finite sequence such that

$$F(k) = c_1 F(k-1) + c_2 F(k-2) + \cdots + c_d F(k-d) \tag{7}$$

holds for any integer $k \geq d$. Then its generating function is of the form

$$f(x) = \sum_{k=0}^{\infty} F(k) x^k = \frac{p(x)}{1 - c_1 x - c_2 x^2 - \cdots - c_d x^d},$$

where $p(x)$ is a polynomial in x of degree less than d. We will say that $\{F(k)\}_{k \geq \mathbb{N}}$ is a C-finite sequence with generating function $p(x)/q(x)$, where $q(x) = 1 - c_1 x - c_2 x^2 - \cdots - c_d x^d$.

It is well-known [Sta86, Section 4.2] that we can uniquely extend the domain of $F(k)$ to $k \in \mathbb{Z}$ by requiring that (7) holds for any $k \in \mathbb{Z}$. The k-th term of the extended sequence can be given by the following lemma.

Lemma 2 *Let $\{F(k)\}_{k\in\mathbb{N}}$ be a C-finite sequence with generating function $p(x)/q(x)$ and $\{F(k)\}_{k\in\mathbb{Z}}$ be its extension. Then*

$$F(k) = \operatorname*{CT}_x \left\{ \frac{x^{-k}p(x)}{q(x)} \right\} = \left\{ \frac{x^{-k}p(x)}{q(x)} \right\} \Big|_{x=0} \quad \forall\, k \in \mathbb{Z}. \tag{8}$$

Proof: Since $q(0) = 1$, the second equality holds trivially. Let

$$G(k) = \operatorname*{CT}_x \left\{ \frac{x^{-k}p(x)}{q(x)} \right\}.$$

Then for $k \geq 0$, applying Lemma 1 gives

$$G(k) = \operatorname*{CT}_x x^{-k} \left\{ \frac{p(x)}{q(x)} \right\} = \operatorname*{CT}_x x^{-k} \frac{p(x)}{q(x)} = [x^k]f(x) = F(k),$$

where $[x^k]f(x)$ means to take the coefficient of x^k in $f(x)$.

Therefore, by the uniqueness of the extension, it suffices to show that $G(k)$ also satisfy the recursion (7) for all $k \in \mathbb{Z}$. We compute as follows:

$$\begin{aligned}
&G(k) - c_1 G(k-1) - \cdots - c_d G(k-d) \\
&\quad = \operatorname*{CT}_x \left\{ \frac{x^{-k}p(x)}{q(x)} \right\} - c_1 \left\{ \frac{x^{-k+1}p(x)}{q(x)} \right\} - \cdots - c_d \left\{ \frac{x^{-k+d}p(x)}{q(x)} \right\} \\
&\quad = \operatorname*{CT}_x \left\{ \frac{x^{-k}p(x)q(x)}{q(x)} \right\} = 0.
\end{aligned}$$

This completes the proof. □

3 Constant term evaluation

Let $\{F_j(k)\}_{k\in\mathbb{Z}}$ be C-finite sequences with generating functions $f_j(x) = p_j(x)/q_j(x)$ for $j = 1, 2, \ldots, r$. We denote the degree of the denominators by $d_j = \deg q_j(x)$. To evaluate the sum

$$S = \sum_{k=0}^{n-1} \prod_{j=1}^{r} F_j(a_j n + b_j k + c_j),$$

we evaluate the more general sum $S_r(z)$ instead, where $S_m(z)$ is defined by

$$S_m(z) = \sum_{k=0}^{n-1} z^k \prod_{j=1}^{m} F_j(a_j n + b_j k + c_j), \quad 0 \leq m \leq r. \tag{9}$$

The advantage is that $S_m(z)$ can be evaluated recursively. Since $a_j \geq 0$ and $a_j + b_j \geq 0$, we have $a_j n + b_j k \geq 0$ for any $n \geq 0$ and $0 \leq k < n$. By Lemmas 1 and 2, we have

$$\begin{aligned} S_m(z) &= \sum_{k=0}^{n-1} z^k \prod_{j=1}^{m-1} F_j(a_j n + b_j k + c_j) \operatorname*{CT}_x \left\{ \frac{x^{-a_m n - b_m k - c_m} p_m(x)}{q_m(x)} \right\} \\ &= \sum_{k=0}^{n-1} z^k \prod_{j=1}^{m-1} F_j(a_j n + b_j k + c_j) \operatorname*{CT}_x x^{-a_m n} x^{-b_m k} \left\{ \frac{x^{-c_m} p_m(x)}{q_m(x)} \right\} \\ &= \operatorname*{CT}_x x^{-a_m n} \left\{ \frac{x^{-c_m} p_m(x)}{q_m(x)} \right\} \sum_{k=0}^{n-1} (z x^{-b_m})^k \prod_{j=1}^{m-1} F_j(a_j n + b_j k + c_j). \end{aligned}$$

Therefore, we obtain the recursion

$$S_m(z) = \operatorname*{CT}_x x^{-a_m n} \left\{ \frac{p_m(x) x^{-c_m}}{q_m(x)} \right\} S_{m-1}(z x^{-b_m}). \tag{10}$$

The initial condition is $S_0(z) = 1 + z + \cdots + z^{n-1} = \frac{1-z^n}{1-z}$.

Let L_m and L'_m be the linear operators acting on Laurent polynomials in $x_1, \ldots, x_m$ by

$$\begin{aligned} L_m \left(\textstyle\prod_{j=1}^m x_j^{\alpha_j} \right) &= \textstyle\prod_{j=1}^m F_j(a_j n - \alpha_j), \\ L'_m \left(\textstyle\prod_{j=1}^m x_j^{\alpha_j} \right) &= \textstyle\prod_{j=1}^m F_j((a_j + b_j) n - \alpha_j). \end{aligned} \tag{11}$$

Then $S_m(z)$ have simple rational function representations.

Theorem 3 *For any $0 \leq m \leq r$, there exist a polynomial $P_m(z)$ with coefficients being Laurent polynomials in $x_1, \ldots, x_m$ and a non-zero polynomial $Q_m(z) \in K[z]$ such that*

$$S_m(z) = \frac{L_m(P_m(z)) - z^n L'_m(P_m(z))}{Q_m(z)}, \tag{12}$$

where L_m, L'_m are defined by (11).

Proof: We prove the theorem by induction on m.

Setting $P_0(z) = 1$ and $Q_0(z) = 1 - z$, we see that the assertion holds for $m = 0$. Suppose that the assertion holds for $m - 1$. We can compute $P_m(z)$ and $Q_m(z)$ as follows. For brevity, we write $R(z) = P_{m-1}(z)/Q_{m-1}(z)$.

By definition $S_{m-1}(z)$ is a polynomial in z of degree less than n. If $b_m \geq 0$, then $-a_m n \leq 0$; If $b_m \leq 0$, then $-a_m n - b_m(n-1) \leq 0$. Thus $x^{-a_m n} S_{m-1}(z x^{-b_m})$ is always a Laurent polynomial of degree no more than 0. Therefore, by Lemma 1 and the recursion (10), we have

$$\begin{aligned} S_m(z) &= \operatorname*{CT}_x \left\{ \frac{p_m(x)x^{-c_m}}{q_m(x)} \right\} x^{-a_m n} S_{m-1}(zx^{-b_m}) \\ &= \operatorname*{CT}_x \left\{ \frac{p_m(x)x^{-c_m} x^{-a_m n} S_{m-1}(zx^{-b_m})}{q_m(x)} \right\} \\ &= \operatorname*{CT}_x \left\{ \frac{p_m(x)x^{-c_m} x^{-a_m n} \left(L_{m-1}(R(zx^{-b_m})) - z^n x^{-b_m n} L'_{m-1}(R(zx^{-b_m})) \right)}{q_m(x)} \right\} \\ &= L_{m-1} \operatorname*{CT}_x x^{-a_m n} G(x,z) - z^n L'_{m-1} \operatorname*{CT}_x x^{-(a_m+b_m)n} G(x,z), \end{aligned}$$

where $G(x,z)$ is given by

$$G(x,z) = \left\{ \frac{p_m(x)x^{-c_m} R(zx^{-b_m})}{q_m(x)} \right\}.$$

Now set

$$\frac{u(x,z)}{w(z)} = \operatorname{rem}(x^{d_m-1-c_m} R(zx^{-b_m}), q_m(x), x), \tag{13}$$

where $u(x,z)$ is a polynomial in x, z and $w(z)$ is a polynomial in z. Then

$$\begin{aligned} G(x,z) &= \left\{ \frac{p_m(x)x^{-d_m+1} \cdot x^{d_m-1-c_m} R(zx^{-b_m})}{q_m(x)} \right\} \\ &= \left\{ \frac{p_m(x)x^{-d_m+1} u(x,z)/w(z)}{q_m(x)} \right\} \\ &= \frac{1}{w(z)} \left\{ \frac{p_m(x) \cdot x^{-d_m+1} u(x,z)}{q_m(x)} \right\}. \end{aligned}$$

Since $x^{-d_m+1}u(x,z)$ is a Laurent polynomial of degree in x less than or equal to 0, we obtain

$$\begin{aligned} S_m(z) &= L_{m-1} \operatorname*{CT}_x \frac{x^{-d_m+1}u(x,z)}{x^{a_m n}} \left\{ \frac{p_m(x)}{q_m(x)} \right\} - z^n L'_{m-1} \operatorname*{CT}_x \frac{x^{-d_m+1}u(x,z)}{x^{(a_m+b_m)n}} \left\{ \frac{p_m(x)}{q_m(x)} \right\} \\ &= L_{m-1} \operatorname*{CT}_x \frac{x^{-d_m+1}u(x,z)}{x^{a_m n}} f_m(x) - z^n L'_{m-1} \operatorname*{CT}_x \frac{x^{-d_m+1}u(x,z)}{x^{(a_m+b_m)n}} f_m(x). \end{aligned}$$

Now set

$$P_m(z) = x_m^{-d_m+1} u(x_m, z), \qquad Q_m(z) = w(z). \tag{14}$$

It is then easy to check that $S_m(z)$ has the desired form. This completes the induction. □

Remark 1. Form the above proof we see that the degree of x_m in $P_m(z)$ is between $-d_m+1$ and 0. Therefore the coefficients of the numerator of $S(z)$ are linear combinations of the form

$$\prod_{j=1}^{r} F_j(a_j n + i_j), \qquad \prod_{j=1}^{r} F_j((a_j+b_j)n + i_j),$$

where $0 \le i_j \le d_j - 1$.

Remark 2. Let $\{F(k)\}$ be a sequence with generating function $p(x)/q(x)$. We call the sequence $\{\overline{F}(k)\}$ with generating function $1/q(x)$ its *primitive* sequence. It is more convenient to represent $S(z)$ in terms of the *primitive* sequences $\{\overline{F}_j(k)\}$ instead of the sequences $\{F_j(k)\}$ themselves. The existence of a such representation is obvious since $F_j(k)$ is a linear combination of $\overline{F}_j(k)$. In this way, the coefficients of the numerator of $S(z)$ will be linear combinations of the form

$$\prod_{j=1}^{r} \overline{F}_j(a_j n - i_j), \qquad \prod_{j=1}^{r} \overline{F}_j((a_j + b_j)n - i_j),$$

where $0 \le i_j \le d_j - 1$. Then we can take advantage of the fact $\bar{F}_j(-i_j) = 0,\ 1 \le i_j \le d_j - 1$ if $a_j = 0$ or $a_j + b_j = 0$. The computation is similar and in a natural way. In fact, if we define

$$\frac{P_m(z)}{Q_m(z)} = \operatorname{rem}\left(p_m(x) x^{-c_m} \frac{P_{m-1}(z x^{-b_m})}{Q_{m-1}(z x^{-b_m})}, q_m(x), x\right),$$

then we have

$$S_m(z) = \frac{\overline{L}_m(P_m(z)) - z^n \overline{L}'_m(P_m(z))}{Q_m(z)},$$

where

$$\overline{L}_m\left(\prod_{j=1}^{m} x_j^{\alpha_j}\right) = \prod_{j=1}^{m} \overline{F}_j(a_j n - \alpha_j),$$

$$\overline{L}'_m\left(\prod_{j=1}^{m} x_j^{\alpha_j}\right) = \prod_{j=1}^{m} \overline{F}_j((a_j + b_j)n - \alpha_j).$$

4 Evaluation of $S_r(z)$ at $z = 1$

In this section, we consider the evaluation of $S_r(z)$ at $z = 1$, which is equals to the sum

$$S = \sum_{k=0}^{n-1} \prod_{j=1}^{r} F_j(a_j n + b_j k + c_j). \tag{15}$$

The evaluation of $S_r(z)$ at $z = 1$ can be obtained by the following lemma.

Lemma 4 *Let $f(z) = \sum_i f_i z^i$, $g(z) = \sum_i g_i z^i$ and $h(z) = \sum_i h_i z^i$ be polynomials in z. Suppose that*

$$S(z) = \frac{f(z) - z^n g(z)}{h(z)}$$

is a polynomial in z and

$$h(z) = \sum_{i \ge e} \tilde{h}_i (z-1)^i, \quad \tilde{h}_e \ne 0. \tag{16}$$

Then

$$S(1) = \frac{1}{\tilde{h}_e} \sum_i \left(f_i \binom{i}{e} - g_i \binom{n+i}{e} \right).$$

Proof: By expanding $f(z) - z^n g(z)$ at the point $z = 1$, we obtain

$$f(z) - z^n g(z) = \sum_i f_i(z-1+1)^i - g_i(z-1+1)^{n+i} = \sum_j (z-1)^j A_j,$$

where

$$A_j = \sum_i \left(f_i \binom{i}{j} - g_i \binom{n+i}{j} \right).$$

Since $S(z)$ is a polynomial in z and $\tilde{h}_e \neq 0$, we have $A_j = 0$ for any $j < e$ and and $S(1) = A_e/\tilde{h}_e$, as desired. □

Remark. Alternatively, we can write

$$S(z) = z^n \frac{z^{-n} f(z) - g(z)}{h(z)}.$$

A similar argument yields

$$S(1) = \frac{1}{\tilde{h}_e} \sum_i \left(f_i \binom{i-n}{e} - g_i \binom{i}{e} \right),$$

The algorithm CFsum for finding a closed form of the sum (15).

Input: The generating functions $p_j(x)/q_j(x)$ of $F_j(k)$ and the parameters (a_j, b_j, c_j)
Output: A closed formula for $S = \sum_{k=0}^{n-1} \prod_{j=1}^{r} F_j(a_j n + b_j k + c_j)$.

1. Initially set $P(z) = 1$ and $Q(z) = 1 - z$.
2. For $j = 1, 2, \ldots, r$ do

   Set $R(z) = P(z)/Q(z)$.
   Let
   $$\frac{u(x,z)}{w(z)} = \operatorname{rem}(p_j(x) x^{-c_j} R(z x^{-b_j}), q_j(x), x).$$
   Set $P(z) = u(x_j, z)$ and $Q(z) = w(z)$.

3. Set $A = B = P$.
4. For $j = 1, 2, \ldots, r$ do

   $$A = \sum_{i=0}^{d_j - 1} \overline{F}_j(a_j n + d_j - i)[x_j^i] A, \quad B = \sum_{i=0}^{d_j - 1} \overline{F}_j((a_j + b_j)n + d_j - i)[x_j^i] B,$$

   where $[x^i]f(x)$ denotes the coefficient of x^i in $f(x)$ and $\{\overline{F}_j(k)\}$ is the primitive sequence corresponding to $\{F_j(k)\}$.

5. Let e be the lowest degree of z in $Q(z+1)$ and $h = [z^e]Q(z+1)$.

6. Finally, return

$$S = \frac{1}{h}\sum_i \left(\binom{i}{e}[z^i]A - \binom{n+i}{e}[z^i]B\right).$$

Our algorithm suggested a new way to look at the degree bound for the coefficients $\phi_{i_1,\dots,i_r}(n)$. One bound is just the multiplicity of 1 as a root of $Q_r(z)$. To study the e described in (16), it is better to use the alternative representation of $S(z) = S_r(z)$:

$$S(z) \quad = \quad \underset{x_1,\dots,x_r}{\mathrm{CT}} \left(\frac{1}{x_1^{a_1}\cdots x_r^{a_r}}\right)^n \frac{1-(\frac{z}{x_1^{b_1}\cdots x_r^{b_r}})^n}{1-\frac{z}{x_1^{b_1}\cdots x_r^{b_r}}} \prod_{j=1}^r \left\{\frac{p_j(x_j)x_j^{-c_j}}{q_j(x_j)}\right\}.$$

Suppose that α_j is a root of $q_j(x)$ with multiplicity $\nu_j(\alpha_j)$. By partial fraction decomposition, $S(z)$ can be written as a linear combination of the terms

$$S(z;\boldsymbol{\alpha},\boldsymbol{s}) = \underset{x_1,\dots,x_r}{\mathrm{CT}} \left(\frac{1}{x_1^{a_1}\cdots x_r^{a_r}}\right)^n \frac{1-(\frac{z}{x_1^{b_1}\cdots x_r^{b_r}})^n}{1-\frac{z}{x_1^{b_1}\cdots x_r^{b_r}}} \prod_{j=1}^r \frac{1}{(1-x_j/\alpha_j)^{s_j}},$$

where $s_j \le \nu_j(\alpha_j)$. From the discussion on Omega operator [Xin04], we see that the denominator of $S(z;\boldsymbol{\alpha},\boldsymbol{s})$ is given by

$$\left(1-\frac{z}{\alpha_1^{b_1}\cdots\alpha_r^{b_r}}\right)^{s_1+s_2+\cdots+s_r-r+1}$$

Therefore, by summing over all $\boldsymbol{\alpha},\boldsymbol{s}$ and take common denominator, we see that

$$e \le \max\{\nu_1+\cdots+\nu_r-r+1:\ \alpha_1^{b_1}\cdots\alpha_r^{b_r}=1 \ \text{ and } \ q_j(\alpha_j)=0\}.$$

5 Examples

We have implement the algorithm *CFSum* in `Maple`, which can be download from
http://www.combinatorics.net.cn/homepage/xin/maple/CFsum.txt .
Example 1. Let

$$f(n) = \sum_{k=0}^{n-1} F(k)^2 F(2n-k),$$

where $\{F(k)\}$ is the Fibonacci sequence defined by

$$F(0)=0,\ F(1)=1,\ F(k)=F(k-1)+F(k-2),\ \forall\, k\ge 2.$$

We see that the generating function for $\{F(k)\}$ is $x/(1-x-x^2)$. Using the package, we immediately derive that

$$f(n) = \frac{1}{2}(-\overline{F}(2n)+\overline{F}(2n-1)+\overline{F}(n)^3+\overline{F}(n)\overline{F}(n-1)^2-\overline{F}(n-1)\overline{F}(n)^2),$$

where $\{\overline{F}(k)\}$ is the primitive sequence of $\{F(k)\}$. In fact, $\overline{F}(k) = F(k+1)$ and hence

$$f(n) = \frac{1}{2}(-F(2n+1) + F(2n) + F(n+1)^3 + F(n+1)F(n)^2 - F(n)F(n+1)^2).$$

Example 2. Let

$$S(z) = \sum_{k=0}^{n-1} F(k)^4 z^k,$$

where $\{F(k)\}$ is the Fibonacci sequence defined as in Example 1. Using the package, we find that

$$\begin{aligned} S(z) &= \frac{\sum_{i=0}^{4} f_i(z) z^n \overline{F}(n-1)^i \overline{F}(n)^{4-i} - z(z+1)(z^2-5z+1)}{(z-1)(z^2+3z+1)(z^2-7z+1)} \\ &= \frac{\sum_{i=0}^{4} f_i(z) z^n F(n)^i F(n+1)^{4-i} - z(z+1)(z^2-5z+1)}{(z-1)(z^2+3z+1)(z^2-7z+1)}, \end{aligned}$$

where

$$f_0(z) = z(z+1)(z^2-5z+1), \quad f_1(z) = -4z^2(z^2-3z-1), \quad f_2(z) = 6z^2(z^2-z+1),$$

and

$$f_3(z) = -4z^2(z^2+3z-1), \quad f_4(z) = z^4 + 11z^3 - 14z^2 - 5z + 1.$$

6 D-finite sequence involved

The readers are referred to [Sta99, Chapter 6.4] for definitions of D-finite generating functions and P-recursive sequence. Let $\{G(k)\}_{k\in\mathbb{N}}$ be a P-recursive sequence with D-finite generating function $g(x)$. We wish to find a similar representation of the sum

$$S = \sum_{k=0}^{n-1} \prod_{j=1}^{r+1} F_j(k)$$

with F_j as before except for $F_{r+1}(k) = G(k)$ being P-recursive. We shall only consider the case $c_r = 0$ for brevity.

Define $S_m(z)$ as in (9). The recursion (10) still holds for $m \leq r$, and a similar calculation yields

$$S_{r+1}(z) = \operatorname*{CT}_x x^{-a_{r+1}n} g(x) S_r(zx^{-b_{r+1}}).$$

By Theorem 3, we can write

$$S_r(z) = \frac{L_r(P_r(z)) - z^n \tilde{L}_r(P_r(z))}{Q_r(z)}. \tag{17}$$

Since $S_r(z)$ is a polynomial in z of degree less than n, $\deg P_r(z) < \deg Q_r(z)$ and $Q_r(0) \neq 0$.

Now put (17) into the recursion and set $z = 1$, we obtain

$$S_{r+1}(1) = \operatorname*{CT}_x x^{-a_{r+1}n} g(x) \frac{L_r(P_r(x^{-b_{r+1}})) - x^{-nb_{r+1}} \tilde{L}_r(P_r(x^{-b_{r+1}}))}{Q_r(x^{-b_{r+1}})}$$

This expression can be written as

$$S_{r+1}(1) = \operatorname*{CT}_x x^{-a_{r+1}n} \bar{g}(x) L_r(\bar{P}(x)) - x^{-n(a_{r+1}+b_{r+1})} \tilde{L}_r(\bar{P}(x)) \bar{g}(x),$$

where $\bar{g}(x) = g(x)\bar{Q}(x)^{-1}$, with

$$\frac{P_r(x^{-b_{r+1}})}{Q_r(x^{-b_{r+1}})} = \frac{\bar{P}(x)}{\bar{Q}(x)}$$

being in its standard representation.

Now if we let $\bar{G}(k) = [x^k]\bar{g}(x)$. Then we have a representation of $S_{r+1}(1)$ by a linear combination of terms of the form

$$\prod_{j=1}^{r+1} \bar{F}_j(u_j n + v_j),$$

where $\bar{F}_j(k)$ is the primitive sequence of $F_j(k)$ as before, except that $\bar{F}_{r+1}(k) = \bar{G}(k)$.

It is clear that $\bar{g}(x)$ is also D-finite and hence $\bar{G}(k)$ is P-recursive. It can be shown that if $G(k)$ satisfy a P-recursion of order e then we can find for $\bar{G}(k)$ a P-recursion of order $e + \deg \bar{Q}(x)$.

Acknowledgements

References

[APR01] G.E. Andrews, P. Paule, and A. Riese. Macmahon's partition analysis iii: the Omega package. *Europ. J. Combin.*, 22:887–904, 2001.

[GW07] C. Greene and H.S. Wilf. Closed form summation of C-finite sequences. *Trans. AMS*, 359(3):1161–1190, 2007.

[Mac16] P.A. MacMahon. *Combinatory Analysis, 2 volumes*. Cambridge University, 1915–1916. (Reprinted: Chelsea, New York, 1960).

[Sta86] R.P. Stanley. *Enumerative Combinatorics, I*. Wadsworth & Brooks/Cole, 1986. (Reprinted: Cambridge University Press, 1997).

[Sta99] R.P. Stanley. *Enumerative Combinatorics, II*. Cambridge University Press, 1999.

[Xin04] G. Xin. A fast algorithm for MacMahon's partition analysis. *Electron. J. Combin.*, 11:R58, 2004.

[Zei90] D. Zeilberger. A holonomic systems approach to special functions identities. *J. Comput. Appl. Math.*, 32:321–368, 1990.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 643–654

The Möbius function of separable permutations (extended abstract)

Vít Jelínek[1†] and Eva Jelínková[2‡] and Einar Steingrímsson[1]

[1]*The Mathematics Institute, School of Computer Science, Reykjavik University, Kringlan 1, IS-103 Reykjavik, Iceland*
[2]*Department of Applied Mathematics, Charles University, Malostranské náměstí 25, 118 00 Praha, Czech Republic*

Abstract. A permutation is separable if it can be generated from the permutation 1 by successive sums and skew sums or, equivalently, if it avoids the patterns 2413 and 3142. Using the notion of separating tree, we give a computationally efficient formula for the Möbius function of an interval (q, p) in the poset of separable permutations ordered by pattern containment. A consequence of the formula is that the Möbius function of such an interval (q, p) is bounded by the number of occurrences of q as a pattern in p. The formula also implies that for any separable permutation p the Möbius function of $(1, p)$ is either 0, 1 or -1.

Résumé. Une permutation est séparable si elle peut être générée á partir de la permutation 1 par des sommes directes et des sommes indirectes, ou de façon équivalente, si elle évite les motifs 2413 et 3142. En utilisant le concepte de l'arbre séparant, nous donnons une formule pour le calcul efficace de la fonction de Möbius d'un intervalle de (q, p) dans l'ensemble partiellement ordonné des permutations séparables. Une conséquence est que la fonction de Möbius de (q, p) pour q et p séparables est bornée par le nombre d'occurrences de q comme un motif en p. Nous montrons aussi que pour une permutation p séparable, la fonction de Möbius de $(1, p)$ est soit 0, 1 ou -1.

Keywords: Möbius function, pattern poset, separable permutations.

1 Introduction

Let $\mathcal{S}_n$ be the set of permutations of the integers $\{1, 2, \ldots, n\}$. The union of all $\mathcal{S}_n$ for $n = 1, 2, \ldots$ forms a poset, which we call $\mathcal{P}$, with respect to *pattern containment*. That is, we define $q \leq p$ in $\mathcal{P}$ if there is a subsequence of p whose letters are in the same order of size as the letters in q. For example, $132 \leq 24153$, because 2,5,3 appear in the same order of size as 132. We denote the number of *occurrences* of q in p by $q(p)$, for example $132(24153) = 3$, since 243, 253 and 153 are all the occurrences of the pattern 132 in 24153.

A classical question to ask for any combinatorially defined poset is what its Möbius function is. For our poset $\mathcal{P}$ this seems to have first been mentioned explicitly by Wilf [8]. The first result in this direction was given by Sagan and Vatter [5], who showed that an interval (q, p) of *layered* permutations is isomorphic to a certain poset of compositions of an integer, and they gave a formula for the Möbius function in this

†Jelínek and Steingrímsson were supported by grant no. 090038011 from the Icelandic Research Fund.
‡Jelínková was supported by project 1M0021620838 of the Czech Ministry of Education.

case. In this paper, we find a (computationally effective) formula for the Möbius function of an interval (q,p), where p is a separable permutation. This is a generalization of the results of Sagan and Vatter, and is based on similar principles.

Recently, Steingrímsson and Tenner [7] exhibited a class of intervals whose Möbius function is zero and described certain other intervals where the Möbius function is either 1 or -1. In addition, they conjectured that for permutations q and p avoiding the pattern 132 the absolute value of the Möbius function of the interval (q,p), denoted $\mu(q,p)$, is bounded by the number of occurrences of q in p.

With the help our formula for $\mu(q,p)$, we prove a more general version of this conjecture. We show that for any interval (q,p) of separable permutations q and p we have $\mu(q,p) \leq q(p)$ (for general p and q this inequality does not hold). In particular, if p has a single occurrence of q then $\mu(q,p)$ is either 1, 0 or -1.

We also prove a conjecture mentioned in [7], showing that for any separable permutation p, $\mu(1,p)$ is either 1, 0 or -1, where 1 in $(1,p)$ is the single permutation of length 1. In addition, we give some results on the Möbius function of various special intervals of arbitrary permutations, which we then use to conclude that $\mu(1,p)$ is bounded on certain classes of permutations. These results show how to express the Möbius function $\mu(1,p)$ for a *decomposable* permutation p in terms of $\mu(1,p_i)$, where p_i are the summands in the decomposition of p. For "most" decomposable permutations p this shows that $\mu(1,p)$ is zero.

2 Definitions and Preliminaries

An *interval* (q,p) in a poset $(\mathcal{P},\leq)$ is the set $\{r\colon q \leq r \leq p\}$. In this paper, we deal exclusively with intervals of the poset of permutations ordered by pattern containment.

The Möbius function $\mu(q,p)$ of an interval (q,p) is uniquely defined by setting $\mu(q,q) = 1$ for all q and requiring that

$$\sum_{r\in(q,p)} \mu(q,r) = 0 \tag{1}$$

for every $q < p$.

An equivalent definition is given by Philip Hall's Theorem [6, Proposition 3.8.5], which says that

$$\mu(q,p) = \sum_{C\in\mathcal{C}(q,p)} (-1)^{|C|} = \sum_i (-1)^i c_i, \tag{2}$$

where $\mathcal{C}$ is the set of chains in (q,p) that contain both q and p, and c_i is the number of such chains of length i in (q,p). For details and further information on this see [6].

The *direct sum*, $a+b$, of two nonempty permutations a and b is the permutation obtained by concatenating a and b', where b' is b with all letters incremented by the number of letters in a. A permutation that can be written as a direct sum is *decomposable*, otherwise it is *indecomposable*. Examples are $2314576 = 231 + 12 + 21$, and 231, which is indecomposable. In the *skew sum* of a and b, denoted by $a * b$, we increment the letters of a by the length of b to obtain a' and then concatenate a' and b. For example, $6743512 = 12 * 213 * 12$.

A *decomposition* of p is an expression $p = p_1 + p_2 + \cdots + p_k$ in which each summand p_i is indecomposable. The summands $p_1,\ldots,p_k$ will be called *the blocks* of p. Every permutation p has a unique decomposition (including an indecomposable permutation p, whose decomposition has a single block p).

A permutation is *separable* if it can be generated by successive sums and skew sums of the permutation 1. Being separable is equivalent to avoiding the patterns 2413 and 3142, that is, containing no occurrences of them. Separable permutations have nice algorithmic properties. For instance, Bose, Buss and Lubiw [2] have shown that it can be decided in polynomial time whether $q \leq p$ when p and q are separable, while for general permutations the problem is NP-hard.

3 Möbius function of separable permutations

Let us now consider the values of $\mu(q, p)$ for separable permutations q and p.

The recursive structure of separable permutations makes it often convenient to represent a separable permutation by a tree that describes how the permutation may be obtained from smaller permutations as a sum or skew sum. Let us formalize this concept. A *separating tree* T is a rooted tree T with the following properties:

- Each internal node of T has one of two types: it is either a *direct node* or a *skew node*.
- Each internal node has at least two children. The children of a given internal node are ordered into a sequence from left to right.

Each separating tree T represents a unique separable permutation p, defined recursively as follows:

- If T has a single vertex, it represents the singleton permutation 1.
- Assume T has more than one vertex. Let $N_1, \ldots, N_k$ be the children of the root in their left-to-right order, and let T_i denote the subtree of T rooted at the node N_i. Let $p_1, \ldots, p_k$ be the permutations represented by the trees $T_1, \ldots, T_k$. If the root of T is a direct node (skew node), then T represents the permutation $p_1 + \cdots + p_k$ ($p_1 * \cdots * p_k$, respectively).

Note that the leaves of T correspond bijectively to the elements of p. In fact, when we perform a depth-first left-to-right traversal of T, we encounter the leaves in the order that corresponds to the left-to-right order of the elements of p.

A given separable permutation may be represented by more than one separating tree. A separating tree is called *reduced tree* if it has the property that the children of a direct node are leaves or skew nodes, and the children of a skew node are leaves or direct nodes.

Each separable permutation p is represented by a unique reduced tree, denoted by $T(p)$. We assume that each leaf of T is labelled by the corresponding element of p.

The (slightly modified) concept of separating tree and its relationship with separable permutations have been previously studied in algorithmic contexts [2, 9]. We will now show that the reduced tree allows us to obtain a simple formula for the Möbius function of separable permutations.

Let $[n]$ denote the set $\{1, \ldots, n\}$. Let $p = (p(1), p(2), \ldots, p(n))$ and $q = (q(1), q(2), \ldots, q(m))$ be two permutations, with $q \leq p$. An *embedding* of q into p is a function $f\colon [m] \to [n]$ with the following two properties:

- for every $i, j \in [m]$, if $i < j$ then $f(i) < f(j)$ (i.e., f is monotone increasing).
- for every $i, j \in [m]$, if $q(i) < q(j)$, then $p(f(i)) < p(f(j))$ (i.e., f is order-preserving).

Let f be an embedding of q into p. We say that a leaf ℓ of $T(p)$ is *covered* by the embedding f if the element of p corresponding to ℓ is in the image of f. A leaf is *omitted* by f if it is not covered by f. An internal node N of $T(p)$ is *omitted* by f if all the leafs in the subtree rooted at N are omitted. A node is *maximal omitted*, if it is omitted but its parent in $T(p)$ is not omitted.

Assume that p is a separable permutation and $T(p)$ its reduced tree. Two nodes N_1 and N_2 of a tree $T(p)$ are called *twins* if they are siblings (i.e., share a common parent P), they appear consecutively in the sequence of children of P, and the two subtrees of T rooted at N_1 and N_2 are isomorphic (i.e., they only differ by the labeling of their leaves, but otherwise have the same structure). In particular, any two adjacent leaves are twins.

A *run* under a node N in T is a maximal sequence $N_1, \ldots, N_k$ of children of N such that each two consecutive elements of the sequence are twins. Note that the sequence of children of each internal node is uniquely partitioned into runs (possibly consisting of a single node). A *leaf run* is a run whose nodes are leaves, and a *non-leaf run* is a run whose nodes are non-leaves. The first (i.e., leftmost) element of each run is called *the leader* of the run, the remaining elements are called *followers*.

Using the tree structure of $T(p)$, we will show that $\mu(q,p)$ can be expressed as a signed sum over a set of particularly 'nice' embeddings of q into p. Following the terminology of Sagan and Vatter [5], we call these nice embeddings 'normal'.

Definition 1 *Let q and p be separable permutations, let $T(p)$ be the reduced tree of p. An embedding f of q into p is called* normal *if it satisfies the following two conditions.*

- *If a leaf ℓ is maximal omitted by f, then ℓ is the leader of its corresponding leaf run.*
- *If an internal node N is maximal omitted by f, then N is a follower in its non-leaf run.*

Let $N(q,p)$ denote the set of normal embeddings of q into p. The *defect* of an embedding $f \in N(q,p)$, denoted by $d(f)$ is the number of leaves that are maximal omitted by f. The *sign* of f, denoted by $\mathrm{sgn}(f)$ is defined as $(-1)^{d(f)}$.

We now present our main result.

Theorem 2 *If q and p are separable permutations, then*

$$\mu(q,p) = \sum_{f \in N(q,p)} sgn(f).$$

Consider, as an example, the two permutations p and q depicted on Figure 1. The children of the root of $T(p)$ are partitioned into three runs, where the first run has three internal nodes, the second run has a single leaf, and the last run has a single internal node. Accordingly, there are five normal embeddings of q into p, depicted in Figure 2. Of these five normal embeddings, two have sign -1 and three have sign 1, giving $\mu(q,p) = 1$.

Although the formula for $\mu(q,p)$ given in Theorem 2 may in general involve exponentially many summands, it allows us to compute $\mu(q,p)$ in time which is polynomial in $|p| + |q|$, using a simple dynamic programming approach which we outline in Subsection 3.1.

To prove Theorem 2, we show that the formula for $\mu(q,p)$ satisfies the Möbius function recurrence $\sum_{r \in (q,p)} \mu(q,r) = 0$ for each nontrivial interval (q,p) of separable permutations. To verify this recurrence, we construct a sign-reversing involution on the set $\bigcup_{r \in (q,p)} N(q,r)$. The construction of this

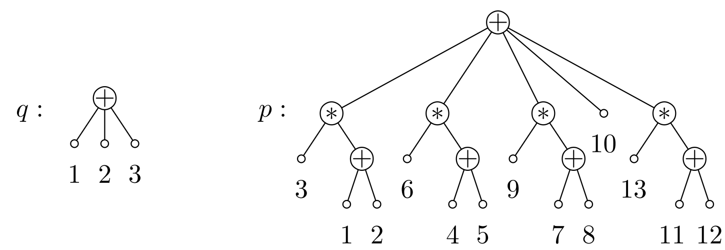


Fig. 1: The separating trees of two permutations q and p

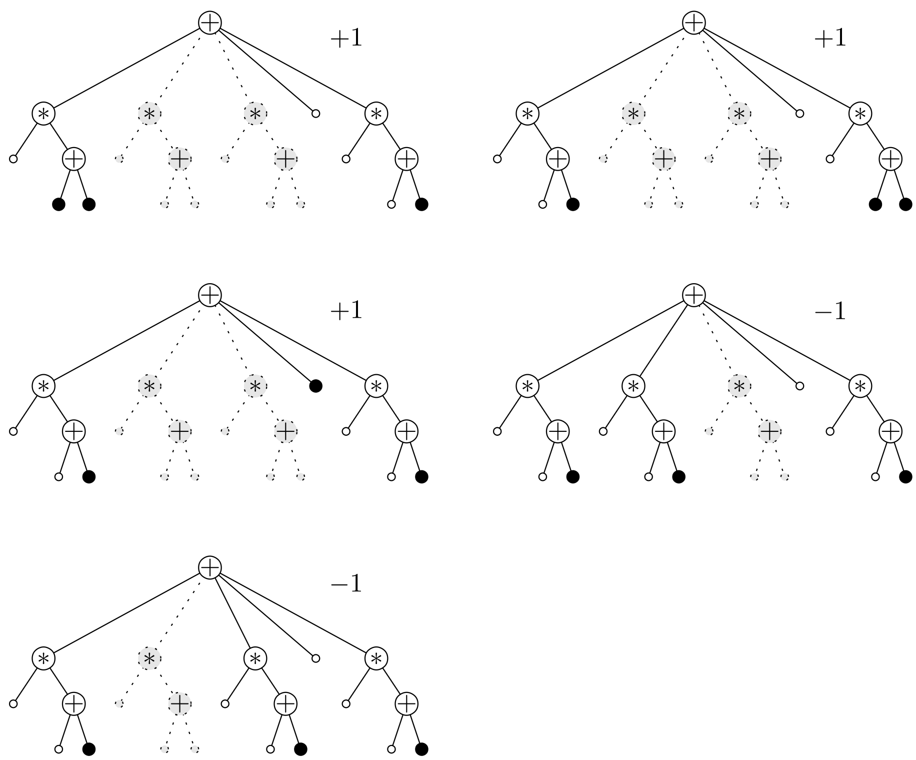


Fig. 2: The normal embeddings of q in p, together with their signs. The leaves covered by the embedding are represented by black disks, the leaves that are maximal omitted are represented by empty circles. Dotted lines represent subtrees rooted at a maximal omitted internal vertex. Note that the leaves of such subtrees do not contribute to the sign of the embedding.

bijection is very technical and relies on a complicated case analysis. We omit the construction from this extended abstract.

Let us now state several consequences of Theorem 2.

Corollary 3 *If p is separable, then $\mu(1,p) \in \{0,1,-1\}$.*

Proof: The permutation 1 can have at most one normal embedding into p. It is in fact easy to observe that if $|p| > 1$, then $T(p)$ has at least one leaf ℓ that is not a leader of its leaf run, but each of its ancestors is a leader of its non-leaf run. Such a leaf ℓ must be covered by any normal embedding of any permutation into p. □

The next Corollary confirms a (more general version of a) conjecture of Steingrímsson and Tenner [7].

Corollary 4 *If p and q are separable permutations, then $|\mu(q,p)|$ is at most the number of occurrences of q in p.*

Proof: This follows from the fact that the number of occurrences of q in p is clearly at least the number of normal embeddings of q into p. □

3.1 The algorithm

As we already pointed out, Theorem 2 allows us to compute $\mu(p,q)$ in polynomial time for any separable permutations p and q. We outline the main ideas of the algorithm.

Given p and q, we construct the trees $T(p)$ and $T(q)$. We then check all the subtrees of $T(p)$ and mark each subtree which is a follower in its non-leaf run. We call such a subtree *marked*.

In order to use dynamic programming to compute the Möbius function, we define an extension of normal embeddings of trees.

Let N be an internal node of a tree, and let $T_1,\dots,T_k$ be the subtrees rooted at the children of N. A *range* (N,i,j) is the subtree of $T(p)$ induced by $N \cup T_i \cup T_{i+1} \cup \cdots \cup T_j$. Observe that a tree has quadratically many ranges.

The leaves of each range in $T(p)$ represent a subsequence of p. We may thus speak of embeddings of a range in $T(q)$ into a range of $T(p)$.

An embedding f of a range in $T(q)$ to a range (N_p,i,j) in $T(p)$ is called *normal* if it satisfies the following conditions.

- If a leaf ℓ in (N_p,i,j) is a maximal omitted node in f, then ℓ is the leader of its corresponding leaf run in p.
- If an internal node N' in (N_p,i,j) is maximal omitted by f, then N' is marked.

We say that a normal embedding to a range (N_p,i,j) is *even* if it has an even number of maximal omitted leaves, and *odd* otherwise.

For every pair of ranges $(N_q,\ell,m) \subseteq T(q)$ and $(N_p,i,j) \subseteq T(q)$, we compute the number of odd and even normal embeddings of (N_q,ℓ,m) to (N_p,i,j), denoted by emb-odd(N_q,ℓ,m,N_p,i,j) and emb-even(N_q,ℓ,m,N_p,i,j) respectively.

For a range (N_p,i,j) with a single leaf, emb-odd(N_q,ℓ,m,N_p,i,j) and emb-even(N_q,ℓ,m,N_p,i,j) can be easily computed. To compute emb-odd(N_q,l,m,N_p,i,j) and emb-even(N_q,ℓ,m,N_p,i,j) for a range (N_p, i, j) that contains more than one leaf, we assume that we already know the values of emb-odd and emb-even for any range of $T(q)$ and for any range of $T(p)$ properly contained in (N_p, i, j). It is not difficult to see that the values of emb-odd$(N_q, \ell, m, N_p, i, j)$ and emb-even$(N_q, \ell, m, N_p, i, j)$ can then be determined in linear time from previously computed values of emb-odd and emb-even.

Having computed all the values of emb-odd and emb-even, we evaluate $\mu(q,p)$. Let R_p and R_q be the root vertices of $T(p)$ and $T(q)$, and assume that R_p has k children and R_q has j children. Then $\mu(q,p) = \text{emb-even}(R_q, 1, j, R_p, 1, k) - \text{emb-odd}(R_q, 1, j, R_p, 1, k)$.

It is clear that the algorithm works in polynomial time.

4 The Möbius function for general permutations

In this section, we demonstrate several basic properties of $\mu(p,q)$, for general (not necessarily separable) permutations p and q.

Lemma 5 *Let q be an indecomposable permutation. Let p be a permutation containing q that has the form $p = r + 1 + s$, where r, s are nonempty permutations. Then $\mu(q,p) = 0$.*

Proof: Proceed by induction on $|p|$. The smallest permutations p satisfying the assumptions of the lemma are $p = q + 1 + 1$ and $p = 1 + 1 + q$. For these permutations, the statement holds. Assume now that $|p| > |q| + 2$, and $p = r + 1 + s$ with r and s nonempty. We have

$$\mu(q,p) = -\sum_{t<p} \mu(q,t) = -\left(\sum_{t \le r+s} \mu(q,t) + \sum_{t<p,\, t \not\le r+s} \mu(q,t)\right).$$

The sum $\sum_{t \le r+s} \mu(q,t)$ is zero by the definition of μ (note that $r + s \neq q$), while every summand in $\sum_{t<p,\, t\not\le r+s} \mu(q,t)$ is zero by induction (note that both $r+1$ and $1+s$ are subpermutations of $r+s$). □

Let us say that a permutation p is *low* if p can be written as $p = 1 + r$, where r is a nonempty permutation. A permutation is *high* if it is not low. Note that the permutation 1 is high, and each indecomposable permutation is high as well.

Lemma 6 *If p and q are high permutations, then $\mu(q, 1+p) = -\mu(q,p)$.*

Proof: We know that $\mu(q, 1+p)$ can be written as

$$\mu(q, 1+p) = \sum_{C \text{ chain from } q \text{ to } 1+p} (-1)^{\ell(C)}, \tag{3}$$

where $\ell(C)$ denotes the length of the chain C.

We now distinguish two types of chains from q to p.

Type 1: A chain from q to $1+p$ is of *Type 1* if it contains the element p (which is high by assumption). Necessarily, p is the penultimate element of the chain. There is a parity-reversing bijection between all chains from q to p and the Type 1 chains: the bijection works by simply adding $1+p$ at the end of the chain from q to p. This shows that Type 1 chains contribute $-\mu(q,p)$ to the right hand side of (3).

Type 2: A chain from q to $1+p$ is of *Type 2* if it does not contain p. We claim that the contributions of Type 2 chains to (3) sum to 0. For such a chain C, let $h(C)$ be the largest high permutation appearing in the chain C. We split Type 2 chains into two groups:

Group A contains the Type 2 chains with the property that the permutation $h(C)$ is followed by the permutation $1+h(C)$ in the chain. Note that by assumption, $h(C)$ is different from p (else the chain would be of Type 1), so $1+h(C)$ is not the last element of the chain.

Group B contains the Type 2 chains where $h(C)$ is followed by a permutation r different from $1+h(C)$. The permutation r must be low, so it is of the form $1+s$, where $s \neq h(C)$. Since $r = 1+s$ contains $h(C)$, and $h(C)$ is high, it follows that $h(C)$ is in fact (properly) contained in s.

There is then a parity-reversing bijection between Group A and Group B, which works by removing the element $1+h(C)$ from a chain C. This shows that Type 2 chains contribute 0 to $\mu(q, 1+p)$. $\square$

Let us make some simple observations about embeddings. Assume that p has a decomposition $p = p_1 + p_2 + \cdots + p_k$, and assume that q is indecomposable. Let f be an embedding of q into p. Since q is indecomposable, it is easy to see that the embedding f must in fact embed q into a single block p_x of p (formally speaking, for every $i \in [m]$, we have $|p_1 + \cdots + p_{x-1}| < f(i) \leq |p_1 + \cdots + p_x|$). More generally, if q has a decomposition $q = q_1 + \cdots + q_\ell$ and f is an embedding of q into p, then each block of q is embedded by f into a single block of p.

For an integer α and a permutation q, let αq denote the sum $q + q + \cdots + q$ with α summands.

Theorem 7 *Let p be a decomposable permutation of order at least 3. Assume p has a decomposition $p_1 + p_2 + \cdots + p_k$, where neither p_1 nor p_k are equal to 1. Assume that q is an indecomposable permutation. If $\mu(q,p) \neq 0$, then all the blocks of p are equal to a single indecomposable permutation r, and in such case $\mu(q,p) = \mu(q,r)$.*

The rest of this text is devoted to the proof of Theorem 7.

The proof proceeds by induction on $|p|$. For $|p| \leq 3$, the statement holds trivially. Let us assume that p is a decomposable permutation of order at least 4, and that neither the first nor the last block of the decomposition of p is equal to 1. If any of the internal blocks of p is equal to 1, we use Lemma 5 to show that $\mu(q,p) = 0$.

In the rest of the proof, we assume that all the blocks of p have order at least 2. To compute $\mu(q,p)$, we use the expression

$$\mu(q,p) = -\sum_{t<p} \mu(q,t) \tag{4}$$

and we will show that all the terms on the right-hand side of (4) cancel out, except at most one term, whose value we will be able to determine.

We classify all the permutations $t < p$ into the following four types:

- The permutations of the form $t = 1+1+r$, where r is not empty. For any such permutation, we have $\mu(q,t) = 0$ by Lemma 5, so these permutations do not contribute anything to the right-hand side of (4).
- The permutations of the form $t = 1+r$, where r is high. We will call such permutations *extended*.
- The permutations t such that t is high and $1+t < p$. We will call such permutations *extendable*.

- The permutations t such that t is high, and $1 + t$ is not contained in p. (Note that we can never have $1 + t = p$, because we assume that all the blocks of p have order at least 2.) We will call these permutations *significant*.

Note that a permutation t is extendable if and only if $1+t$ is extended, and that this provides a bijection between extendable and extended proper subpermutations of p. Moreover, for any extendable t, we have $\mu(q, t) = -\mu(q, 1 + t)$ by Lemma 6. This shows that the contribution of extendable permutations on the right side of (4) cancels exactly with the contribution of the extended permutations. Consequently, we have

$$\mu(q,p) = -\sum_{t \text{ significant}} \mu(q,t) \tag{5}$$

Applying induction, we see that a significant permutation t that has nonzero $\mu(q, t)$ can be of one of the following two forms:

- t has the decomposition $t = \alpha r$ for some $\alpha \geq 1$ and some indecomposable permutation r of order at least 2. In such case, we know from induction hypothesis that $\mu(q, t) = \mu(q, r)$. We call such a permutation t a *significant permutation of type 1* and we say that t is an *r-permutation*. Note that we allow the possibility that $t = r$.

- t has the decomposition $t = \beta r + 1$, for some $\beta \geq 1$ and some indecomposable permutation r of order at least 2. In such case, we know from induction hypothesis and from Lemma 6 that $\mu(q, t) = -\mu(q, r)$. We call such a permutation t a *significant permutation of type 2*, and we again say that t is an *r-permutation*.

To complete the proof of Theorem 7, we plan to show that for any given indecomposable permutation $r < p$, the contribution of r-permutations of type 1 on the right side of (5) cancels with the contribution of r-permutations of type 2. The only exception to this exact cancellation will occur when the permutation p has the decomposition $p = \gamma r$ for some $\gamma \geq 2$ — in such case, there is a type-2 r-permutation $(\gamma-1)r+1$, but there is no significant type-1 r-permutation. To prove that the cancellations work as required, we first need to prove several claims.

Let us introduce some more notation. For an indecomposable permutation $r < p$, we define $\alpha_r = \max\{\alpha;\ \alpha r \leq p\}$ and $\beta_r = \max\{\beta;\ \beta r + 1 \leq p\}$. Observe that $\alpha_r - 1 \leq \beta_r \leq \alpha_r$.

Claim 8 *Let $r < p$ be an indecomposable permutation. If t is a significant r-permutation of type 1, then $t = \alpha_r r$. If t is a significant r-permutation of type 2, then $t = \beta_r r + 1$.*

Proof: Assume that t is an r-permutation of type 1. By definition, $t = \alpha r$ for some $\alpha \geq 1$. If α were greater than α_r, then t would not be a subpermutation of p. If α were smaller than α_r, then t would be extendable, because $1 + t < \alpha_r r \leq p$. Thus, if t is to be significant, it must be of the form $t = \alpha_r r$. The argument for permutations of type 2 is the same. □

Claim 8 shows that for a given r there can be at most one significant r-permutation of each type. However, for some r it might happen that even the permutations of the 'correct' form $\alpha_r r$ and $\beta_r r + 1$ are extendable, and hence not significant.

From the definition of α_r, it is clear that $\alpha_r > 0$ for any $r < p$. However, we might have $\beta_r = 0$ even when $r < p$. The next claim allows us to avoid such degenerate cases in our considerations.

Claim 9 *Let $r < p$ be an indecomposable permutation. If $\beta_r = 0$, then there is no significant r-permutation of any kind.*

Proof: Assume that $\beta_r = 0$ for some indecomposable permutation $r < p$. This implies that $\alpha_r = 1$, and that the permutation $r + 1$ is not contained in p. Recall that p has the decomposition $p_1 + \cdots + p_k$. Since r is indecomposable, r must be contained in at least one block of p.

Since $r + 1$ is not contained in p, we know that the last block p_k is the only block of p that contains r. Since p is assumed to have at least two blocks, this shows that p contains $1 + r$. This means that the permutation $\alpha_r r = r$ is extendable, hence not significant. The permutation $\beta_r r+1 = 1$ is also extendable. By Claim 8, there can be no significant r-permutations different from $\alpha_r r$ and $\beta_r r + 1$. Thus, there are no significant r-permutations. □

Before we state our next claim, we need some preparation. Consider the decomposition $p_1 + \cdots + p_k$ of p. For an integer $i \leq k$ and for an indecomposable permutation $r < p$, we define

$$\alpha_r^{(i)} = \max\{\alpha;\ \alpha r \leq p_i\}$$

and

$$\beta_r^{(i)} = \max\{\beta;\ \beta r + 1 \leq p_i\}.$$

Notice that $\alpha_r = \alpha_r^{(1)} + \alpha_r^{(2)} + \cdots + \alpha_r^{(k)}$ and $\beta_r = \alpha_r^{(1)} + \alpha_r^{(2)} + \cdots + \alpha_r^{(k-1)} + \beta_r^{(k)}$.

Claim 10 *Let $r < p$ be an indecomposable permutation. The following statements are equivalent:*

1. *The permutation $\alpha_r r$ is extendable.*
2. *The permutation $\beta_r r + 1$ is extendable.*
3. *The block p_1 contains the permutation $1 + \alpha_r^{(1)} r$.*

Proof: Let us prove that 1 implies 3. Assume that $\alpha_r r$ is extendable. This means that the permutation $t = 1 + \alpha_r r$ is contained in p. Consider an embedding f of t into p. Choose f in such a way that the value of $f(1)$ is as small as possible.

The embedding f cannot embed into any given block p_i more than $\alpha_r^{(i)}$ copies of r, because that would contradict the definition of $\alpha_r^{(i)}$. In particular, f embeds $\alpha_r^{(1)}$ copies of r inside p_1.

Note also that the leftmost element of t (which is equal to 1) is embedded inside p_1 by f. If f would embed the leftmost element of q into any other block, it would mean that f does not embed anything into p_1, and this would contradict the choice of f. We conclude that f embeds the permutation $1 + \alpha_r^{(1)} r$ inside p_1, which shows that 1 implies 3, as claimed. By an analogous argument, we may also see that 2 implies 3.

Let us now prove that 3 implies 1. Assume that $1 + \alpha_r^{(1)} r$ is contained in p_1. By definition, we also know that for each $i \geq 2$, the block p_i contains the permutation $\alpha_r^{(i)} r$. Putting this together, we see that p contains $1 + \alpha_r^{(1)} r + \alpha_r^{(2)} r + \cdots + \alpha_r^{(k)} r = 1 + \alpha_r r$. This shows that $\alpha_r r$ is extendable, and 1 holds. By the same reasoning, we see that 3 implies 2. This completes the proof of the claim. □

Let r be again an indecomposable permutation contained in p. It is clear that the permutation $\beta_r r+1$ is either significant or extendable. It is also clear that the permutation $\alpha_r r$ is either significant, or extendable, or equal to p. In view of Claim 10, for each r, exactly one of the following three cases must occur:

- The permutation $\alpha_r r$ is extendable, and so is $\beta_r r + 1$. In this case there is no significant r-permutation of any type.
- The permutation $\alpha_r r$ is significant, and so is $\beta_r r + 1$. In this case, these two permutations are exactly the only two significant r-permutations, and their contributions to the sum on the right-hand side of Equation (5) cancel out.
- The permutation $\alpha_r r$ is equal to p. In this case, it is easy to see that $\beta_r = \alpha_r - 1$, and that the permutation $\beta_r r+1$ is significant. This case arises if and only if all the blocks in the decomposition of p are equal to r. From the induction hypothesis and from Lemma 6, we know that $\mu(q, \beta_r r+1) = -\mu(q, r)$.

We conclude that if the blocks in the decomposition of p are all equal to r, then $\mu(q, p) = \mu(q, r)$. On the other hand, if the blocks of p are not all equal, then $\mu(q, p) = 0$. This completes the proof of Theorem 7.

Let us say that a class of permutations $\mathcal{C}$ is *sum-closed* if for each $p, q \in \mathcal{C}$, the class $\mathcal{C}$ also contains $p+q$. Similarly, $\mathcal{C}$ is *skew-closed* if $p, q \in \mathcal{C}$ implies $p * q \in \mathcal{C}$. For a set $\mathcal{P}$ of permutations, the *closure* of $\mathcal{P}$, denoted by cl$(\mathcal{P})$, is the smallest sum-closed and skew-closed class of permutations that contains $\mathcal{P}$. Notice that cl$(\{1\})$ is exactly the set of separable permutations.

The next corollary is an immediate consequence of Lemma 6 and Theorem 7.

Corollary 11 *Suppose that q is a permutation that is neither decomposable nor skew-decomposable. Let $\mathcal{P}$ be any set of permutations. Then*

$$\max\{|\mu(q, p)|;\ p \in \mathcal{P}\} = \max\{|\mu(q, p)|;\ p \in cl(\mathcal{P})\}.$$

Moreover, the computation of $\mu(q, p)$ for $p \in cl(\mathcal{P})$ can be efficiently reduced to the computation of the values $\mu(q, r)$ for $r \in \mathcal{P}$.

5 Concluding remarks and open problems

We have shown that the Möbius function $\mu(q, p)$ can be computed efficiently whenever p and q are separable permutations. With some additional arguments (omitted from this extended abstract), we can in fact show that the values of $\mu(q, p)$ can be computed in polynomial time whenever p belongs to a fixed class of permutations that is the closure of a finite set. We do not know whether this result can be extended to larger classes of permutations.

Bose, Buss and Lubiw [2] have shown that it is NP-hard, for given permutations p and q, to decide whether p contains q. In view of this, it seems unlikely that $\mu(q, p)$ could be computed efficiently for general permutations q and p.

Our results imply that for a separable permutation p, the Möbius function $\mu(1, p)$ has absolute value at most 1. In fact, the class of separable permutations is the largest hereditary class with this property, since any hereditary class not contained in the class of separable permutations must contain 2413 or 3142, and $\mu(1, 2413) = \mu(1, 3142) = -3$. We have also seen that $\mu(1, p)$ is bounded on any permutation class that is a closure of finitely many permutations. Is there another example of a class on which $\mu(1, p)$ is bounded?

References

[1] M. H. Albert, M. D. Atkinson, and M. Klazar, The enumeration of simple permutations, *J. Integer Seq.* **6** (2003), 03.4.4.

[2] P. Bose, J. F. Buss, A. Lubiw, Pattern-matching for permutations, *Information Processing Letters* **65** (1998), 277–283.

[3] A. Björner, The Möbius function of subword order, in *Invariant Theory and Tableaux* (Minneapolis, MN, 1988), vol. 19 of IMA Vol. Math. Appl. Springer, New York (1990) 118-124.

[4] A. Björner, The Möbius function of factor order, *Theoretical Computer Sci.* **117** (1993), 91–98.

[5] B. E. Sagan and V. Vatter, The Möbius function of a composition poset, *J. Algebraic Combin.* **24** (2006), 117–136.

[6] R. P. Stanley, *Enumerative Combinatorics, vol. 1*, Cambridge Studies in Advanced Mathematics, no. 49, Cambridge University Press, Cambridge, 1997.

[7] E. Steingrímsson, B. E. Tenner, The Möbius function of the permutation pattern poset, `arXiv:0902.4011`.

[8] H. Wilf, The patterns of permutations, *Discrete Math.* **257** (2002), 575–583.

[9] V. Yugandhar, S. Saxena, Parallel algorithms for separable permutations, *Discrete Applied Mathematics* **146** (2005), 343–364.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 655–666

Affine structures and a tableau model for E_6 crystals

Brant Jones[1†] and Anne Schilling[1‡]

[1] *Department of Mathematics, One Shields Avenue, University of California, Davis, CA 95616*

Abstract.

We provide the unique affine crystal structure for type $E_6^{(1)}$ Kirillov–Reshetikhin crystals corresponding to the multiples of fundamental weights $s\Lambda_1, s\Lambda_2$, and $s\Lambda_6$ for all $s \geq 1$ (in Bourbaki's labeling of the Dynkin nodes, where 2 is the adjoint node). Our methods introduce a generalized tableaux model for classical highest weight crystals of type E and use the order three automorphism of the affine $E_6^{(1)}$ Dynkin diagram. In addition, we provide a conjecture for the affine crystal structure of type $E_7^{(1)}$ Kirillov–Reshetikhin crystals corresponding to the adjoint node.

Résumé.

Nous donnons l'unique structure cristalline affine pour les cristaux de Kirillov–Reshetikhin de type $E_6^{(1)}$ correspondant aux multiples des poids fondamentaux $s\Lambda_1, s\Lambda_2$ et $s\Lambda_6$ pour tout $s \geq 1$ (dans l'étiquetage de Bourbaki des noeuds de Dynkin, où 2 est le noeud adjoint). Pour ceci, nous introduisons un modèle de tableaux généralisés pour les cristaux classiques du plus haut poids de type E et nous employons l'automorphisme d'ordre trois du diagramme de Dynkin du type $E_6^{(1)}$. En outre, nous fournissons une conjecture pour la structure affine pour les cristaux de Kirillov–Reshetikhin de type $E_7^{(1)}$ correspondant au noeud adjoint.

Keywords: Affine crystals, Kirillov–Reshetikhin crystals, type E_6

This document is an extended abstract of Jones and Schilling (2009). Please see the full paper for complete proofs.

1 Introduction

Let $\mathfrak{g}$ be an affine Kac–Moody algebra and $U_q'(\mathfrak{g})$ be the associated quantized affine algebra. Kirillov–Reshetikhin modules are finite dimensional $U_q'(\mathfrak{g})$-modules labeled by a node r of the Dynkin diagram together with a nonnegative integer s. It is expected that each Kirillov–Reshetikhin module has a crystal basis.

[†] BJ was partially supported by NSF grant DMS-0636297.

[‡] AS was partially supported by the NSF grants DMS–0501101, DMS–0652641, and DMS–0652652.

In this paper, we provide the unique affine crystal structure for the Kirillov–Reshetikhin crystals $B^{r,s}$ of type $E_6^{(1)}$ for the Dynkin nodes $r = 1, 2$, and 6 in the Bourbaki labeling, where node 2 corresponds to the adjoint node (see Figure 1). In addition, we provide a conjecture for the affine crystal structure for type $E_7^{(1)}$ Kirillov–Reshetikhin crystals of level s corresponding to the adjoint node.

Our construction of the affine crystals uses the classical decomposition given by Chari (2001) together with a promotion operator. Combinatorial models of all Kirillov–Reshetikhin crystals of nonexceptional types were constructed using promotion and similarity methods in Schilling (2008); Okado and Schilling (2008); Fourier et al. (2009) and perfectness was proven in Fourier et al. (2010). Affine crystals of type $E_6^{(1)}$ and $E_7^{(1)}$ of level 1 corresponding to minuscule coweights ($r = 1, 6$) have also been studied in Magyar (2006) using the Littelmann path model. Hernandez and Nakajima (2006) gave a construction of the Kirillov–Reshetihkin crystals $B^{r,1}$ for all r for type $E_6^{(1)}$ and most nodes r in type $E_7^{(1)}$.

For nonexceptional types, the classical crystals appearing in the decomposition can be described using Kashiwara–Nakashima tableaux Kashiwara and Nakashima (1994). We provide a similar construction for general types (see Theorem 2.6). This involves the explicit construction of the highest weight crystals $B(\Lambda_i)$ corresponding to fundamental weights Λ_i using the Lenart–Postnikov Lenart and Postnikov (2008) model and the notion of pairwise weakly increasing columns (see Definition 2.1).

This paper is structured as follows. In Section 2, the fundamental crystals $B(\Lambda_1)$ and $B(\Lambda_6)$ are constructed explicitly for type E_6 and it is shown that all other highest weight crystals $B(\lambda)$ of type E_6 can be constructed from these. In Section 2.4, a generalized tableaux model is given for $B(\lambda)$ for general types. These results are used to construct the affine crystals in Section 3. Our main results are stated in Theorem 3.10 and Conjecture 3.11.

2 A tableau model for finite-dimensional highest weight crystals

In this section, we describe a model for the classical highest weight crystals in type E. In Section 2.1, we introduce our notation and give the axiomatic definition of a crystal. The tensor product rule for crystals is reviewed in Section 2.2. In Section 2.3, we give an explicit construction of the highest weight crystals associated to the fundamental weights in types E_6 and E_7. In Section 2.4, we give a generalized tableaux model to realize all of the highest weight crystals in these types. The generalized tableaux are type-independent, and can be viewed as an extension of the Kashiwara–Nakashima tableaux Kashiwara and Nakashima (1994) to type E. For a general introduction to crystals we refer to Hong and Kang (2002).

2.1 Axiomatic definition of crystals

Denote by $\mathfrak{g}$ a symmetrizable Kac-Moody algebra, P the weight lattice, I the index set for the vertices of the Dynkin diagram of $\mathfrak{g}$, $\{\alpha_i \in P \mid i \in I\}$ the simple roots, and $\{\alpha_i^\vee \in P^* \mid i \in I\}$ the simple coroots. Let $U_q(\mathfrak{g})$ be the quantized universal enveloping algebra of $\mathfrak{g}$. A $U_q(\mathfrak{g})$-*crystal* Kashiwara (1995) is a nonempty set B equipped with maps $\mathrm{wt} : B \to P$ and $e_i, f_i : B \to B \cup \{\mathbf{0}\}$ for all $i \in I$, satisfying

$$
\begin{aligned}
f_i(b) = b' \Leftrightarrow e_i(b') = b \text{ if } b, b' \in B \\
\mathrm{wt}(f_i(b)) = \mathrm{wt}(b) - \alpha_i \text{ if } f_i(b) \in B \\
\langle \alpha_i^\vee, \mathrm{wt}(b) \rangle = \varphi_i(b) - \varepsilon_i(b).
\end{aligned}
$$

Here, we have

$$\varepsilon_i(b) = \max\{n \geq 0 \mid e_i^n(b) \neq \mathbf{0}\}$$
$$\varphi_i(b) = \max\{n \geq 0 \mid f_i^n(b) \neq \mathbf{0}\}$$

for $b \in B$, and we denote $\langle \alpha_i^\vee, \mathrm{wt}(b)\rangle$ by $\mathrm{wt}_i(b)$. A $U_q(\mathfrak{g})$-crystal B can be viewed as a directed edge-colored graph called the *crystal graph* whose vertices are the elements of B, with a directed edge from b to b' labeled $i \in I$, if and only if $f_i(b) = b'$. Given $i \in I$ and $b \in B$, the *i-string through b* consists of the nodes $\{f_i^m(b) : 0 \leq m \leq \varphi_i(b)\} \cup \{e_i^m(b) : 0 < m \leq \varepsilon_i(b)\}$.

Let $\{\Lambda_i \mid i \in I\}$ be the fundamental weights of $\mathfrak{g}$. For every $b \in B$ define $\varphi(b) = \sum_{i \in I} \varphi_i(b)\Lambda_i$ and $\varepsilon(b) = \sum_{i \in I} \varepsilon_i(b)\Lambda_i$. An element $b \in B$ is called *highest weight* if $e_i(b) = \mathbf{0}$ for all $i \in I$. We say that B is a *highest weight crystal* of highest weight λ if it has a unique highest weight element of weight λ. For a dominant weight λ, we let $B(\lambda)$ denote the unique highest-weight crystal with highest weight λ.

An *isomorphism* of crystals is a bijection $\Psi : B \cup \{\mathbf{0}\} \to B' \cup \{\mathbf{0}\}$ such that $\Psi(\mathbf{0}) = \mathbf{0}$, $\varepsilon(\Psi(b)) = \varepsilon(b)$, $\varphi(\Psi(b)) = \varphi(b)$, $f_i\Psi(b) = \Psi(f_i(b))$, and $\Psi(e_i(c)) = e_i\Psi(c)$ for all $b, c \in B$, $\Psi(b), \Psi(c) \in B'$ where $f_i(b) = c$.

When $\widetilde{\lambda}$ is a weight in an affine type, we call

$$\langle \widetilde{\lambda}, c\rangle = \sum_{i \in I} a_i^\vee \langle \widetilde{\lambda}, \alpha_i^\vee\rangle \tag{1}$$

the *level* of $\widetilde{\lambda}$, where c is the canonical central element and $\widetilde{\lambda} = \sum_{i \in I} \lambda_i \Lambda_i$ is the affine weight. In our work, we will often compute the 0-weight $\lambda_0\Lambda_0$ at level 0 for a node b in a classical crystal from the classical weight $\lambda = \sum_{i \in I \setminus \{0\}} \lambda_i \Lambda_i = \mathrm{wt}(b)$ by setting $\langle \lambda_0\Lambda_0 + \lambda, c\rangle = 0$ and solving for λ_0.

When $\mathfrak{g}$ is a finite-dimensional Lie algebra, every integrable $U_q(\mathfrak{g})$-module decomposes as a direct sum of highest weight modules. On the level of crystals, this implies that every crystal graph B corresponding to an integrable module is a union of connected components, and each connected component is the crystal graph of a highest weight module. We denote this by $B = \bigoplus B(\lambda)$ for some set of dominant weights λ, and we call these $B(\lambda)$ the *components* of the crystal.

Suppose that $\mathfrak{g}$ is a symmetrizable Kac–Moody algebra and let $U_q'(\mathfrak{g})$ be the corresponding quantum algebra without derivation. The goal of this work is to study crystals $B^{r,s}$ that correspond to certain finite dimensional $U_q'(\mathfrak{g})$-modules known as Kirillov–Reshetikhin modules. Here, r is a node of the Dynkin diagram and s is a nonnegative integer. The existence of the crystals $B^{r,s}$ that we study follows from results in (Kang et al., 1992, Proposition 3.4.4) for $r = 1, 6$ and (Kang et al., 1992, Proposition 3.4.5) for $r = 2$, while the classical decomposition of these crystals is given in Chari (2001).

2.2 Tensor products of crystals

Let $B_1, B_2, \ldots, B_L$ be $U_q(\mathfrak{g})$-crystals. The Cartesian product $B_1 \times B_2 \times \cdots \times B_L$ has the structure of a $U_q(\mathfrak{g})$-crystal using the so-called signature rule. The resulting crystal is denoted $B = B_1 \otimes B_2 \otimes \cdots \otimes B_L$ and its elements $(b_1, \ldots, b_L)$ are written $b_1 \otimes \cdots \otimes b_L$ where $b_j \in B_j$. The reader is warned that our convention is opposite to that of Kashiwara Kashiwara (1995). Fix $i \in I$ and $b = b_1 \otimes \cdots \otimes b_L \in B$. The *$i$-signature* of b is the word consisting of the symbols $+$ and $-$ given by

$$\underbrace{- \cdots -}_{\varphi_i(b_1)\text{ times}} \quad \underbrace{+ \cdots +}_{\varepsilon_i(b_1)\text{ times}} \cdots \underbrace{- \cdots -}_{\varphi_i(b_L)\text{ times}} \quad \underbrace{+ \cdots +}_{\varepsilon_i(b_L)\text{ times}} .$$

The *reduced i-signature* of b is the subword of the i-signature of b, given by the repeated removal of adjacent symbols $+-$ (in that order); it has the form

$$\underbrace{-\cdots-}_{\varphi_i \text{ times}} \quad \underbrace{+\cdots+}_{\varepsilon_i \text{ times}}.$$

If $\varphi_i = 0$ then $f_i(b) = \mathbf{0}$; otherwise

$$f_i(b_1 \otimes \cdots \otimes b_L) = b_1 \otimes \cdots \otimes b_{j-1} \otimes f_i(b_j) \otimes \cdots \otimes b_L$$

where the rightmost symbol $-$ in the reduced i-signature of b comes from b_j. Similarly, if $\varepsilon_i = 0$ then $e_i(b) = \mathbf{0}$; otherwise

$$e_i(b_1 \otimes \cdots \otimes b_L) = b_1 \otimes \cdots \otimes b_{j-1} \otimes e_i(b_j) \otimes \cdots \otimes b_L$$

where the leftmost symbol $+$ in the reduced i-signature of b comes from b_j. It is not hard to verify that this defines the structure of a $U_q(\mathfrak{g})$-crystal with $\varphi_i(b) = \varphi_i$ and $\varepsilon_i(b) = \varepsilon_i$ in the above notation, and weight function

$$\mathrm{wt}(b_1 \otimes \cdots \otimes b_L) = \sum_{j=1}^{L} \mathrm{wt}(b_j).$$

2.3 Fundamental crystals for type E_6 and E_7

Let $I = \{1, 2, 3, 4, 5, 6\}$ denote the classical index set for E_6. We number the nodes of the affine Dynkin diagram as in Figure 1.

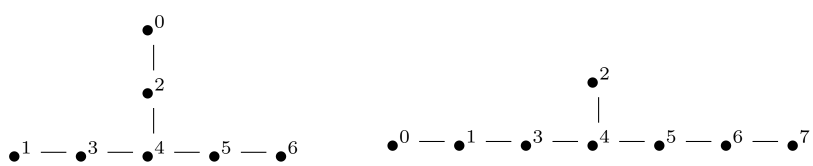


Fig. 1: Affine $E_6^{(1)}$ and $E_7^{(1)}$ Dynkin diagrams

Classical highest-weight crystals $B(\lambda)$ for E_6 can be realized by the Lenart–Postnikov alcove path model described in Lenart and Postnikov (2008). We implemented this model in `Sage` and have recorded the crystal $B(\Lambda_1)$ in Figure 2. This crystal has 27 nodes.

To describe our labeling of the nodes, observe that all of the i-strings in $B(\Lambda_1)$ have length 1 for each $i \in I$. Therefore, the crystal admits a transitive action of the Weyl group. Also, it is straightforward to verify that all of the nodes in $B(\Lambda_1)$ are determined by weight. For our work in Section 3, we also compute the 0-weight at level 0 of a node b in any classical crystal from the classical weight as described in Remark 3.4.

Thus, we label the nodes of $B(\Lambda_1)$ by weight, which is equivalent to recording which i-arrows come in and out of b. The i-arrows into b are recorded with an overline to indicate that they contribute negative weight, while the i-arrows out of b contribute positive weight.

By the symmetry of the Dynkin diagram, we have that $B(\Lambda_6)$ also has 27 nodes and is dual to $B(\Lambda_1)$ in the sense that its crystal graph is obtained from $B(\Lambda_1)$ by reversing all of the arrows. Reversing the arrows requires us to label the nodes of $B(\Lambda_6)$ by the weight that is the negative of the weight of the corresponding node in $B(\Lambda_1)$. Moreover, observe that $B(\Lambda_1)$ contains no pair of nodes with weights μ, $-\mu$, respectively. Hence, we can unambiguously label any node of $B(\Lambda_1) \cup B(\Lambda_6)$ by weight.

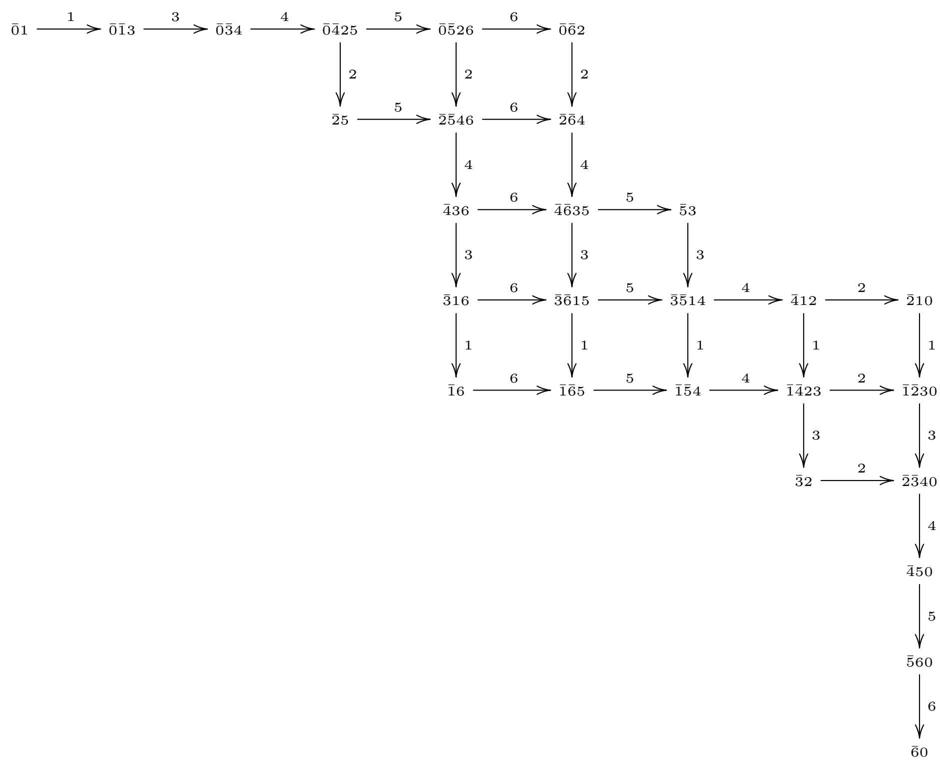


Fig. 2: Crystal graph for $B(\Lambda_1)$ of type E_6

It is straightforward to show using characters that every classical highest-weight representation $B(\Lambda_i)$ for $i \in I$ can be realized as a component of some tensor product of $B(\Lambda_1)$ and $B(\Lambda_6)$ factors. On the level of crystals, the tensor products $B(\Lambda_1)^{\otimes k}$, $B(\Lambda_6)^{\otimes k}$ and $B(\Lambda_6) \otimes B(\Lambda_1)$ are defined for all k by the tensor product rule of Section 2.2. Therefore, we can realize the other classical fundamental crystals $B(\Lambda_i)$ as shown in Table 1. There are additional realizations for these crystals obtained by dualizing.

There is a similar construction for the fundamental crystals for type E_7. The highest weight crystal $B(\Lambda_7)$ has 56 nodes and these nodes all have distinct weights. Also, $\varphi_i(b) \leq 1$ and $\varepsilon_i(b) \leq 1$ for all $i \in \{1, 2, \ldots, 7\}$ and $b \in B(\Lambda_7)$. Using character calculations, we can show that every classical highest-weight representation $B(\Lambda_i)$ appears in some tensor product of $B(\Lambda_7)$ factors.

Tab. 1: Fundamental realizations for E_6

	Generator	in	Dimension
$B(\Lambda_2)$	$2\bar{1}\bar{0} \otimes \bar{0}1$	$B(\Lambda_6) \otimes B(\Lambda_1)$	78
$B(\Lambda_3)$	$\bar{0}\bar{1}3 \otimes \bar{0}1$	$B(\Lambda_1)^{\otimes 2}$	351
$B(\Lambda_4)$	$\bar{0}\bar{3}4 \otimes \bar{0}\bar{1}3 \otimes \bar{0}1$	$B(\Lambda_1)^{\otimes 3}$	2925
$B(\Lambda_5)$	$5\bar{6}\bar{0} \otimes 6\bar{0}$	$B(\Lambda_6)^{\otimes 2}$	351

2.4 Generalized tableaux

In this section, we describe how to realize the crystal $B(\Lambda_{i_1} + \Lambda_{i_2} + \cdots + \Lambda_{i_k})$ inside the tensor product $B(\Lambda_{i_1}) \otimes B(\Lambda_{i_2}) \otimes \cdots \otimes B(\Lambda_{i_k})$, where the Λ_i are all fundamental, or more generally dominant weights. Our arguments use only abstract crystal properties, so the results in this section apply to any finite type.

If b is the unique highest weight node in $B(\lambda)$ and c is the unique highest weight node in $B(\mu)$, then $B(\lambda+\mu)$ is generated by $b \otimes c \in B(\lambda) \otimes B(\mu)$. Iterating this procedure provides a recursive description of any highest-weight crystal embedded in a tensor product of crystals. Our goal is to give a non-recursive description of the nodes of $B(\Lambda_{i_1} + \Lambda_{i_2} + \cdots + \Lambda_{i_k})$ for any collection of fundamental weights Λ_i.

For an ordered set of dominant weights $(\mu_1, \mu_2, \ldots, \mu_k)$ and for each permutation w in the symmetric group S_k, define

$$B_w(\mu_1, \ldots, \mu_k) = B(\mu_{w(1)}) \otimes B(\mu_{w(2)}) \otimes \cdots \otimes B(\mu_{w(k)})$$

so $B_e(\mu_1, \ldots, \mu_k)$ is $B(\mu_1) \otimes \cdots \otimes B(\mu_k)$ where $e \in S_k$ is the identity.

Definition 2.1 *Let $(\mu_1, \mu_2, \ldots, \mu_k)$ be dominant weights. Then, we say that*

$$b_1 \otimes b_2 \otimes \cdots \otimes b_k \in B(\mu_1) \otimes B(\mu_2) \otimes \cdots \otimes B(\mu_k)$$

is pairwise weakly increasing *if*

$$b_j \otimes b_{j+1} \in B(\mu_j + \mu_{j+1}) \subset B(\mu_j) \otimes B(\mu_{j+1})$$

for each $1 \leq j < k$.

Next, we fix an isomorphism of crystals

$$\Phi_w^{(\mu_1,\ldots,\mu_k)} : B_w(\mu_1, \ldots, \mu_k) \to B_e(\mu_1, \ldots, \mu_k)$$

for each $w \in S_k$. Observe that each choice of $\Phi_w^{(\mu_1,\ldots,\mu_k)}$ corresponds to a choice for the image of each of the highest-weight nodes in $B_w(\mu_1, \ldots, \mu_k)$.

Let b_j^* denote the unique highest weight node of the jth factor $B(\mu_j)$. Since we are fixing the dominant weights $(\mu_1, \ldots, \mu_k)$, we will sometimes drop the notation $(\mu_1, \ldots, \mu_k)$ from B_w and Φ_w.

Definition 2.2 *Let w be a permutation and choose j to be the maximal integer such that w that fixes $\{1, 2, \ldots, j\}$. We say that $\Phi_w^{(\mu_1,\ldots,\mu_k)}$ is a* lazy isomorphism *if the image of every highest weight node of the form*

$$b_1 \otimes b_2 \otimes \cdots \otimes b_j \otimes b_{j+1}^* \otimes \cdots \otimes b_k^*$$

under $\Phi_w^{(\mu_1,\ldots,\mu_k)}$ is equal to

$$b_1 \otimes b_2 \otimes \cdots \otimes b_j \otimes b^*_{w^{-1}(j+1)} \otimes \cdots \otimes b^*_{w^{-1}(k)} \,.$$

We want to choose our isomorphisms $\Phi_w^{(\mu_1,\ldots,\mu_k)}$ to be lazy, but our results do not otherwise depend upon the choice of $\Phi_w^{(\mu_1,\ldots,\mu_k)}$.

Definition 2.3 *Let T be any subset of S_k, and $\{\Phi_w^{(\mu_1,\ldots,\mu_k)}\}_{w\in T}$ be a collection of lazy isomorphisms. We define $I^{(\mu_1,\ldots,\mu_k)}(T)$ to be*

$$\bigcap_{w\in T} \Phi_w^{(\mu_1,\ldots,\mu_k)}(\{\textit{pairwise weakly increasing nodes of } B_w(\mu_1,\ldots,\mu_k)\ \}) \subset B_e(\mu_1,\ldots,\mu_k).$$

Proposition 2.4 *Let T be any subset of S_k. Then, whenever $b \in I^{(\mu_1,\ldots,\mu_k)}(T)$ we have $e_i(b), f_i(b) \in I^{(\mu_1,\ldots,\mu_k)}(T)$.*

Corollary 2.5 *For any subset T of S_k, we have that $I^{(\mu_1,\ldots,\mu_k)}(T)$ is a direct sum of highest weight crystals $\bigoplus_\lambda B(\lambda)$ for some collection of weights λ.*

Proof: Proposition 2.4 implies that whenever $b \in I^{(\mu_1,\ldots,\mu_k)}(T)$, the entire connected component of the crystal graph containing b is in $I^{(\mu_1,\ldots,\mu_k)}(T)$. □

Theorem 2.6 *Fix a sequence $(\mu_1,\ldots,\mu_k)$ of dominant weights. Then,*

$$I^{(\mu_1,\ldots,\mu_k)}(S_k) \cong B(\mu_1+\mu_2+\ldots+\mu_k).$$

Proof: Let b_j^* be the unique highest weight node of B_j with highest weight μ_j for each $j = 1,\ldots,k$. Then $b^* = b_1^* \otimes b_2^* \otimes \cdots \otimes b_k^*$ generates $B(\mu_1+\ldots+\mu_k)$ and this node lies in $I^{(\mu_1,\ldots,\mu_k)}(S_k)$.

The proof proceeds to show that b^* is the only highest weight node of $I^{(\mu_1,\ldots,\mu_k)}(S_k)$ using calculations involving the tensor product rule. □

Remark 2.7 *The condition that there is a unique highest weight element that we used in the proof of Theorem 2.6 is equivalent to the hypothesis of (Kashiwara and Nakashima, 1994, Proposition 2.2.1) from which the desired conclusion also follows.*

Remark 2.8 *Because we only require a constant amount of data to check the pairwise weakly increasing condition for each pair of tensor factors, Theorem 2.6 and its refinements will allow us to formulate arguments that apply to all highest-weight crystals simultaneously, regardless of the number of tensor factors.*

When we are considering a specific highest-weight crystal, it may be computationally easier to generate $B(\mu_1+\cdots+\mu_k)$ by simply applying f_i operations to the highest-weight node in all possible ways.

We say that any node of $I^{(\mu_1,\ldots,\mu_k)}(S_k)$ is *weakly increasing*. It turns out that we can often take T to be much smaller than S_k by starting with $T = \{e\}$ and adding permutations to T until $I^{(\mu_1,\ldots,\mu_k)}(T)$ contains a unique highest weight node. In particular, the next result shows that we can take $T = \{e\}$ when we are considering a linear combination of two distinct fundamental weights.

Lemma 2.9 *Let Λ_{i_1} and Λ_{i_2} be distinct fundamental weights, and $k_1, k_2 \in \mathbb{Z}_{\geq 0}$ with $k = k_1 + k_2$. Then, the nodes of*

$$B(k_1\Lambda_{i_1} + k_2\Lambda_{i_2}) \subset B(\Lambda_{i_1})^{\otimes k_1} \otimes B(\Lambda_{i_2})^{\otimes k_2}$$

are precisely the pairwise weakly increasing tensor products $b_1 \otimes b_2 \otimes \cdots \otimes b_k$ of $B(\Lambda_{i_1})^{\otimes k_1} \otimes B(\Lambda_{i_2})^{\otimes k_2}$.

All of the crystals in our work have classical decompositions that have been given by Chari (2001). These crystals satisfy the requirement of Lemma 2.9 that at most two fundamental weights appear. On the other hand, there exist examples showing that no ordering of the factors in $B(\Lambda_2) \otimes B(\Lambda_1) \otimes B(\Lambda_6)$ in type E_6 admits an analogous weakly increasing condition that is defined using only pairwise comparisons.

We now restrict to type E_6. Lemma 2.9 implies that we have a non-recursive description of all $B(k\Lambda_i)$ determined by the finite information in $B(2\Lambda_i)$. In the case of particular fundamental representations, we can be more specific about how to test for the weakly increasing condition.

Proposition 2.10 *We have that $b_1 \otimes b_2 \in B(2\Lambda_1) \subset B(\Lambda_1)^{\otimes 2}$ if and only if b_2 can be reached from b_1 by a sequence of f_i operations in $B(\Lambda_1)$.*

Proof: This is a finite computation on $B(2\Lambda_1)$. □

The crystal graph for $B(\Lambda_1)$ of Figure 2 can be viewed as a poset. Then Proposition 2.10 implies in particular that incomparable pairs in $B(\Lambda_1)$ are not weakly increasing.

There are 78 nodes in $B(\Lambda_2)$. We construct $B(\Lambda_2)$ as the highest weight crystal graph generated by $2\bar{1}\bar{0} \otimes \bar{0}1$ inside $B(\Lambda_6) \otimes B(\Lambda_1)$. Note that we only need to use the nodes in the "top half" of Figure 2 and their duals. There are 2430 nodes in $B(2\Lambda_2)$.

Proposition 2.11 *We have that*

$$(b_1 \otimes c_1) \otimes (b_2 \otimes c_2) \in B(2\Lambda_2) \subset (B(\Lambda_6) \otimes B(\Lambda_1))^{\otimes 2}$$

if and only if

(1) b_2 can be reached from b_1 by f_i operations in $B(\Lambda_6)$, and c_2 can be reached from c_1 by f_i operations in $B(\Lambda_1)$, and

(2) Whenever c_1 is dual to b_2, we have that there is a path of f_i operations from $(b_1 \otimes c_1)$ to $(b_2 \otimes c_2)$ of length at least 1 (so in particular, the elements are not equal) in $B(\Lambda_2)$.

Proof: This is a finite computation on $B(2\Lambda_2)$. □

3 Affine structures

In this section, we study the affine crystals of type $E_6^{(1)}$. We introduce the method of promotion to obtain a combinatorial affine crystal structure in Section 3.1 and the notion of composition graphs in Section 3.2. It is shown in Theorem 3.7 that order three twisted isomorphisms yield regular affine crystals. This is used to construct $B^{r,s}$ of type $E_6^{(1)}$ for the minuscule nodes $r = 1, 6$ and the adjoint node $r = 2$. We summarize these results in Section 3.3 along with a conjecture for $B^{1,s}$ of type $E_7^{(1)}$.

3.1 Combinatorial affine crystals and twisted isomorphisms

The following concept is fundamental to this work.

Definition 3.1 *Let $\widetilde{C}$ be an affine Dynkin diagram and C the associated finite Dynkin diagram (obtained by removing node 0) with index set I. Let $\dot{p}$ be an automorphism of $\widetilde{C}$, and B be a classical crystal of type C. We say that $\dot{p}$* induces a twisted isomorphism of crystals *if there exists a bijection of crystals $p : B \cup \{\mathbf{0}\} \to B' \cup \{\mathbf{0}\}$ satisfying*

$$p(b) = \mathbf{0} \text{ if and only if } b = \mathbf{0}, \text{ and} \tag{2}$$

$$p \circ f_i(b) = f_{\dot{p}(i)} \circ p(b) \text{ and } p \circ e_i(b) = e_{\dot{p}(i)} \circ p(b) \tag{3}$$

for all $i \in I \setminus \{\dot{p}^{-1}(0)\}$ and all $b \in B$.

We frequently abuse notation and denote B' by $p(B)$ even though the isomorphism $p : B \to p(B)$ may not be unique.

If we are given two classical crystals B and B', and there exists a Dynkin diagram automorphism $\dot{p}$ that induces a twisted isomorphism between B and B', then we say that B and B' are twisted-isomorphic.

Definition 3.2 *Let B be a directed graph with edges labeled by I. Then B is called* regular *if for any 2-subset $J \subset I$, we have that the restriction of B to its J-arrows is a classical rank two crystal.*

Definition 3.3 *Let B be a classical crystal with index set I. Suppose $\widetilde{B}$ is a labeled directed graph on the same nodes as B and with the same I-arrows, but with an additional set of 0-arrows. If $\widetilde{B}$ is regular with respect to $I \cup \{0\}$, then we say that $\widetilde{B}$ is a* combinatorial affine structure *for B.*

Remark 3.4 *Although we do not assume that $\widetilde{B}$ is a crystal graph for a $U_q'(\mathfrak{g})$-module, Kashiwara (2002, 2005) has shown that the crystals of such modules must be regular and have weights at level 0. Therefore, we compute the 0-weight $\lambda_0\Lambda_0$ of the nodes b in a classical crystal from the classical weight $\lambda = \sum_{i\in I} \lambda_i\Lambda_i = \mathrm{wt}(b)$ using the formula given in Equation (1) (recall that I in this section is the index set of the Dynkin diagram without 0).*

Remark 3.5 *Here are some consequences of Definitions 3.1 and 3.3.*

(1) Any crystal $p(B)$ induced by $\dot{p}$ is just a classical crystal that is isomorphic to B up to relabeling. In particular, any graph automorphism $\dot{p}$ induces at least one twisted isomorphism p: If we view B as an edge-labeled directed graph, the image of p is given on the same nodes as B by relabeling all of the arrows according to $\dot{p}$. On the other hand, it is important to emphasize that there is no canonical labeling for the nodes of $p(B)$. Also, some crystal graphs may have additional symmetry which lead to multiple twisted isomorphisms of crystals associated with a single graph automorphism $\dot{p}$.

(2) For $b \in B$, we have $\varphi(p(b)) = \sum_{i\in I} \varphi_{\dot{p}^{-1}(i)}(b)\Lambda_i$ and $\varepsilon(p(b)) = \sum_{i\in I} \varepsilon_{\dot{p}^{-1}(i)}(b)\Lambda_i$. In addition, we can compute the 0-weight of any node in B by Remark 3.4. Therefore, $\dot{p}$ permutes all of the affine weights, in the sense that

$$\mathrm{wt}_i(b) = \mathrm{wt}_{\dot{p}(i)}(p(b)) \quad \text{for all } b \in B \text{ and } i \in I \cup \{0\} .$$

(3) Since the node $\dot{p}(0)$ becomes the affine node in $p(B)$, it is sometimes possible to define a combinatorial affine structure for B "by promotion." Namely, we define f_0 on B to be $p^{-1} \circ f_{\dot{p}(0)} \circ p$. Note that in order for this to succeed, we must take the additional step of identifying the image $p(B)$ with a canonically labeled classical crystal so that we can infer the $f_{\dot{p}(0)}$ edges.

Example 3.6 *The E_6 Dynkin diagram automorphism of order two that interchanges nodes 1 and 6 induces the dual map between $B(\Lambda_1)$ and $B(\Lambda_6)$.*

The Dynkin diagram of $E_6^{(1)}$ has an automorphism of order three that we can use to construct combinatorial affine structures by promotion.

Theorem 3.7 *Let B be a classical E_6 crystal. Suppose there exists a bijection $p : B \to B$ that is a twisted isomorphism satisfying $p \circ f_1 = f_6 \circ p$, and suppose that p has order three. Then, there exists a combinatorial affine structure on B. This structure is given by defining f_0 to be $p^2 \circ f_1 \circ p$.*

Proof: If we apply p on the left and right of $pf_1 = f_6p$, we obtain $ppf_1p = pf_6pp$. Since p has order three, this is

$$p^{-1} f_1 p = p f_6 p^{-1}. \tag{4}$$

Because p is a bijection on B, we may define 0-arrows on B by the map $p^{-1}f_1p$. By the hypotheses, p must be induced by the unique Dynkin diagram automorphism $\dot{p}$ of order three that sends node 0 to 1.

To verify that this affine structure satisfies Definition 3.3, we need to check that restricting B to $\{0, i\}$-arrows is a crystal for all $i \in I$. Each of these restrictions corresponds to a rank 2 classical crystal, and Stembridge has given local rules in Stembridge (2003) that characterize such classical crystals in simply laced types. These rules depend only on calculations involving $\varphi_i(b)$ and $\varepsilon_i(b)$ at each node $b \in B$, and these quantities are preserved by twisted isomorphism.

Hence, we obtain a combinatorial affine structure for B. □

From now on, we use the notation p to denote a twisted isomorphism induced by $\dot{p}$ sending

$$0 \mapsto 1 \mapsto 6 \mapsto 0, 2 \mapsto 3 \mapsto 5 \mapsto 2, 4 \mapsto 4.$$

Also, we let $\dot{p}$ act on the affine weight lattice as in Remark 3.5(2).

3.2 Composition graphs

Let $I = \{1, 2, \dots, 6\}$ be the index set for the Dynkin diagram of E_6, and $\widetilde{I} = I \cup \{0\}$ be the index set of $E_6^{(1)}$. Suppose $J \subset I$. Consider a classical crystal B of the form $\bigoplus B(k\Lambda)$ where Λ is a fundamental weight and we sum over some collection of nonnegative integers k. Let $H^J(B)$ denote the $(I \setminus J)$-highest weight nodes of B. We will study affine crystals with B as underlying classical crystal. For a given such affine crystal, let $H^{J;0}(B)$ be the $(\widetilde{I} \setminus J)$-highest weight nodes. Using the level 0 hypothesis of Remark 3.4, we can prove properties of $H^{J;0}(B)$ for any given affine crystal with B as underlying classical crystal.

Our general strategy to define a twisted isomorphism p on a classical crystal B is to first define p on $H^J(B)$, and then extend this definition to the rest of B using Equation (3). To accomplish this, we introduce the following model for the nodes in $H^J(B)$ and $H^{J;0}(B)$.

Definition 3.8 *Fix $J \subset I$ and form directed graphs G_J and $G_{J;0}$ as follows.*

We construct the vertices of G_J and $G_{J;0}$ iteratively, beginning with all of the $(I \setminus J)$-highest weight nodes of $B(\Lambda)$. Then, we add all of the vertices $b \in B(\Lambda)$ such that

$$\{i \in I : \varepsilon_i(b) > 0\} \subset J \cup \{i \in I : \text{there exists } b' \in G_J \text{ with } b \otimes b' \text{ pairwise weakly increasing and } \varphi_i(b') > 0\}$$

to G_J. Moreover, if b also satisfies the property that there exists $b' \in G_{J;0}$ with $b \otimes b'$ pairwise weakly increasing and $\mathrm{wt}_0(b') > 0$ whenever $\mathrm{wt}_0(b) < 0$, then we add b to $G_{J;0}$. We repeat this construction until no new vertices are added. This process eventually terminates since $B(\Lambda)$ is finite.

The edges of G_J and $G_{J;0}$ are determined by the pairwise weakly increasing condition described in Definition 2.1. Note that some nodes may have loops. We call G_J and $G_{J;0}$ the complete composition graph *for J and $J;0$, respectively.*

Lemma 3.9 *Every element of $H^J(B)$ and $H^{J;0}(B)$ is a pairwise weakly increasing tensor product of vertices that form a directed path in G_J, respectively $G_{J;0}$, where the element in $B(0) \subset H^J(B)$ is identified with the empty tensor product.*

3.3 Further results

Using composition graphs and the tableau model, we are able to prove the following result which gives an affine structure for the Kirillov–Reshetikhin crystal $B^{2,s}$.

Theorem 3.10 *There exists a unique twisted isomorphism $p : \bigoplus_{k=0}^{s} B(k\Lambda_2) \to \bigoplus_{k=0}^{s} B(k\Lambda_2)$ of order three. This isomorphism sends an $I \setminus \{6\}$-highest weight node b from component k to the unique $I \setminus \{1\}$-highest weight node b' in component $(s-k) + (\mathrm{wt}_2(b) + \mathrm{wt}_3(b) + \mathrm{wt}_5(b))$ satisfying $\mathrm{wt}_{\dot{p}(i)}(b') = \mathrm{wt}_i(b)$ for each $i \in \{2,3,5\}$.*

We also obtain analogous results for $B^{1,s}$ and $B^{6,s}$. Furthermore, we provide a conjecture for the adjoint crystal $B^{1,s}$ in type $E_7^{(1)}$.

Conjecture 3.11 *Define $p : \bigoplus_{k=0}^{s} B(k\Lambda_1) \to \bigoplus_{k=0}^{s} B(k\Lambda_1)$ on the $I \setminus \{7\}$-highest weight nodes by sending $b \in B(k\Lambda_1)$ to the unique $I \setminus \{7\}$-highest weight node b' in component $(s-k) + (\mathrm{wt}_1(b) + \mathrm{wt}_2(b) + \mathrm{wt}_6(b))$ satisfying $\mathrm{wt}_{\dot{p}(i)}(b') = \mathrm{wt}_i(b)$ for each $i \in \{1,2,6\}$.*

Let $f_0 = p \circ f_7 \circ p$. Then f_0 commutes with f_7 so we obtain a combinatorial affine structure on $\bigoplus_{k=0}^{s} B(k\Lambda_1)$, which is isomorphic to $B^{1,s}$ of type $E_7^{(1)}$.

We have verified this conjecture for $s \leq 2$.

Acknowledgements

We thank Daniel Bump, Jesus DeLoera, Matthias Koeppe and Masato Okado for helpful conversations about this work.

This research was driven by computer exploration using the open-source mathematical software `Sage` and its algebraic combinatorics features developed by the `Sage-Combinat` community; we are grateful to Nicolas M. Thiéry for all his support.

References

V. Chari. On the fermionic formula and the Kirillov-Reshetikhin conjecture. *Internat. Math. Res. Notices*, (12):629–654, 2001. ISSN 1073-7928.

G. Fourier, M. Okado, and A. Schilling. Kirillov-Reshetikhin crystals for nonexceptional types. *Advances in Mathematics*, 222:1080–1116, 2009.

G. Fourier, M. Okado, and A. Schilling. Perfectness of Kirillov–Reshetikhin crystals for nonexceptional types. *Contemp. Math.*, 506:127–143, 2010.

D. Hernandez and H. Nakajima. Level 0 monomial crystals. *Nagoya Math. J.*, 184:85–153, 2006. ISSN 0027-7630. URL `http://projecteuclid.org/getRecord?id=euclid.nmj/1167159343`.

J. Hong and S.-J. Kang. *Introduction to quantum groups and crystal bases*, volume 42 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002. ISBN 0-8218-2874-6.

B. Jones and A. Schilling. Affine structures and a tableau model for E_6 crystals. `arXiv:0909.2442`, 2009.

S.-J. Kang, M. Kashiwara, K. C. Misra, T. Miwa, T. Nakashima, and A. Nakayashiki. Perfect crystals of quantum affine Lie algebras. *Duke Math. J.*, 68(3):499–607, 1992. ISSN 0012-7094.

M. Kashiwara. On crystal bases. In *Representations of groups (Banff, AB, 1994)*, volume 16 of *CMS Conf. Proc.*, pages 155–197. Amer. Math. Soc., Providence, RI, 1995.

M. Kashiwara. On level-zero representations of quantized affine algebras. *Duke Math. J.*, 112(1):117–175, 2002. ISSN 0012-7094.

M. Kashiwara. Level zero fundamental representations over quantized affine algebras and Demazure modules. *Publ. Res. Inst. Math. Sci.*, 41(1):223–250, 2005. ISSN 0034-5318.

M. Kashiwara and T. Nakashima. Crystal graphs for representations of the q-analogue of classical Lie algebras. *J. Algebra*, 165(2):295–345, 1994. ISSN 0021-8693.

C. Lenart and A. Postnikov. A combinatorial model for crystals of Kac-Moody algebras. *Trans. Amer. Math. Soc.*, 360(8):4349–4381, 2008. ISSN 0002-9947.

P. Magyar. Littelmann paths for the basic representation of an affine Lie algebra. *J. Algebra*, 305(2): 1037–1054, 2006. ISSN 0021-8693.

M. Okado and A. Schilling. Existence of Kirillov-Reshetikhin crystals for nonexceptional types. *Represent. Theory*, 12:186–207, 2008. ISSN 1088-4165.

A. Schilling. Combinatorial structure of Kirillov-Reshetikhin crystals of type $D_n^{(1)}, B_n^{(1)}, A_{2n-1}^{(2)}$. *J. Algebra*, 319(7):2938–2962, 2008. ISSN 0021-8693.

J. R. Stembridge. A local characterization of simply-laced crystals. *Trans. Amer. Math. Soc.*, 355(12): 4807–4823 (electronic), 2003. ISSN 0002-9947.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 667–677

Denominator formulas for Lie superalgebras (extended abstract)

Victor Kac[1] and Pierluigi Möseneder Frajria[2] and Paolo Papi[3]

[1]*Department of Mathematics, Rm 2-178, MIT, 77 Mass. Ave, Cambridge, MA 02139;*
[2]*Politecnico di Milano, Polo regionale di Como, Via Valleggio 11, 22100 Como, ITALY;*
[3]*Dipartimento di Matematica, Sapienza Università di Roma, P.le A. Moro 2, 00185, Roma , ITALY.*

Abstract. We provide formulas for the Weyl-Kac denominator and superdenominator of a basic classical Lie superalgebra for a distinguished set of positive roots.

Résumé. Nous donnons les formules pour les dénominateurs et super-dénominateurs de Weyl-Kac d'une superalgèbre de Lie basique classique pour un ensemble distingué de racines positives.

Keywords: Lie superalgebra, denominator identity, dual pair

1 Introduction

The Weyl denominator identity

$$\prod_{\alpha\in\Delta^+}(1-e^{-\alpha})=\sum_{w\in W} sgn(w)\, e^{w(\rho)-\rho} \tag{1.1}$$

is one of the most intriguing combinatorial identities in the character ring of a complex finite dimensional simple Lie algebra. It admits far reaching generalizations to the Kac-Moody setting, where it provides a proof for the Macdonald's identities (including, as easiest cases, the Jacobi triple and quintuple product identities). Its role in representation theory is well-understood, since the inverse of the l.h.s of (1.1) is the character of the Verma module $M(0)$ with highest weight 0.

The goal of the present paper is to provide an expression of the character $M(0)$ in the case of a basic classical Lie superalgebra; the analog of the l.h.s of (1.1) is the Weyl-Kac denominator [6]

$$R=\frac{\prod_{\alpha\in\Delta_0^+}(1-e^{-\alpha})}{\prod_{\alpha\in\Delta_1^+}(1+e^{-\alpha})}. \tag{1.2}$$

Here and in the remaining part of the Introduction we refer the reader to Section 2 for undefined notation. Generalizations of formulas for R to affine superalgebras and their connection with number theory and the theory of special functions are thoroughly discussed in [7]. The striking differences which make the super case very different from the purely even one are the following. First, it is no more true that the sets of positive roots are conjugate under the Weyl group (to get transitivity on the set of set of positive rootss one has to consider Serganova's odd reflections, which however play no role in this paper). In particular, the denominator identity looks very different according to the chosen set of positive roots. Moreover the restriction of the supersymmetric nondegenerate invariant bilinear form to the real span of roots may be indefinite, hence isotropic sets of roots appear naturally. Indeed, one defines the *defect* d of $\mathfrak{g}$ (notation $\operatorname{def}\mathfrak{g}$) as the dimension of a maximal isotropic subspace of $\sum_{\alpha\in\Delta}\mathbb{R}\alpha$. It is shown in [7] that d equals the cardinality of a *maximal isotropic subset of* Δ^+ (a subset $S\subset\Delta^+$ is isotropic if it is formed by linearly independent pairwise orthogonal isotropic roots).

Definition 1.1 *We call a set of positive roots distinguished if the corresponding set of simple roots has exactly one odd root.*

Distinguished sets of positive roots exist for any basic classical Lie superalgebra; they are implicitly classified in [5]. The main result of the paper is the following theorem.

Theorem 1.1 *Let $\mathfrak{g}=\mathfrak{g}_0\oplus\mathfrak{g}_1$ be a basic classical Lie superalgebra of defect d, where $\mathfrak{g}=A(d-1,d-1)$ is replaced by $gl(d,d)$. Then, for any distinguished set of positive roots, we have*

$$e^\rho R=\frac{1}{C}\sum_{w\in W}sgn(w)w\frac{e^\rho}{(1+e^{-\gamma_1})(1-e^{-\gamma_1-\gamma_2})\cdots(1+(-1)^{n+1}e^{-\gamma_1-\gamma_2-\ldots-\gamma_d})},\tag{1.3}$$

$$e^\rho \check{R}=\frac{1}{C}\sum_{w\in W}sgn'(w)w\frac{e^\rho}{(1-e^{-\gamma_1})(1-e^{-\gamma_1-\gamma_2})\cdots(1-e^{-\gamma_1-\gamma_2-\ldots-\gamma_d})}\tag{1.4}$$

where W is the Weyl group of $\mathfrak{g}$, $\{\gamma_1,\ldots,\gamma_d\}$ is an explicitly defined maximal isotropic subset of Δ^+ and C is the following constant:

$$C=\begin{cases}1 & \text{if } \mathfrak{g}=A(n,m),\\ 2^{\min\{m,n\}} & \text{if } \mathfrak{g}=B(m,n),\\ 2^n & \text{if } \mathfrak{g}=D(m,n),\ m>n,\\ 2^{m-1} & \text{if } \mathfrak{g}=D(m,n),\ n\geq m,\\ 2 & \text{if } \mathfrak{g}=D(2,1,\alpha),\ F(4),\ G(3).\end{cases}\tag{1.5}$$

A suitable modification of the previous statement holds for $\mathfrak{g}$ of type $A(d-1,d-1)$ too: see Remark 3.1. The elements γ_i are defined in (3.11), (4.9) for types $A,\ B$, respectively.

Theorem 1.1 has been proved by Kac and Wakimoto in the defect 1 case [7] (see Theorem 3.1 below), so we are reduced to discuss the cases in which the defect of $\mathfrak{g}$ is greater than 1. Hence we have to deal with superalgebras of type $A(m,n),\ B(m,n),\ D(m,n)$. Our approach to these cases is based on the analysis of the $\mathfrak{g}_0$-module structure of the oscillator representation of the Weyl algebra $W(\mathfrak{g}_1)$ of $\mathfrak{g}_1$. This is done in Sections (3) and (4) relying on methods coming from the theory of Lie groups. More precisely we use Howe theory of dual pairs, and results of Kashiwara-Vergne and Li-Paul-Than-Zhu which provide explictly the Theta correspondence. The key result in this respect is Theorem 4.1. Proofs are only outlined and sometimes skipped (mainly in the case of purely representation-theoretical results). We work out in some detail the case $A(m,n)$, where a simpler treatment using Cauchy formulas in place of Howe duality is available: see Section 3.1. We also explain our general approach in type $B(m,n)$, providing a complete proof when $m \geq n$ and presenting an example to give the flavour of the general case when $m < n$. Type D is not treated here at all. Finally, in Section 5 we reformulate our main Theorem in a form which leads to a general conjecture for the expression of R, $\check{R}$ for *any* set of positive roots, involving certain maximal isotropic subsets S of positive roots. A purely combinatorial proof of this conjecture for some special choices of S will appear in a forthcoming publication, publication, where we actually derive the theta correspondence of [8] and [9] from the denominator identity.

2 Setup

In this Subsection we collect some notation and definitions which will be constantly used throughout the paper. Let $\mathfrak{g}$ be a basic classical Lie superalgebra. This means that $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1$ is a finite dimensional simple Lie superalgebra such that $\mathfrak{g}_0$ is a reductive Lie algebra and that $\mathfrak{g}$ admits a nondegenerate invariant supersymmetric bilinear form $(\cdot,\cdot)$ [5].

Recall that for a Lie superalgebra $\mathfrak{g}$ the Casimir operator is defined as $\Omega_{\mathfrak{g}} = \sum_i x^i x_i$ if $\{x_i\}$ is a basis of $\mathfrak{g}$ and $\{x^i\}$ its dual basis w.r.t. $(\cdot,\cdot)$ (see [5, pag. 85]).Then $\Omega_{\mathfrak{g}}$ acts on $\mathfrak{g}$ as $2gI_{\mathfrak{g}}$, where g is a constant. Choose a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}_0$, and let $\Delta, \Delta_0, \Delta_1$ be the set of roots, even roots, odd roots, respectively. Let $W \subset GL(\mathfrak{h}^*)$ be the group generated by the reflections w.r.t. even roots. Choose a set of positive roots $\Delta^+ \subset \Delta$ and set $\Delta_i^+ = \Delta_i \cap \Delta^+$, $i = 0, 1$. Set also, as usual, for $i = 0, 1$, $\rho_i = \frac{1}{2}\sum_{\alpha \in \Delta_i^+} \alpha$, $\rho = \rho_0 - \rho_1$. Assume that $g \neq 0$. Then set $\Delta_0^\sharp = \{\alpha \in \Delta_0 \mid g \cdot (\alpha,\alpha) > 0\}$ and let $W^\sharp$ be the subgroup of W generated by the reflections in roots from $\Delta_0^\sharp$. We refer to [7, Remark 1.1, b)] for the definition of $W^\sharp$ when $g = 0$. Set

$$\overline{\Delta}_0 = \{\alpha \in \Delta_0 \mid \tfrac{1}{2}\alpha \notin \Delta\}, \qquad \overline{\Delta}_1 = \{\alpha \in \Delta_1 \mid (\alpha,\alpha) = 0\}. \tag{2.1}$$

Finally, for $w \in W$, set

$$sgn(w) = (-1)^{\ell(w)}, \qquad sgn'(w) = (-1)^m \tag{2.2}$$

where ℓ is the usual length function on W and m is the number of reflections from $\overline{\Delta}_0^+$ occurring in an expression of w.

Beyond the Weyl denominator R defined in (1.2) it will be very important for us the Weyl-Kac superdenominator, defined as

$$\check{R} = \frac{\prod_{\alpha \in \Delta_0^+}(1 - e^{-\alpha})}{\prod_{\alpha \in \Delta_1^+}(1 - e^{-\alpha})}. \tag{2.3}$$

As a notational convention, we denote by $L^X(\mu)$ the irreducible highest weight module of highest weight μ for a Lie algebra of type X.

3 Denominator formulas for distinguished set of positive roots

Kac and Wakimoto provided an expression for R, $\check{R}$ for certain systems of positive roots.

Theorem 3.1 *Let $\mathfrak{g}$ be a classical Lie superalgebras and let Δ^+ be any set of positive roots such that a maximal isotropic subset S of Δ^+ is contained in the set of simple roots Π corresponding to Δ^+. Then*

$$e^\rho R = \sum_{w \in W^\sharp} sgn(w) w \frac{e^\rho}{\prod_{\beta \in S}(1 + e^{-\beta})}, \tag{3.1}$$

$$e^\rho \check{R} = \sum_{w \in W^\sharp} sgn'(w) w \frac{e^\rho}{\prod_{\beta \in S}(1 - e^{-\beta})}. \tag{3.2}$$

This result has been stated in [7], and fully proved if $|S| = 1$ by using representation theoretical methods. A complete combinatorial proof has recently been obtained by Gorelik [3]. Note that a distinguished set of positive roots verifies the hypothesis of Theorem 3.1 if and only if $\operatorname{def} \mathfrak{g} = 1$ (i.e., $|S| = 1$).

The choice of a set of positive roots Δ^+determines a polarization $\mathfrak{g}_1 = \mathfrak{g}_1^+ + \mathfrak{g}_1^-$, where $\mathfrak{g}_1^\pm = \bigoplus_{\alpha \in \Delta_1^\pm} \mathfrak{g}_\alpha$. Hence we can consider the Weyl algebra $W(\mathfrak{g}_1)$ of $(\mathfrak{g}_1, (\ ,\ )_{|\mathfrak{g}_1})$ and construct the $W(\mathfrak{g}_1)$-module

$$M^{\Delta^+}(\mathfrak{g}_1) = W(\mathfrak{g}_1)/W(\mathfrak{g}_1)\mathfrak{g}_1^+, \tag{3.3}$$

with action by left multiplication. The module $M^{\Delta^+}(\mathfrak{g}_1)$ is also a $sp(\mathfrak{g}_1, (\ ,\ ))$–module with $T \in sp(\mathfrak{g}_1, (\ ,\ ))$ acting by left multiplication by

$$\theta(T) = -\frac{1}{2} \sum_{i=1}^{\dim \mathfrak{g}_1} T(x_i) x^i, \tag{3.4}$$

where $\{x_i\}$ is any basis of $\mathfrak{g}_1$ and $\{x^i\}$ is its dual basis w.r.t. $(\ ,\ )$. It is easy to check that, in $W(\mathfrak{g}_1)$, relation

$$[\theta(T), x] = T(x) \tag{3.5}$$

holds for any $x \in \mathfrak{g}_1$. This implies that we have a $\mathfrak{h}$-module isomorphism

$$M^{\Delta^+}(\mathfrak{g}_1) \cong S(\mathfrak{g}_1^-) \otimes \mathbb{C}_{-\rho_1} \tag{3.6}$$

where ρ_1 is the half sum of positive odd roots and $S(\mathfrak{g}_1^-)$ is the symmetric algebra of $\mathfrak{g}_1^-$. Hence its $\mathfrak{h}$-character is given by

$$chM^{\Delta^+}(\mathfrak{g}_1) = \frac{e^{-\rho_1}}{\prod_{\alpha \in \Delta_1^+}(1 - e^{-\alpha})}. \tag{3.7}$$

The key of our approach to the denominator formula is the following observation: since $ad(\mathfrak{g}_0) \subset sp(\mathfrak{g}_1, (\ ,\ ))$, we obtain an action of $\mathfrak{g}_0$ on $M^{\Delta^+}(\mathfrak{g}_1)$. Upon multiplication by $e^{\rho_0} \prod_{\alpha \in \Delta_0^+}(1 - e^{-\alpha})$ the r.h.s. of (3.7) becomes $e^\rho \check{R}$ and equating it with the $\mathfrak{g}_0$-character of $M^{\Delta^+}(\mathfrak{g}_1)$ one obtains our formula.

Our approach to the calculation of the $\mathfrak{g}_0$-character of $M^{\Delta^+}(\mathfrak{g}_1)$ is outlined in Section 4. Next we deal the special case of type I Lie superalgebras (cf. [5]).

3.1 Type I superalgebras

The key of our approach is the following fact, which is easily proved.

Lemma 3.2 *Let $\mathfrak{g}$ be a type I basic classical Lie superalgebra and let Δ^+ be a distinguished set of positive roots. Then $\mathfrak{g}_1^+$ and $\mathfrak{g}_1^-$ are $\mathfrak{g}_0$-modules.*

Corollary 3.3 *For type I superalgebras, we have*

$$M^{\Delta^+}(\mathfrak{g}_1) \cong S(\mathfrak{g}_1^-) \otimes \mathbb{C}_{-\rho_1} \tag{3.8}$$

as $\mathfrak{g}_0$-modules.

The previous corollary reduces the problem of computing the $\mathfrak{g}_0$-character of $M^{\Delta^+}(\mathfrak{g}_1)$ to the calculation of the $\mathfrak{g}_0$-character of $S(\mathfrak{g}_1^-)$. This latter character is well-known: for a uniform approach one might e.g. refer to the work of Schmid [10].

We start discussing the denominator formula in type $A(m,n), m \neq n$. Introduce the following notation: $\mathfrak{h}$ is the set of diagonal matrices in $gl(m+1|n+1)$ with zero supertrace, $\{\epsilon_i\}$ is the standard basis of $(\mathbb{C}^{m+n+2})^*$ and $\delta_i = \epsilon_{m+i+1}$, $1 \leq i \leq n+1$.

It follows from the analysis made in [5, 2.5.4] that in this case there are two distinguished sets of positive roots up to W-action: if we fix $\Delta_0^+ = \{\epsilon_i - \epsilon_j \mid 1 \leq i < j \leq m+1\} \cup \{\delta_i - \delta_j \mid 1 \leq i < j \leq n+1\}$, and we set $\Delta_1^+ = \{\epsilon_i - \delta_j \mid 1 \leq i \leq m+1, 1 \leq j \leq n+1\}$, then the only distinguished sets of positive roots containing Δ_0^+ are $\Delta_0^+ \cup \Delta_1^+$ and $\Delta_0^+ \cup -\Delta_1^+$. The arguments which follow clearly hold for both systems, hence we deal only with $\Delta_0^+ \cup \Delta_1^+$ which we denote by Δ_A^+ (or just by Δ^+). Its corresponding set of simple roots is $\Pi = \{\epsilon_1 - \epsilon_2, \epsilon_2 - \epsilon_3, \ldots, \epsilon_{m+1} - \delta_1, \delta_1 - \delta_2, \ldots, \delta_n - \delta_{n+1}\}$.

By Corollary 3.3 we have to calculate the $\mathfrak{g}_0$-character of $S(\mathfrak{g}_1^-)$. Note that, according to our identifications, the action of $\mathfrak{g}_0$ on $\mathfrak{g}_1^-$ is the natural action of $\{(A,B) \in gl(n+1) \times gl(m+1) \mid tr(A) + tr(B) = 0\}$ on $(\mathbb{C}^{n+1})^* \otimes \mathbb{C}^{m+1}$. Assume $m > n$. Cauchy formulas in our setting give

$$ch(S(\mathfrak{g}_1^-)) = ch(S((\mathbb{C}^{n+1})^* \otimes \mathbb{C}^{m+1})) = \sum_{\lambda} L^{A_m}(\tau(\lambda)) L^{A_n}(\lambda) \tag{3.9}$$

where for $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{n+1}$

$$\lambda = \sum_{i=1}^{n+1} \lambda_i \delta_i, \quad \tau(\lambda) = -w_0(\sum_{i=1}^{n+1} \lambda_i \epsilon_i). \tag{3.10}$$

and w_0 is the longest element in the symmetric group $W(A_m)$. Set

$$\gamma_1 = \epsilon_{m+1} - \delta_1, \quad \gamma_2 = \epsilon_m - \delta_2, \ldots\ldots, \gamma_{n+1} = \epsilon_{m-n+1} - \delta_{n+1}. \tag{3.11}$$

Then (3.8) and (3.9) imply

$$S(\mathfrak{g}_1^-) \otimes \mathbb{C}_{-\rho_1} = \bigoplus_{s_1 \geq s_2 \geq \ldots \geq s_{n+1}} L^{A_m \times A_n}(-\rho_1 - s_1\gamma_1 - \ldots - s_{n+1}\gamma_{n+1}). \tag{3.12}$$

Denote by $\lambda_{s_1,\dots,s_{n+1}}$ the $\mathfrak{g}_0$-dominant weight appearing in the r.h.s. of the above expression. Taking the $\mathfrak{g}_0$-supercharacter of both sides of (3.12) and using the Weyl character formula, we have

$$\frac{e^{-\rho_1}}{\prod\limits_{\beta\in\Delta_1^+}(1+e^{-\beta})}=\sum_{s_1\geq s_2\geq\dots\geq s_{n+1}}(-1)^{s_1+s_2+\dots+s_{n+1}}chL^{A_m\times A_n}(\lambda_{s_1,\dots,s_{n+1}})= \tag{3.13}$$
$$\sum_{s_1\geq s_2\geq\dots\geq s_{n+1}}(-1)^{s_1+s_2+\dots+s_{n+1}}\sum_{w\in W}sgn(w)\frac{e^{w(\lambda_{s_1,\dots,s_{n+1}}+\rho_0)}}{e^{\rho_0}\prod\limits_{\beta\in\Delta_0^+}(1-e^{-\alpha})}.$$

Then, multiplying the first and last member of the equalities in (3.13) by $e^{\rho_0}\prod_{\beta\in\Delta_0^+}(1-e^{-\beta})$, we obtain

$$e^\rho R=\sum_{s_1\geq s_2\geq\dots\geq s_{n+1}}(-1)^{s_1+s_2+\dots+s_{n+1}}\sum_{w\in W}sgn(w)e^{w(\rho-s_1\gamma_1-\dots-s_{n+1}\gamma_{n+1})}.$$

Hence we have proven formula (3.14) below, which is an instance of (1.3). Deriving the companion formula (3.15) is even easier: start from

$$\frac{e^{-\rho_1}}{\prod\limits_{\beta\in\Delta_1^+}(1-e^{-\beta})}=\sum_{s_1\geq s_2\geq\dots\geq s_{n+1}}chL^{A_m\times A_n}(\lambda_{s_1,\dots,s_{n+1}})$$

and proceed as above. So we have proved the following proposition.

Proposition 3.4 *Let $\mathfrak{g}$ be a Lie superalgebra of type $A(m,n), m>n$. Then for a distinguished set of positive roots we have:*

$$e^\rho R=\sum_{w\in W}sgn(w)w\frac{e^\rho}{(1+e^{-\gamma_1})(1-e^{-\gamma_1-\gamma_2})\cdots(1+(-1)^{n+1}e^{-\gamma_1-\gamma_2-\dots-\gamma_{n+1}})}, \tag{3.14}$$
$$e^\rho \check{R}=\sum_{w\in W}sgn(w)w\frac{e^\rho}{(1-e^{-\gamma_1})(1-e^{-\gamma_1-\gamma_2})\cdots(1-e^{-\gamma_1-\gamma_2-\dots-\gamma_{n+1}})}. \tag{3.15}$$

Remark 3.1 *The above formulas hold clearly in $gl(n+1,n+1)$, but do not restrict to $sl(n+1,n+1)$, since the last factor in the r.h.s. of* (3.15) *has a pole. Note that this factor is W-invariant, hence can be taken out of the sum. Since the left hand side restricts to $sl(n+1,n+1)$, the sum*

$$\sum_{w\in W}sgn(w)w\frac{e^\rho}{(1-e^{-\gamma_1})(1-e^{-\gamma_1-\gamma_2})\cdots(1-e^{-\gamma_1-\gamma_2-\dots-\gamma_n})}$$

is divisible by $1-e^{-\gamma_1-\gamma_2-\dots-\gamma_{n+1}}$. After simplifying, we may restrict to the Cartan subalgebra of $A(n,n)$ getting a superdenominator formula in this type too.

Remark 3.2 *The above reasoning works also in type C. There are two distinguished sets of positive roots (cf. [5, 2.5.4]), one being the opposite of the other. Using Corollary 3.3 and a theorem of Schmid [10] in place of Cauchy formulas we get* (1.3) *and* (1.4) *in this case.*

4 The $\mathfrak{g}_0$-character of $M^{\Delta^+}(\mathfrak{g}_1)$ via compact dual pairs

We start discussing the possible distinguished root systems up to W-equivalence for type II Lie superalgebras of defect greater than 1, following [5].

In type $B(m,n)$ there is a unique distinguished set of positive roots Δ^+_B, which, with notation as in [5], can be described as follows. We have, for $1 \le i \ne j \le m$, $1 \le k \ne l \le n$,

$$\Delta_0^+ = \{\epsilon_i \pm \epsilon_j, \epsilon_i, \delta_k \pm \delta_l, 2\delta_k\}, \qquad \Delta_1^+ = \{\delta_k \pm \epsilon_i, \delta_k\}, \tag{4.1}$$

$$\overline{\Delta}_0^+ = \{\epsilon_i \pm \epsilon_j, \epsilon_i, \delta_k \pm \delta_l\}, \qquad \overline{\Delta}_1^+ = \{\delta_k \pm \epsilon_i\}, \tag{4.2}$$

$$\Pi = \{\delta_1 - \delta_2, \dots, \delta_n - \epsilon_1, \epsilon_1 - \epsilon_2, \dots, \epsilon_{m-1} - \epsilon_m, \epsilon_m\}, \tag{4.3}$$

$$2\rho_1 = (2m+1)(\delta_1 + \ldots + \delta_n). \tag{4.4}$$

Note that $\pm\{\epsilon_i \pm \epsilon_j, \epsilon_i\}$ is a root system of type B_m (which will be denoted by $\Delta(B_m)$), that $\pm\{\delta_k \pm \delta_l, 2\delta_k\}$ is a root system of type C_n (which will be denoted by $\Delta(C_n)$) and that $\pm\{\delta_k - \delta_l \mid 1 \le k \ne l \le n\}$ is a root system of type A_{n-1} (which will be denoted by $\Delta(A_{n-1})$).

In type $D(m,n)$ there are three distinguished sets of positive roots $\Delta^+_{D1}, \Delta^+_{D2}, \Delta^+_{D2'}$. The corresponding sets of simple roots are

$$\Pi_1 = \{\delta_1 - \delta_2, \dots, \delta_n - \epsilon_1, \epsilon_1 - \epsilon_2, \dots, \epsilon_{m-1} - \epsilon_m, \epsilon_{m-1} + \epsilon_m\},$$
$$\Pi_2 = \{\epsilon_1 - \epsilon_2, \dots, \epsilon_{m-1} - \epsilon_m, \epsilon_m - \delta_1, \delta_1 - \delta_2, \dots, \delta_{n-1} - \delta_n, 2\delta_n\},$$
$$\Pi_2' = \{\epsilon_1 - \epsilon_2, \dots, \epsilon_{m-1} + \epsilon_m, -\epsilon_m - \delta_1, \delta_1 - \delta_2, \dots, \delta_{n-1} - \delta_n, 2\delta_n\}.$$

Theorem 4.1 *The character of $M^{\Delta^+}(\mathfrak{g}_1)$ as a $\mathfrak{g}_0$-module is afforded by the Theta correspondence for the compact dual pairs (G_1, G_2) as in the following table*

Δ^+	(G_1, G_2)
Δ^+_B	$(O(2m+1), Sp(2n,\mathbb{R}))$
Δ^+_A	$(U(m), U(n))$
Δ^+_{D1}	$(O(2m), Sp(2n,\mathbb{R}))$
Δ^+_{D2}	$(Sp(m), O^*(2n))$
$\Delta^+_{D2'}$	$(Sp(m), O^*(2n))$

(4.5)

For a quick review of the Theta correspondence see e.g. [1]. The explicit Theta correspondence is provided in [8] for the first, second and third dual pairs and in [9] for the fourth and fifth.

4.1 $B(m,n), m \ge n$.

Theorem 7.2 of [8] and (3.7) give

$$ch\, M^{\Delta^+}(\mathfrak{g}_1) = \frac{e^{-\rho_1}}{\prod_{\beta\in\Delta_1^+}(1-e^{-\beta})} = \tag{4.6}$$
$$\sum_{a_1\ge a_2\ge\ldots\ge a_n\ge 0} chL^{C_n}(-(a_n+m+\tfrac{1}{2})\delta_1 - \ldots - (a_1+m+\tfrac{1}{2})\delta_n)\ chL^{B_m}(a_1\epsilon_1 + \ldots + a_n\epsilon_n).$$

By [4, Theorem 9.2 a)], $L^{C_n}(-(a_n+m+\frac{1}{2})\delta_1-\ldots-(a_1+m+\frac{1}{2})\delta_n)$ is an irreducible parabolic Verma module w.r.t. $\Delta(A_{n-1})$. To prove irreducibility we have to show that if $\lambda=-(a_n+m+\frac{1}{2})\delta_1-\ldots-(a_1+m+\frac{1}{2})\delta_n$ then

$$(\lambda+\rho^{C_n},\psi)\notin\mathbb{Z}^{>0}\qquad\forall\,\psi\in\,\Delta^+(C_n)\setminus\Delta^+(A_{n-1}).$$

Since $\lambda+\rho^{C_n}=(n-a_n-m-\frac{1}{2})\delta_1+(n-a_{n-1}-m-\frac{3}{2})\delta_2+\ldots+(-a_1-m-\frac{1}{2})\delta_n$, the condition $m\geq n$ implies that all coefficients of the $\delta_i,\delta_i+\delta_j$ are not positive integers and the claim follows. Therefore the character is given by

$$\begin{aligned}
&chL^{C_n}(-(a_n+m+\tfrac{1}{2})\delta_1-\ldots-(a_1+m+\tfrac{1}{2})\delta_n)\\
&=\frac{chL^{A_{n-1}}(-(a_n+m+\frac{1}{2})\delta_1-\ldots-(a_1+m+\frac{1}{2})\delta_n)}{\prod\limits_{1\leq k,\,l\leq n}(1-e^{-(\delta_k+\delta_l)})}\\
&=\frac{\sum_{w\in W(A_{n-1})}sgn(w)we^{\rho^{A_{n-1}}-(a_n+m+\frac{1}{2})\delta_1-\ldots-(a_1+m+\frac{1}{2})\delta_n}}{\prod\limits_{1\leq k,\,l\leq n}(1-e^{-(\delta_k+\delta_l)})\cdot\prod\limits_{1\leq k<l\leq n}(1-e^{-(\delta_k-\delta_l)})}
\end{aligned}\tag{4.7}$$

where the second equality has been obtained using the Weyl character formula. Again Weyl formula allows us to make explicit the character of $L^{B_m}(a_1\epsilon_1+\ldots+a_n\epsilon_n)$:

$$chL^{B_m}(a_1\epsilon_1+\ldots+a_n\epsilon_n)=\frac{\sum_{w\in W(B_m)}sgn(w)we^{\rho^{B_m}+a_1\epsilon_1+\ldots+a_n\epsilon_n}}{\prod\limits_{1\leq i<j\leq m}(1-e^{-(\epsilon_i-\epsilon_j)})(1-e^{-(\epsilon_i+\epsilon_j)})\prod\limits_{i=1}^{m}(1-e^{-\epsilon_i})}.\tag{4.8}$$

Set now

$$\gamma_1=\delta_n-\epsilon_1,\,\gamma_2=\delta_{n-1}-\epsilon_2,\ldots,\gamma_n=\delta_1-\epsilon_n.\tag{4.9}$$

Combining (4.6), (4.7),(4.8),(4.4) we obtain

Proposition 4.2 *If $\gamma_1,\ldots,\gamma_n$ are defined by* (4.9)*, we have*

$$e^\rho R=\sum_{w\in W(A_{n-1})\times W(B_m)}sgn(w)w\frac{e^\rho}{(1+e^{-\gamma_1})(1-e^{-(\gamma_1+\gamma_2)})\cdots(1+(-1)^{n+1}e^{-(\gamma_1+\ldots+\gamma_n)})},\tag{4.10}$$

$$e^\rho\check{R}=\sum_{w\in W(A_{n-1})\times W(B_m)}sgn(w)w\frac{e^\rho}{(1-e^{-\gamma_1})(1-e^{-(\gamma_1+\gamma_2)})\cdots(1-e^{-(\gamma_1+\ldots+\gamma_n)})}.\tag{4.11}$$

Remark 4.1 *We want to prove that* (4.11) *coincides with* (1.4)*. Recall that $e^\rho\check{R}$ is such that $w(e^\rho\check{R})=sgn'(w)e^\rho\check{R}$. Take $g\in\Gamma=W(C_n)/W(A_{n-1})$, i.e., a sign change on the δ_i, and compute:*

$$\sum_{g\in\Gamma}sgn'(g)g(e^\rho\check{R})=2^ne^\rho\check{R}.$$

On the other hand, note that $\Gamma W(A_{n-1})=W(C_n)$, therefore if we apply $\sum_{g\in\Gamma}sgn'(g)g$ we get the (suitably signed) summation over the full Weyl group W, and (4.11) *becomes* (1.4)*.*

4.2 $B(2,4)$.

This is a defect 2 case. By Kashiwara-Vergne theorem,

$$M^{\Delta^+}(\mathfrak{g}_1) = \sum_{a_1 \geq a_2 \geq 0} L(-\tfrac{5}{2}\delta_1 - \tfrac{5}{2}\delta_2 - (\tfrac{5}{2}+a_2)\delta_3 - (\tfrac{5}{2}+a_1)\delta_4) \otimes L(a_1\epsilon_1 + a_2\epsilon_2) + \\ \sum_{a_1 \geq a_2 \geq 1} L(-\tfrac{5}{2}\delta_1 - \tfrac{7}{2}\delta_2 - (\tfrac{5}{2}+a_2)\delta_3 - (\tfrac{5}{2}+a_1)\delta_3) \otimes L(a_1\epsilon_1 + a_2\epsilon_2) \tag{4.12}$$

A computation with Kazhdan-Lusztig polynomials shows that we can write the $sp(8,\mathbb{C})$-modules appearing in terms of the $\Delta(A_3)$-parabolic Verma modules whose highest weights shifted by ρ^{C_4} are

$$\tfrac{3}{2}\delta_1 + \tfrac{1}{2}\delta_2 - (\tfrac{1}{2}+a_2)\delta_3 - (\tfrac{3}{2}+a_1)\delta_4, \quad -\tfrac{1}{2}\delta_1 - \tfrac{3}{2}\delta_2 - (\tfrac{1}{2}+a_2)\delta_3 - (\tfrac{3}{2}+a_1)\delta_4, \\ \tfrac{3}{2}\delta_1 - \tfrac{1}{2}\delta_2 - (\tfrac{1}{2}+a_2)\delta_3 - (\tfrac{3}{2}+a_1)\delta_4, \quad \tfrac{1}{2}\delta_1 - \tfrac{3}{2}\delta_2 - (\tfrac{1}{2}+a_2)\delta_3 - (\tfrac{3}{2}+a_1)\delta_4. \tag{4.13}$$

Hence we have

$$e^{\rho}\check{R} = \sum_{w \in W(A_3)} \sum_{u \in A} \sum_{v \in W(B_2)} sgn(w) sgn(v)\, w\, u\, v\, \frac{e^{\rho}}{(1-e^{-\delta_3+\epsilon_1})(1-e^{-\delta_3-\delta_4+\epsilon_1+\epsilon_2})} \tag{4.14}$$

where A is a set of coset representatives related to the list (4.13). Now argue as in Remark 4.1. Take $g \in \Gamma = W(C_4)/W(A_3)$. On the one hand $\sum_{g\in\Gamma} sgn'(g) g(e^{\rho}\check{R}) = 16 e^{\rho}\check{R}$. On the other hand, note that $\Gamma W(A_3) = \Gamma W(A_3) A = W(C_4)$, therefore if we apply $\sum_{g\in\Gamma} sgn'(g) g$ to the r.h.s. of (4.14) we get four times the r.h.s. of (4.14). So

$$e^{\rho}\check{R} = \frac{1}{4}\sum_{w \in W} sgn'(w) w \frac{e^{\rho}}{(1-e^{-\delta_3+\epsilon_1})(1-e^{-\delta_3-\delta_4+\epsilon_1+\epsilon_2})}$$

proving (1.4) in this case. In the general case, the calculation of the KL-polynomials is replaced by the use of a result of Enright on the $\mathfrak{u}$-homology of unitary highest weight modules (cf. [2]).

5 Final remarks

We would like to rephrase our main theorem in a form which seems most suitable for a generalization. We need to single out a special maximal isotropic subset S of positive roots. Fix a distinguished set of positive roots Δ^+. Construct $S = S_1 \cup \ldots \cup S_m = \{\gamma_1, \ldots, \gamma_d\}$, $d = \operatorname{def} \mathfrak{g}$, as follows: S_1 is an isotropic subset having maximal cardinality in the set of simple roots, and inductively S_i is such a subset in the set of indecomposable roots of $S_{i-1}^{\perp} \setminus S_{i-1}$. Define

$$\gamma_i^{\leq} = \{\beta \in S,\, \beta \leq \gamma_i\}, \quad \langle \gamma_i \rangle = \sum_{\beta \in \gamma_i^{\leq}} \beta, \quad sgn(\gamma_i) = (-1)^{|\gamma_i^{\leq}|+1} \tag{5.1}$$

where as usual $\alpha \leq \beta$ if $\beta - \alpha$ is a sum of positive roots. This procedure determines uniquely S once Δ^+ is fixed (up to a mild exception in type D) and gives rise to the set $\{\gamma_1, \ldots, \gamma_d\}$ of Theorem 1.1.

Theorem 5.1 *Let $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1$ be a basic classical Lie superalgebra of defect d, where $\mathfrak{g} = A(d-1, d-1)$ is replaced by $gl(d, d)$. Then, for any distinguished set of positive roots, if S is as above, we have*

$$e^{\rho} R = \frac{1}{C} \sum_{w \in W} sgn(w) w \frac{e^{\rho}}{\prod_{i=1}^{d} (1 + sgn(\gamma_i) e^{-\langle \gamma_i \rangle})}, \tag{5.2}$$

$$e^{\rho} \check{R} = \frac{1}{C} \sum_{w \in W} sgn'(w) w \frac{e^{\rho}}{\prod_{i=1}^{d} (1 - e^{-\langle \gamma_i \rangle})}, \tag{5.3}$$

where C is defined in (1.5).

We would like to remark that the above statement holds true in the hypothesis of Kac-Wakimoto-Gorelik theorem (in which case $e^{-\langle \gamma_i \rangle} = e^{-\gamma_i}$ and $C = |W/W^{\sharp}|$).

Denote by Q, Q_0 the lattices generated by all roots and even roots, respectively. Set

$$\varepsilon(\eta) = \begin{cases} 1 & \text{if } \eta \in Q_0 \\ -1 & \text{if } \eta \in Q \setminus Q_0 \end{cases}, \quad ||\gamma|| = \sum_{\beta \in \gamma^{\leq}} \varepsilon(\gamma - \beta) \beta.$$

Note that for the γ_i appearing in (5.2), (5.3) the equality $\langle \gamma_i \rangle = ||\gamma_i||$ holds. We modify the construction of S as follows: S_1 is an isotropic subset having maximal cardinality in a maximal subdiagram of type A of odd cardinality having only odd simple roots, and inductively S_i is such a subset in the set of indecomposable roots of $S_{i-1}^{\perp} \setminus S_{i-1}$. This time the choice of S is not unique.

Conjecture 5.2 *Let $\mathfrak{g}$ be a basic classical Lie superalgebra of defect d, where $\mathfrak{g} = A(d-1, d-1)$ is replaced by $gl(d, d)$, and Δ^+ any set of positive roots. Let S be any maximal isotropic subset of Δ^+ built up as above. Then*

$$e^{\rho} R = \frac{1}{K} \sum_{w \in W} sgn(w) w \frac{e^{\rho}}{\prod_{i=1}^{d} (1 + sgn(\gamma_i) e^{-||\gamma_i||})},$$

$$e^{\rho} \check{R} = \frac{1}{K} \sum_{w \in W} sgn'(w) w \frac{e^{\rho}}{\prod_{i=1}^{d} (1 - e^{-||\gamma_i||})},$$

where

$$K = \frac{C\, d!}{\prod_{\gamma \in S} \frac{ht(\gamma)+1}{2}},$$

C is defined in (1.5) *and $ht(\gamma)$ denotes the height of the root γ w.r.t. to the simple roots corresponding to Δ^+. Moreover, there exists a choice of S for which $||\gamma_i||$ is a linear combination with non negative coefficients of positive roots for any $i = 1, \ldots, d$.*

Acknowledgement

We would like to thank Corrado De Concini for useful conversations.

References

[1] J. Adams, *The Theta correspondence over* $\mathbb{R}$, Harmonic analysis, group representations, automorphic forms and invariant theory, Lecture Notes Series, Institut for Mathematical Sciences, National University of Singapore, Vol. 12, p. 1–37

[2] T. Enright, *Analogues of Kostant's* $\mathfrak{u}$*-cohomology formulas for unitary highest weight modules.* J. Reine Angew. Math. 392 (1988), 27–36.

[3] M. Gorelik, *Weyl denominator identity for finite-dimensional Lie superalgebras*, arXiv:0905.1181

[4] J. Humphreys, *Representations of semisimple Lie algebras in the BGG category* $\mathcal{O}$. Graduate Studies in Mathematics, 94. American Mathematical Society, Providence, RI, 2008.

[5] V. G. Kac, *Lie superalgebras*, Advances in Math. **26** (1977), no. 1, 8–96.

[6] V. G. Kac, *Representations of classical Lie superalgebras.* Differential geometrical methods in mathematical physics, II (Proc. Conf., Univ. Bonn, Bonn, 1977), pp. 597–626, Lecture Notes in Math., 676, Springer, Berlin, 1978.

[7] V. G. Kac, W. Wakimoto, *Integrable highest weight modules over affine superalgebras and number theory.* Lie theory and geometry, 415–456, Progr. Math., 123, Birkhäuser Boston, 1994.

[8] M. Kashiwara, M. Vergne, *On the Segal-Shale-Weil representations and harmonic polynomials.* Invent. Math. **44** (1978), no. 1, 1–47.

[9] J.-S. Li, A. Paul, E.-C. Tan, C.-B. Zhu, *The explicit duality correspondence of* $(\mathrm{Sp}(p,q), \mathrm{O}^*(2n))$. J. Funct. Anal. 200 (2003), no. 1, 71–100.

[10] W. Schmid, *Die Randwerte holomorpher Funktionen auf hermitesch symmetrischen Raumen*, Invent. Math. **9** (1969/70), 61-80.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 678–689

Chain enumeration of k-divisible noncrossing partitions of classical types

Jang Soo Kim†

LIAFA, Université Paris Diderot, 175 rue du Chevaleret, 75013 Paris, France

Abstract. We give combinatorial proofs of the formulas for the number of multichains in the k-divisible noncrossing partitions of classical types with certain conditions on the rank and the block size due to Krattenthaler and Müller. We also prove Armstrong's conjecture on the zeta polynomial of the poset of k-divisible noncrossing partitions of type A invariant under the 180° rotation in the cyclic representation.

Résumé. Nous donnons une preuve combinatoire de la formule pour le nombre de multichaînes dans les partitions k-divisibles non-croisées de type classique avec certaines conditions sur le rang et la taille du bloc due à Krattenthaler et Müller. Nous prouvons aussi la conjecture d'Amstrong sur le polynôme zeta du poset des partitions k-divisibles non-croisées de type A invariantes par la rotation de 180° dans la représentation cyclique.

Keywords: noncrossing partitions, chain enumeration

1 Introduction

For a finite set X, a *partition* of X is a collection of mutually disjoint nonempty subsets, called *blocks*, of X whose union is X. Let $\Pi(n)$ denote the poset of partitions of $[n] = \{1, 2, \dots, n\}$ ordered by refinement, i.e. $\pi \le \sigma$ if each block of σ is a union of blocks of π. There is a natural way to identify $\pi \in \Pi(n)$ with an intersection of reflecting hyperplanes of the Coxeter group A_{n-1}. For this reason, we will call $\pi \in \Pi(n)$ a *partition of type* A_{n-1}. With this observation Reiner [12] defined partitions of type B_n and type D_n as follows. A *partition of type* B_n is a partition π of $[\pm n] = \{1, 2, \dots, n, -1, -2, \dots, -n\}$ such that if B is a block of π then $-B = \{-x : x \in B\}$ is also a block of π, and there is at most one block, called *zero block*, which satisfies $B = -B$. A *partition of type* D_n is a partition of type B_n such that its zero block, if exists, has more than two elements. Let $\Pi_B(n)$ (resp. $\Pi_D(n)$) denote the poset of type B_n (resp. type D_n) partitions ordered by refinement.

A *noncrossing partition of type* A_{n-1}, or simply a noncrossing partition, is a partition $\pi \in \Pi(n)$ with the following property: if integers a, b, c and d with $a < b < c < d$ satisfy $a, c \in B$ and $b, d \in B'$ for some blocks B and B' of π, then $B = B'$.

Let k be a positive integer. A noncrossing partition is called *k-divisible* if the size of each block is divisible by k. Let $\mathrm{NC}(n)$ (resp. $\mathrm{NC}^{(k)}(n)$) denote the subposet of $\Pi(n)$ (resp. $\Pi(kn)$) consisting of the noncrossing partitions (resp. k-divisible noncrossing partitions).

†The author was supported by the grant ANR08-JCJC-0011.

Bessis [4], Brady and Watt [5] defined the generalized noncrossing partition poset $\mathrm{NC}(W)$ for each finite Coxeter group W, which satisfies $\mathrm{NC}(A_{n-1}) \cong \mathrm{NC}(n)$. Armstrong [1] defined the poset $\mathrm{NC}^{(k)}(W)$ of generalized k-divisible noncrossing partitions for each finite Coxeter group W, which reduces to $\mathrm{NC}(W)$ for $k = 1$ and satisfies $\mathrm{NC}^{(k)}(A_{n-1}) \cong \mathrm{NC}^{(k)}(n)$.

For each classical Coxeter group W, we have a combinatorial realization of $\mathrm{NC}^{(k)}(W)$. In other words, similar to $\mathrm{NC}(n)$ and $\mathrm{NC}^{(k)}(n)$, there are combinatorial posets $\mathrm{NC}_B(n) \subset \Pi_B(n)$, $\mathrm{NC}_B^{(k)}(n) \subset \Pi_B(kn)$, $\mathrm{NC}_D(n) \subset \Pi_D(n)$ and $\mathrm{NC}_D^{(k)}(n) \subset \Pi_D(kn)$, which are isomorphic to $\mathrm{NC}(B_n)$, $\mathrm{NC}^{(k)}(B_n)$, $\mathrm{NC}(D_n)$ and $\mathrm{NC}^{(k)}(D_n)$ respectively. Reiner [12] defined the poset $\mathrm{NC}_B(n)$ of noncrossing partitions of type B_n, which turned out to be isomorphic to $\mathrm{NC}(B_n)$. This poset is naturally generalized to the poset $\mathrm{NC}_B^{(k)}(n)$ of k-divisible noncrossing partitions of type B_n. Armstrong [1] showed that $\mathrm{NC}_B^{(k)}(n) \cong \mathrm{NC}^{(k)}(B_n)$. Athanasiadis and Reiner [3] defined the poset $\mathrm{NC}_D(n)$ of noncrossing partitions of type D_n and showed that $\mathrm{NC}_D(n) \cong \mathrm{NC}(D_n)$. Krattenthaler [10] defined the poset $\mathrm{NC}_D^{(k)}(n)$ of the k-divisible noncrossing partitions of type D_n using annulus and showed that $\mathrm{NC}_D^{(k)}(n) \cong \mathrm{NC}^{(k)}(D_n)$; see also [9].

In this paper we are mainly interested in the number of multichains in $\mathrm{NC}^{(k)}(n)$, $\mathrm{NC}_B^{(k)}(n)$ and $\mathrm{NC}_D^{(k)}(n)$ with some conditions on the rank and the block size.

Definition 1. For a multichain $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in a graded poset P with the maximum element $\hat{1}$, the *rank jump vector* of this multichain is the vector $(s_1, s_2, \ldots, s_{\ell+1})$, where $s_1 = \mathrm{rank}(\pi_1)$, $s_{\ell+1} = \mathrm{rank}(\hat{1}) - \mathrm{rank}(\pi_\ell)$ and $s_i = \mathrm{rank}(\pi_i) - \mathrm{rank}(\pi_{i-1})$ for $2 \le i \le \ell$.

We note that all the posets considered in this paper are graded with the maximum element, however, they do not necessarily have the minimum element. We also note that the results in this introduction have certain 'obvious' conditions on the rank jump vector or the block size, which we will omit for simplicity.

Edelman [6, Theorem 4.2] showed that the number of multichains in $\mathrm{NC}^{(k)}(n)$ with rank jump vector $(s_1, s_2, \ldots, s_{\ell+1})$ is equal to

$$\frac{1}{n}\binom{n}{s_1}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}}. \tag{1}$$

Modifying Edelman's idea of the proof of (1), Reiner found an analogous formula for the number of multichains in $\mathrm{NC}_B(n)$ with given rank jump vector. Later, Armstrong generalized Reiner's idea to find the following formula [1, Theorem 4.5.7] for the number of multichains in $\mathrm{NC}_B^{(k)}(n)$ with rank jump vector $(s_1, s_2, \ldots, s_{\ell+1})$:

$$\binom{n}{s_1}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}}. \tag{2}$$

Athanasiadis and Reiner [3, Theorem 1.2] proved that the number of multichains in $\mathrm{NC}_D(n)$ with rank jump vector $(s_1, s_2, \ldots, s_{\ell+1})$ is equal to

$$2\binom{n-1}{s_1}\binom{n-1}{s_2}\cdots\binom{n-1}{s_{\ell+1}} + \sum_{i=1}^{\ell+1}\binom{n-1}{s_1}\cdots\binom{n-2}{s_i-2}\cdots\binom{n-1}{s_{\ell+1}}. \tag{3}$$

To prove (3), they [3, Lemma 4.4] showed the following using incidence algebras and the Lagrange inversion formula: the number of multichains $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}_B(n)$ with rank jump vector

$(s_1, s_2, \dots, s_{\ell+1})$ such that i is the smallest integer for which π_i has a zero block is equal to

$$\frac{s_i}{n}\binom{n}{s_1}\binom{n}{s_2}\cdots\binom{n}{s_{\ell+1}}. \tag{4}$$

Since (4) is quite simple and elegant, it deserves a combinatorial proof. In this paper we prove a generalization of (4) combinatorially; see Lemma 11.

The number of noncrossing partitions with given block sizes has been studied as well. In the literature, for instance [1, 2, 3], $\operatorname{type}(\pi)$ for $\pi \in \Pi(n)$ is defined to be the integer partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_\ell)$ where the number of parts of size i is equal to the number of blocks of size i of π. However, to state the results in a uniform way, we will use the following different definition of $\operatorname{type}(\pi)$.

Definition 2. The *type* of a partition $\pi \in \Pi(n)$, denoted by $\operatorname{type}(\pi)$, is the sequence $(b; b_1, b_2, \dots, b_n)$ where b_i is the number of blocks of π of size i and $b = b_1 + b_2 + \cdots + b_n$. The *type* of $\pi \in \Pi_B(n)$ (or $\pi \in \Pi_D(n)$), denoted by $\operatorname{type}(\pi)$, is the sequence $(b; b_1, b_2, \dots, b_n)$ where b_i is the number of unordered pairs $(B, -B)$ of nonzero blocks of π of size i and $b = b_1 + b_2 + \cdots + b_n$. For a partition π in either $\Pi(kn)$, $\Pi_B(kn)$ or $\Pi_D(kn)$, if the size of each block of π is divisible by k, then we define the *k-type* $\operatorname{type}^{(k)}(\pi)$ of π to be $(b; b_k, b_{2k}, \dots, b_{kn})$ where $\operatorname{type}(\pi) = (b; b_1, b_2, \dots, b_{kn})$.

Kreweras [11, Theorem 4] proved that the number of $\pi \in \mathrm{NC}(n)$ with $\operatorname{type}(\pi) = (b; b_1, b_2, \dots, b_n)$ is equal to

$$\frac{n!}{b_1! b_2! \cdots b_n! (n-b+1)!} = \frac{1}{b}\binom{b}{b_1, b_2, \dots, b_n}\binom{n}{b-1}. \tag{5}$$

Athanasiadis [2, Theorem 2.3] proved that the number of $\pi \in \mathrm{NC}_B(n)$ with $\operatorname{type}(\pi) = (b; b_1, b_2, \dots, b_n)$ is equal to

$$\frac{n!}{b_1! b_2! \cdots b_n! (n-b)!} = \binom{b}{b_1, b_2, \dots, b_n}\binom{n}{b-1}. \tag{6}$$

Athanasiadis and Reiner [3, Theorem 1.3] proved that the number of $\pi \in \mathrm{NC}_D(n)$ with $\operatorname{type}(\pi) = (b; b_1, b_2, \dots, b_n)$ is equal to

$$\frac{(n-1)!}{b_1! b_2! \cdots b_n! (n-1-b)!} = \binom{b}{b_1, b_2, \dots, b_n}\binom{n-1}{b-1}, \tag{7}$$

if $b_1 + 2b_2 + \cdots + nb_n \le n-2$, and

$$(2(n-b)+b_1)\frac{(n-1)!}{b_1! b_2! \cdots b_n! (n-b)!} = 2\binom{b}{b_1, b_2, \dots, b_n}\binom{n-1}{b} + \binom{b-1}{b_1-1, b_2, \dots, b_n}\binom{n-1}{b-1}, \tag{8}$$

if $b_1 + 2b_2 + \cdots + nb_n = n$.

Armstrong [1, Theorem 4.4.4 and Theorem 4.5.11] generalized (5) and (6) as follows: the number of multichains $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}^{(k)}(n)$ and in $\mathrm{NC}_B^{(k)}(n)$ with $\operatorname{type}^{(k)}(\pi_1) = (b; b_1, b_2, \dots, b_n)$ are equal to, respectively,

$$\frac{(\ell kn)!}{b_1! b_2! \cdots b_n! (\ell kn - b + 1)!} = \frac{1}{b}\binom{b}{b_1, b_2, \dots, b_n}\binom{\ell kn}{b-1}, \tag{9}$$

and

$$\frac{(\ell kn)!}{b_1!b_2!\cdots b_n!(\ell kn-b)!} = \binom{b}{b_1, b_2, \ldots, b_n}\binom{\ell kn}{b}. \tag{10}$$

Krattenthaler and Müller [9] generalized all the above known results except (4) in the following three theorems.

Theorem 1. *[9, Corollary 12] Let $b, b_1, b_2, \ldots, b_n$ and $s_1, s_2, \ldots, s_{\ell+1}$ be nonnegative integers satisfying $\sum_{i=1}^n b_i = b$, $\sum_{i=1}^n i \cdot b_i = n$, $\sum_{i=1}^{\ell+1} s_i = n-1$ and $s_1 = n-b$. Then the number of multichains $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}^{(k)}(n)$ with rank jump vector $(s_1, s_2, \ldots, s_{\ell+1})$ and $\mathrm{type}^{(k)}(\pi_1) = (b; b_1, b_2, \ldots, b_n)$ is equal to*

$$\frac{1}{b}\binom{b}{b_1, b_2, \ldots, b_n}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}}.$$

Theorem 2. *[9, Corollary 14] Let $b, b_1, b_2, \ldots, b_n$ and $s_1, s_2, \ldots, s_{\ell+1}$ be nonnegative integers satisfying $\sum_{i=1}^n b_i = b$, $\sum_{i=1}^n i \cdot b_i \le n$, $\sum_{i=1}^{\ell+1} s_i = n$ and $s_1 = n-b$. Then the number of multichains $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}_B^{(k)}(n)$ with rank jump vector $(s_1, s_2, \ldots, s_{\ell+1})$ and $\mathrm{type}^{(k)}(\pi_1) = (b; b_1, b_2, \ldots, b_n)$ is equal to*

$$\binom{b}{b_1, b_2, \ldots, b_n}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}}.$$

Theorem 3. *[9, Corollary 16] Let $b, b_1, b_2, \ldots, b_n$ and $s_1, s_2, \ldots, s_{\ell+1}$ be nonnegative integers satisfying $\sum_{i=1}^n b_i = b$, $\sum_{i=1}^n i \cdot b_i \le n$, $\sum_{i=1}^n i \cdot b_i \ne n-1$, $\sum_{i=1}^{\ell+1} s_i = n$ and $s_1 = n-b$. Then the number of multichains $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}_D^{(k)}(n)$ with rank jump vector $(s_1, s_2, \ldots, s_{\ell+1})$ and $\mathrm{type}^{(k)}(\pi_1) = (b; b_1, b_2, \ldots, b_n)$ is equal to*

$$\binom{b}{b_1, b_2, \ldots, b_n}\binom{k(n-1)}{s_2}\cdots\binom{k(n-1)}{s_{\ell+1}},$$

if $b_1 + 2b_2 + \cdots + nb_n \le n-2$, and

$$2\binom{b}{b_1, b_2, \ldots, b_n}\binom{k(n-1)}{s_2}\cdots\binom{k(n-1)}{s_{\ell+1}} \\ + \frac{s_i-1}{b-1}\binom{b-1}{b_1-1, b_2, \ldots, b_n}\sum_{i=2}^{\ell+1}\binom{k(n-1)}{s_2}\cdots\binom{k(n-1)}{s_i-1}\cdots\binom{k(n-1)}{s_{\ell+1}},$$

if $b_1 + 2b_2 + \cdots + nb_n = n$.

Krattenthaler and Müller's proofs of Theorems 1, 2 and 3 were not combinatorial. Especially, in the introduction, they wrote that Theorems 1 and 2 seem amenable to combinatorial proofs, however, to find a combinatorial proof of Theorem 3 seems rather hopeless. In this paper, we will give combinatorial proofs of Theorems 1 and 2. For a combinatorial proof of Theorem 3, see the full version [8] of this paper.

This paper is organized as follows. In Section 2 we recall the definition of $\mathrm{NC}_B(n)$ and $\mathrm{NC}_D(n)$. In Section 3 we recall the bijection ψ in [7] between $\mathrm{NC}_B(n)$ and the set of pairs (σ, x), where $\sigma \in \mathrm{NC}(n)$ and x is either $\emptyset$, an edge or a block of σ. Then we find a necessary and sufficient condition for the two

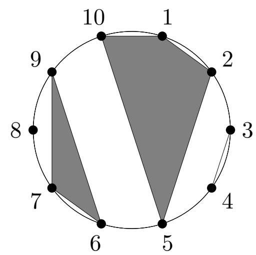


Fig. 1: The circular representation of $\{\{1,2,5,10\},\{3,4\},\{6,7,9\},\{8\}\}$.

pairs (σ_1,x_1) and (σ_2,x_2) to be $\psi^{-1}(\sigma_1,x_1)\le\psi^{-1}(\sigma_2,x_2)$ in the poset $\mathrm{NC}_B(n)$. This property will play a crucial role to prove Theorem 3. In Section 4 we prove Theorem 1 by modifying the argument of Edelman [6]. For $0<r<k$, we consider the subposet $\mathrm{NC}^{(k)}(n;r)$ of $\mathrm{NC}(nk+r)$ consisting of the partitions π such that all but one blocks of π have sizes divisible by k. Then we prove similar chain enumeration results for $\mathrm{NC}^{(k)}(n;r)$. We also prove that the poset $\widetilde{\mathrm{NC}}^{(2k)}(2n+1)$ suggested by Armstrong is isomorphic to $\mathrm{NC}^{(2k)}(n;k)$. With this, we prove Armstrong's conjecture on the zeta polynomial of $\widetilde{\mathrm{NC}}^{(2k)}(2n+1)$ and answer the question on rank-, type-selection formulas [1, Conjecture 4.5.14 and Open Problem 4.5.15]. In Section 5 we prove a generalization of (4) and Theorem 2. All the arguments in this paper are purely combinatorial.

2 Noncrossing partitions of classical types

Recall that $\Pi(n)$ denotes the poset of partitions of $[n]$ and $\Pi_B(n)$ (resp. $\Pi_D(n)$) denotes the poset of partitions of type B_n (resp. D_n). For simplicity, we will write a partition of type B_n or D_n in the following way:

$$\{\pm\{1,-3,6\},\{2,4,-2,-4\},\pm\{5,8\},\pm\{7\}\},$$

which means

$$\{\{1,-3,6\},\{-1,3,-6\},\{2,4,-2,-4\},\{5,8\},\{-5,-8\},\{7\},\{-7\}\}.$$

The *circular representation* of $\pi\in\Pi(n)$ is the drawing obtained as follows. Arrange n vertices around a circle which are labeled with the integers $1,2,\dots,n$. For each block B of π, draw the convex hull of the vertices whose labels are the integers in B. For example, see Figure 1. It is easy to see that the following definition coincides with the definition of a noncrossing partition in the introduction: π is a *noncrossing partition* if the convex hulls in the circular representation of π do not intersect.

Let $\pi\in\Pi_B(n)$. The *circular representation* of π is the drawing obtained as follows. Arrange $2n$ vertices in a circle which are labeled with the integers $1,2,\dots,n,-1,-2,\dots,-n$. For each block B of π, draw the convex hull of the vertices whose labels are the integers in B. For example, see Figure 2. Note that the circular representation of $\pi\in\Pi_B(n)$ is invariant, if we do not concern the labels, under the 180° rotation, and the zero block of π, if exists, corresponds to the convex hull containing the center. Then π is a *noncrossing partition of type* B_n if the convex hulls do not intersect.

Let $\pi\in\Pi_D(n)$. The *circular representation* of π is the drawing obtained as follows. Arrange $2n-2$ vertices labeled with $1,2,\dots,n-1,-1,-2,\dots,-(n-1)$ in a circle and put a vertex labeled with $\pm n$

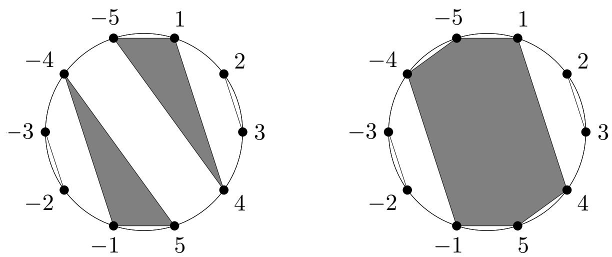


Fig. 2: The circular representations of $\{\pm\{1,4,-5\},\pm\{2,3\}\}$ and $\{\{1,4,5,-1,-4,-5\},\pm\{2,3\}\}$.

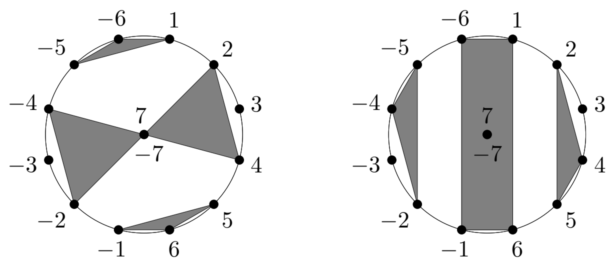


Fig. 3: The type D circular representations of $\{\pm\{1,-5,-6\},\pm\{2,4,-7\},\pm\{3\}\}$ and $\{\{1,6,-1,-6\},\pm\{2,4,5\},\pm\{3\}\}$.

at the center. For each block B of π, draw the convex hull of the vertices whose labels are in B. Then π is a *noncrossing partition of type* D_n if the convex hulls do not intersect in their interiors. For example, see Figure 3.

Let $\pi \in \Pi(n)$. An *edge* of π is a pair (i,j) of integers with $i<j$ such that $i,j \in B$ for a block B of π and there is no other integer k in B with $i<k<j$. The *standard representation* of π is the drawing obtained as follows. Arrange the integers $1,2,\ldots,n$ in a horizontal line. For each edge (i,j) of π, connect the integers i and j with an arc above the horizontal line. For example, see Figure 4. Then π is a noncrossing partition if and only if the arcs in the standard representation do not intersect.

Let $\pi \in \Pi_B(n)$. The *standard representation* of π is the drawing obtained as follows. Arrange the integers $1,2,\ldots,n,-1,-2,\ldots,-n$ in a horizontal line. Then connect the integers i and j with an arc above the horizontal line for each pair (i,j) of integers such that i,j are in the same block B of π and there is no other integer in B between i and j in the horizontal line. For example, see Figure 5. Then π is a noncrossing partition of type B_n if and only if the arcs in the standard representation do not intersect.

Let $\mathrm{NC}_B(n)$ denote the subposet of $\Pi_B(n)$ consisting of the noncrossing partitions of type B_n. Note

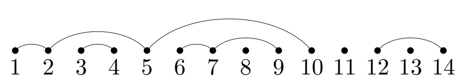


Fig. 4: The standard representation of $\{\{1,2,5,10\},\{3,4\},\{6,7,9\},\{8\},\{11\},\{12,14\},\{13\}\}$.

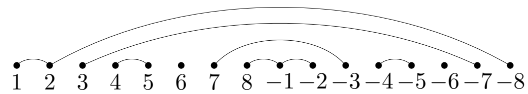


Fig. 5: The standard representation of $\{\pm\{1,2,-8\}, \pm\{3,-7\}, \pm\{4,5\}, \pm\{6\}\}$.

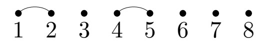


Fig. 6: The partition τ obtained from the partition π in Figure 5 by removing all the negative integers. Then $X = \{\{1,2\}, \{3\}, \{7\}, \{8\}\}$ is the set of blocks of τ which are not blocks of π.

that for $\pi \in \mathrm{NC}_B(n)$, $\mathrm{rank}(\pi) = n - \mathrm{nz}(\pi)$, where $\mathrm{nz}(\pi)$ denotes the number of unordered pairs $(B, -B)$ of nonzero blocks of π.

3 Interpretation of noncrossing partitions of type B_n

Let $\mathfrak{B}(n)$ denote the set of pairs (σ, x), where $\sigma \in \mathrm{NC}(n)$ and x is either $\emptyset$, an edge or a block of σ. Note that since for each $\sigma \in \mathrm{NC}(n)$, we have $n+1$ choices for x with $(\sigma, x) \in \mathfrak{B}(n)$, one may consider $\mathfrak{B}(n)$ as $\mathrm{NC}(n) \times [n+1]$.

Let us recall the bijection $\psi : \mathrm{NC}_B(n) \to \mathfrak{B}(n)$ in [7].

Let $\pi \in \mathrm{NC}_B(n)$. Then let τ be the partition of $[n]$ obtained from π by removing all the negative integers and let X be the set of blocks of τ which are not blocks of π. For example, see Figure 6. Now assume that X has k blocks $A_1, A_2, \ldots, A_k$ with $\max(A_1) < \max(A_2) < \cdots < \max(A_k)$. Let σ be the partition obtained from τ by unioning A_r and A_{k+1-r} for all $r = 1, 2, \ldots, \lfloor (k-1)/2 \rfloor$. Let

$$x = \begin{cases} \emptyset, & \text{if } k = 0; \\ (\min(A_{k/2}), \max(A_{k/2+1})), & \text{if } k \neq 0 \text{ and } k \text{ is even;} \\ A_{(k+1)/2}, & \text{if } k \text{ is odd.} \end{cases}$$

Then we define $\psi(\pi)$ to be the pair (σ, x). For example, see Figure 7.

Theorem 4. *[7] The map* $\psi : \mathrm{NC}_B(n) \to \mathfrak{B}(n)$ *is a bijection. Moreover, for* $\pi \in \mathrm{NC}_B(n)$ *with* $\mathrm{type}(\pi) = (b; b_1, b_2, \ldots, b_n)$ *and* $\psi(\pi) = (\sigma, x)$*, we have* $\mathrm{type}(\sigma) = \mathrm{type}(\pi)$ *if* π *does not have a zero block; and* $\mathrm{type}(\sigma) = (b+1; b_1, \ldots, b_i + 1, \ldots, b_n)$ *if* π *has a zero block of size* $2i$.

Now we will find a necessary and sufficient condition for $(\sigma_1, x_1), (\sigma_2, x_2) \in \mathfrak{B}(n)$ to be $\psi^{-1}(\sigma_1, x_1) \leq \psi^{-1}(\sigma_2, x_2)$ in $\mathrm{NC}_B(n)$.

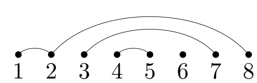


Fig. 7: The partition σ obtained from the partition τ in Figure 5 by unioning $\{1,2\}$, $\{8\}$ and $\{3\}$, $\{7\}$ which are the blocks in $X = \{\{1,2\}, \{3\}, \{7\}, \{8\}\}$. Since X has even number of blocks, x is the edge $(3,7)$. Then $\psi(\pi) = (\sigma, x)$ for the partition π in Figure 5.

For a partition π (either in $\Pi(n)$ or in $\Pi_B(n)$), we write $i \overset{\pi}{\sim} j$ if i and j are in the same block of π and $i \overset{\pi}{\nsim} j$ otherwise. Note that if $\psi(\pi) = (\sigma, x)$, then we have $i \overset{\sigma}{\sim} j$ if and only if $i \overset{\pi}{\sim} j$ or $i \overset{\pi}{\sim} -j$. The following lemmas are clear from the construction of ψ.

Proposition 5. *Let $\psi(\pi_1) = (\sigma_1, x_1)$ and $\psi(\pi_2) = (\sigma_2, x_2)$. Then $\pi_1 \le \pi_2$ if and only if $\sigma_1 \le \sigma_2$ and one of the following holds:*

1. *$x_1 = x_2 = \emptyset$,*
2. *x_2 is an edge (a, b) of σ_2 and x_1 is the unique minimal length edge (i, j) of σ_1 with $i \le a < b \le j$ if such an edge exists; and $x_1 = \emptyset$ otherwise.*
3. *x_2 is a block of σ_2, and x_1 is a block of σ_1 with $x_1 \subset x_2$,*
4. *x_2 is a block of σ_2 and x_1 is an edge (i, j) of σ_1 with $i, j \in x_2$.*
5. *x_2 is a block of σ_2 and x_1 is the minimal length edge (i, j) of σ_1 with $i < \min(x_2) \le \max(x_2) < j$ if such an edge exists; and $x_1 = \emptyset$ otherwise.*

4 k-divisible noncrossing partitions of type A

Let k be a positive integer. A noncrossing partition $\pi \in \mathrm{NC}(kn)$ is *k-divisible* if the size of each block is divisible by k. Let $\mathrm{NC}^{(k)}(n)$ denote the subposet of $\mathrm{NC}(kn)$ consisting of k-divisible noncrossing partitions. Then $\mathrm{NC}^{(k)}(n)$ is a graded poset with the rank function $\mathrm{rank}(\pi) = n - \mathrm{bk}(\pi)$, where $\mathrm{bk}(\pi)$ is the number of blocks of π.

To prove (1), Edelman [6] found a bijection between the set of pairs $(\mathbf{c}, a)$ of a multichain $\mathbf{c} : \pi_1 \le \pi_2 \le \cdots \le \pi_{\ell+1}$ in $\mathrm{NC}^{(k)}(n)$ with rank jump vector $(s_1, s_2, \ldots, s_{\ell+1})$ and an integer $a \in [n]$ and the set of $(\ell+1)$-tuples $(L, R_1, R_2, \ldots, R_\ell)$ with $L \subset [n]$, $|L| = n - s_1$, $R_i \subset [kn]$, and $|R_i| = s_i$ for $i \in [\ell]$. This bijection has been extended to the noncrossing partitions of type B_n [1, 12] and type D_n [3].

In this section we prove Theorem 1 by modifying the idea of Edelman. Let us first introduce several notations.

4.1 The cyclic parenthesization

Let $P(n)$ denote the set of pairs (L, R) of subsets $L, R \subset [n]$ with the same cardinality. Let $(L, R) \in P(n)$. We can identify (L, R) with the *cyclic parenthesization* of (L, R) defined as follows. We place a left parenthesis before the occurrence of i for each $i \in L$ and a right parenthesis after the occurrence of i for each $i \in R$ in the sequence $1, 2, \ldots, n$. We consider this sequence in cyclic order.

For $x \in R$, the *size of* x is defined to be the number of integers enclosed by x and its corresponding left parenthesis, which are not enclosed by any other matching pair of parentheses. The *type* of (L, R), denoted by $\mathrm{type}(L, R)$, is defined to be $(b; b_1, b_2, \ldots, b_n)$, where b_i is the number of $x \in R$ whose sizes are equal to i and $b = b_1 + b_2 + \cdots + b_n$.

Example 1. Let $(L, R) = (\{2, 3, 9, 11, 15, 16\}, \{1, 4, 5, 8, 9, 12\}) \in P(16)$. Then the cyclic parenthesization is the following:

$$1)\ (2\ (3\ 4)\ 5)\ 6\ 7\ 8)\ (9)\ 10\ (11\ 12)\ 13\ 14\ (15\ (16 \tag{11}$$

Since we consider (11) in the cyclic order, the right parenthesis of 1 is matched with the left parenthesis of 16 and the right parenthesis of 8 is matched with the left parenthesis of 15. The sizes of 5 and 8 in R are 2 and 4 respectively. We have $\mathrm{type}(L, R) = (6; 1, 4, 0, 1, 0, \dots, 0)$.

Let $\overline{P}(n)$ denote the set of elements $(L, R) \in P(n)$ such that the type $(b; b_1, b_2, \dots, b_n)$ of (L, R) satisfies $\sum_{i=1}^{n} ib_i < n$. Thus we have $(L, R) \in \overline{P}(n)$ if and only if there is at least one integer in the cyclic parenthesization of (L, R) which is not enclosed by any matching pair of parentheses.

We define a map τ from $\overline{P}(n)$ to the set of pairs (B, π), where $\pi \in \mathrm{NC}(n)$ and B is a block of π as follows. Let $(L, R) \in \overline{P}(n)$. Find a matching pair of parentheses in the cyclic parenthesization of (L, R) which do not enclose any other parenthesis. Remove the integers enclosed by these parentheses, and make a block of π with there integers. Repeat this procedure until there is no parenthesis. Since $(L, R) \in \overline{P}(n)$, we have several remaining integers after removing all the parentheses. These integers also form a block of π and B is defined to be this block.

Example 2. Let (L, R) be the pair in Example 1 represented by (11). Note that $(L, R) \in \overline{P}(16)$. Then $\tau(L, R) = (B, \pi)$, where π consists of the blocks $\{1, 16\}$, $\{2, 5\}$, $\{3, 4\}$, $\{6, 7, 8, 15\}$, $\{9\}$, $\{11, 12\}$ and $\{10, 13, 14\}$, and $B = \{10, 13, 14\}$.

Proposition 6. *The map τ is a bijection between $\overline{P}(n)$ and the set of pairs (B, π), where $\pi \in \mathrm{NC}(n)$ and B is a block of π. Moreover, if $\tau(L, R) = (B, \pi)$, $\mathrm{type}(\pi) = (b; b_1, b_2, \dots, b_n)$ and $|B| = j$, then* $\mathrm{type}(L, R) = (b - 1, b_1, \dots, b_j - 1, \dots, b_n)$.

We define $P(n, \ell)$ to be the set of $(\ell + 1)$-tuples $(L, R_1, R_2, \dots, R_\ell)$ such that $L, R_1, R_2, \dots, R_\ell \subset [n]$ and $|L| = |R_1| + |R_2| + \cdots + |R_\ell|$. Similarly, we can consider the *labeled cyclic parenthesization* of $(L, R_1, R_2, \dots, R_\ell)$ by placing a left parenthesis before i for each $i \in L$ and right parentheses $)_{j_1})_{j_2} \cdots)_{j_t}$ labeled with $j_1 < j_2 < \cdots < j_t$ after i if $R_{j_1}, R_{j_2}, \dots, R_{j_t}$ are the sets containing i among $R_1, R_2, \dots, R_\ell$. For each element $x \in R_i$, the *size of x* is defined in the same way as in the case of (L, R). We define the *type* of $(L, R_1, R_2, \dots, R_\ell)$ similarly to the type of (L, R).

Example 3. Let $T = (L, R_1, R_2) = (\{2, 4, 5\}, \{2\}, \{2, 6\}) \in P(7, 2)$. Then the labeled cyclic parenthesization of T is the following:

$$1\ (2)_1)_2\ 3\ (4\ (5\ 6)_2\ 7 \tag{12}$$

Then the size of $2 \in R_1$ is 1, the size of $2 \in R_2$ is 3 and the size of $6 \in R_2$ is 2. Thus the type of T is $(3; 1, 1, 1, 0, \dots, 0)$.

Lemma 7. *Let $b, b_1, b_2, \dots, b_n$ and $c_1, c_2, \dots, c_\ell$ be nonnegative integers with $b = b_1 + b_2 + \cdots + b_n = c_1 + c_2 + \cdots + c_\ell$. Then the number of elements $(L, R_1, R_2, \dots, R_\ell)$ in $P(n, \ell)$ with type $(b; b_1, b_2, \dots, b_n)$ and $|R_i| = c_i$ for $i \in [\ell]$ is equal to*

$$\binom{b}{b_1, b_2, \dots, b_n}\binom{n}{c_1}\binom{n}{c_2}\cdots\binom{n}{c_\ell}.$$

Let $\overline{P}(n, \ell)$ denote the set of $(L, R_1, R_2, \dots, R_\ell) \in P(n, \ell)$ such that the type $(b; b_1, b_2, \dots, b_n)$ of $(L, R_1, R_2, \dots, R_\ell)$ satisfies $\sum_{i=1}^{n} ib_i < n$.

Using τ, we define a map τ' from $\overline{P}(n, \ell)$ to the set of pairs $(B, \mathbf{c})$, where $\mathbf{c} : \pi_1 \leq \pi_2 \leq \cdots \leq \pi_\ell$ is a multichain in $\mathrm{NC}(n)$ and B is a block of π_1 as follows. Let $P = (L, R_1, R_2, \dots, R_\ell) \in \overline{P}(n, \ell)$. Applying the same argument as in the case of τ to the labeled cyclic parenthesization of P, we get (B_1, π_1). Then remove all the right parentheses in R_1 from the cyclic parenthesization and their corresponding left parentheses. By repeating this procedure, we get (B_i, π_i) for $i = 2, 3, \dots, \ell$. Then we obtain a multichain $\mathbf{c} : \pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}(n)$. We define $\tau'(P) = (B_1, \mathbf{c})$.

Proposition 8. *The map τ' is a bijection between $\overline{P}(n,\ell)$ and the set of pairs $(B, \mathbf{c})$ where $\mathbf{c} : \pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ is a multichain in $\mathrm{NC}(n)$ and B is a block of π_1. Moreover, if $\tau'(L, R_1, R_2, \dots, R_\ell) = (B, \mathbf{c})$, the rank jump vector of $\mathbf{c}$ is $(s_1, s_2, \dots, s_{\ell+1})$, $\mathrm{type}(\pi_1) = (b; b_1, b_2, \dots, b_n)$ and $|B| = j$, then the type of $(L, R_1, R_2, \dots, R_\ell)$ is $(b-1; b_1, \dots, b_j - 1, \dots, b_n)$ and $(|R_1|, |R_2|, \dots, |R_\ell|) = (s_2, s_3, \dots, s_{\ell+1})$.*

Theorem 9. *Let $b, b_1, b_2, \dots, b_n$ and $s_1, s_2, \dots, s_{\ell+1}$ be nonnegative integers satisfying $\sum_{i=1}^{n} b_i = b$, $\sum_{i=1}^{n} i \cdot b_i = n$, $\sum_{i=1}^{\ell+1} s_i = n - 1$ and $s_1 = n - b$. Then the number of multichains $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}(n)$ with rank jump vector $(s_1, s_2, \dots, s_{\ell+1})$, $\mathrm{type}(\pi_1) = (b; b_1, b_2, \dots, b_n)$ is equal to*

$$\frac{1}{b}\binom{b}{b_1, b_2, \dots, b_n}\binom{n}{s_2}\cdots\binom{n}{s_{\ell+1}}.$$

Proof. By Lemma 7 and Proposition 8, the number of pairs $(B, \mathbf{c})$, where $\mathbf{c}$ is a multichain satisfying the conditions and B is a block of π_1, is equal to

$$\sum_{j=1}^{n}\binom{b-1}{b_1, \dots, b_j - 1, \dots, b_n}\binom{n}{s_2}\cdots\binom{n}{s_{\ell+1}} = \binom{b}{b_1, b_2, \dots, b_n}\binom{n}{s_2}\cdots\binom{n}{s_{\ell+1}}.$$

Since there are $b = \mathrm{bk}(\pi_1)$ choices of B for each $\mathbf{c}$, we get the theorem. Note that Lemma 7 states the number of elements in $P(n,\ell)$. However, by the condition on the type, all the elements in consideration are in $\overline{P}(n,\ell)$. □

Now we can prove Theorem 1.

Proof of Theorem 1. Let $\pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ be a multichain in $\mathrm{NC}^{(k)}(n)$ with rank jump vector $(s_1, s_2, \dots, s_{\ell+1})$ and $\mathrm{type}^{(k)}(\pi_1) = (b; b_1, b_2, \dots, b_n)$. Then this is a multichain in $\mathrm{NC}(kn)$ with rank jump vector $(kn - 1 - b, s_2, \dots, s_{\ell+1})$ and $\mathrm{type}(\pi_1) = (b; b'_1, b'_2, \dots, b'_{kn})$ where $b'_{ki} = b_i$ for $i \in [n]$ and $b'_j = 0$ if j is not divisible by k. By Theorem 9, the number of such multichains is equal to

$$\frac{1}{b}\binom{b}{b'_1, b'_2, \dots, b'_{kn}}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}} = \frac{1}{b}\binom{b}{b_1, b_2, \dots, b_n}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}}.$$

□

4.2 Augmented k-divisible noncrossing partitions of type A

If all the block sizes of a partition π are divisible by k then the size of π must be divisible by k. Thus k-divisible partitions can be defined only on $[kn]$. We extend this definition to partitions of size not divisible by k as follows.

Let k and r be integers with $0 < r < k$. A partition π of $[kn + r]$ is *augmented k-divisible* if the sizes of all but one of the blocks are divisible by k.

Let $\mathrm{NC}^{(k)}(n; r)$ denote the subposet of $\mathrm{NC}(kn + r)$ consisting of the augmented k-divisible noncrossing partitions. Then $\mathrm{NC}^{(k)}(n; r)$ is a graded poset with the rank function $\mathrm{rank}(\pi) = n - \mathrm{bk}'(\pi)$, where $\mathrm{bk}'(\pi)$ is the number of blocks of π whose sizes are divisible by k. We define $\mathrm{type}^{(k)}(\pi)$ to be $(b; b_1, b_2, \dots, b_n)$ where b_i is the number of blocks B of size ki and $b = b_1 + b_2 + \cdots + b_n$.

Theorem 10. *Let $0 < r < k$. Let $b, b_1, b_2, \dots, b_n$ and $s_1, s_2, \dots, s_{\ell+1}$ be nonnegative integers satisfying $\sum_{i=1}^n b_i = b$, $\sum_{i=1}^n i \cdot b_i \le n$, $\sum_{i=1}^{\ell+1} s_i = n$ and $s_1 = n - b$. Then the number of multichains $\mathbf{c} : \pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}^{(k)}(n; r)$ with rank jump vector $(s_1, s_2, \dots, s_{\ell+1})$ and $\mathrm{type}^{(k)}(\pi_1) = (b; b_1, b_2, \dots, b_n)$ is equal to*

$$\binom{b}{b_1, b_2, \dots, b_n}\binom{kn+r}{s_2}\cdots\binom{kn+r}{s_{\ell+1}}.$$

5 k-divisible noncrossing partitions of type B

Let $\pi \in \mathrm{NC}_B(kn)$. We say that π is a *k-divisible noncrossing partition of type B_n* if the size of each block of π is divisible by k.

Let $\mathrm{NC}_B^{(k)}(n)$ denote the subposet of $\mathrm{NC}_B(kn)$ consisting of k-divisible noncrossing partitions of type B_n. Then $\mathrm{NC}_B^{(k)}(n)$ is a graded poset with the rank function $\mathrm{rank}(\pi) = n - \mathrm{nz}(\pi)$, where $\mathrm{nz}(\pi)$ denotes the number of unordered pairs $(B, -B)$ of nonzero blocks of π.

We can prove Theorem 2 using a similar method in the proof of Theorem 1. Instead of doing this, we will prove the following lemma which implies Theorem 2. Note that the following lemma is also a generalization of (4).

For a multichain $\mathbf{c} : \pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}_B^{(k)}(n)$, the *index* $\mathrm{ind}(\mathbf{c})$ of $\mathbf{c}$ is the smallest integer i such that π_i has a zero block. If there is no such integer i, then $\mathrm{ind}(\mathbf{c}) = \ell + 1$.

Lemma 11. *Let $b, b_1, b_2, \dots, b_n$ and $s_1, s_2, \dots, s_{\ell+1}$ be nonnegative integers satisfying $\sum_{i=1}^n b_i = b$, $\sum_{i=1}^n i \cdot b_i \le n$, $\sum_{i=1}^{\ell+1} s_i = n$ and $s_1 = n - b$. Then the number of multichains $\mathbf{c} : \pi_1 \le \pi_2 \le \cdots \le \pi_\ell$ in $\mathrm{NC}_B^{(k)}(n)$ with rank jump vector $(s_1, s_2, \dots, s_{\ell+1})$, $\mathrm{type}^{(k)}(\pi_1) = (b; b_1, b_2, \dots, b_n)$ and $\mathrm{ind}(\mathbf{c}) = i$ is equal to*

$$\binom{b}{b_1, b_2, \dots, b_n}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}},$$

if $i = 1$, and

$$\frac{s_i}{b}\binom{b}{b_1, b_2, \dots, b_n}\binom{kn}{s_2}\cdots\binom{kn}{s_{\ell+1}},$$

if $i \ge 2$.

5.1 Armstrong's conjecture

Let $\widetilde{\mathrm{NC}}^{(k)}(n)$ denote the subposet of $\mathrm{NC}^{(k)}(n)$ whose elements are fixed under the $180°$ rotation in the circular representation.

Armstrong [1, Conjecture 4.5.14] conjectured the following. Let n and k be integers such that n is even and k is arbitrary, or n is odd and k is even. Then

$$Z(\widetilde{\mathrm{NC}}^{(k)}(n), \ell) = \binom{\lfloor (k\ell+1)n/2 \rfloor}{\lfloor n/2 \rfloor}.$$

If n is even then $\widetilde{\mathrm{NC}}^{(k)}(n)$ is isomorphic to $\mathrm{NC}_B^{(k)}(n/2)$, whose zeta polynomial is already known. If both n and k are odd, then $\widetilde{\mathrm{NC}}^{(k)}(n)$ is empty. Thus the conjecture is only for n and k such that n is odd and k is even.

Armstrong's conjecture is a consequence from the following theorem and Theorem 10.

Theorem 12. *Let n and k be positive integers. Then*

$$\widetilde{\mathrm{NC}}^{(2k)}(2n+1) \cong \mathrm{NC}^{(2k)}(n;k).$$

References

[1] Drew Armstrong, *Generalized noncrossing partitions and combinatorics of Coxeter groups*, Mem. Amer. Math. Soc. **2002** (2009), no. 949.

[2] Christos A. Athanasiadis, *On noncrossing and nonnesting partitions for classical reflection groups*, Electron. J. Combin. **5** (1998), R42.

[3] Christos A. Athanasiadis and Victor Reiner, *Noncrossing partitions for the group D_n*, SIAM J. Discrete Math. **18** (2005), no. 2, 397–417.

[4] David Bessis, *The dual braid monoid*, Annales scientifiques de l'Ecole normale supérieure **36** (2003), 647–683.

[5] T. Brady and C. Watt, *Non-crossing partition lattices in finite reflection groups*, Trans. Amer. Math. Soc. **360** (2008), 1983–2005.

[6] Paul H. Edelman, *Chain enumeration and non-crossing partitions*, Discrete Math. **31** (1980), 171–180.

[7] Jang Soo Kim, *New interpretations for noncrossing partitions of classical types*, `http://arxiv.org/abs/0910.2036`.

[8] Jang Soo Kim, *Chain enumeration of k-divisible noncrossing partitions of classical types*, `http://arxiv.org/abs/0908.2641`.

[9] C. Krattenthaler and T.W. Müller, *Decomposition numbers for finite Coxeter groups and generalised non-crossing partitions*, `http://arxiv.org/abs/0704.0199`.

[10] Christian Krattenthaler, *Non-crossing partitions on an annulus*, in preparation.

[11] G. Kreweras, *Sur les partitions non croisées d'un cycle*, Discrete Math. **1** (1972), 333–350.

[12] V. Reiner, *Non-crossing partitions for classical reflection groups*, Discrete Math. **177** (1997), 195–222.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 690–701

Enumerating (2+2)-free posets by the number of minimal elements and other statistics

Sergey Kitaev[1†] and Jeffrey Remmel[2‡]

[1]*The Mathematics Institute, School of Computer Science, Reykjavík University, IS-103 Reykjavík, Iceland*
[2]*Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112. USA*

Abstract. A poset is said to be $(\mathbf{2}+\mathbf{2})$-free if it does not contain an induced subposet that is isomorphic to $\mathbf{2}+\mathbf{2}$, the union of two disjoint 2-element chains. In a recent paper, Bousquet-Mélou et al. found, using so called ascent sequences, the generating function for the number of $(\mathbf{2}+\mathbf{2})$-free posets: $P(t) = \sum_{n\geq 0}\prod_{i=1}^{n}\left(1-(1-t)^i\right)$. We extend this result by finding the generating function for $(\mathbf{2}+\mathbf{2})$-free posets when four statistics are taken into account, one of which is the number of minimal elements in a poset. We also show that in a special case when only minimal elements are of interest, our rather involved generating function can be rewritten in the form $P(t,z) = \sum_{n,k\geq 0} p_{n,k}t^n z^k = 1+\sum_{n\geq 0}\frac{zt}{(1-zt)^{n+1}}\prod_{i=1}^{n}(1-(1-t)^i)$ where $p_{n,k}$ equals the number of $(\mathbf{2}+\mathbf{2})$-free posets of size n with k minimal elements.

Résumé. Un poset sera dit $(\mathbf{2}+\mathbf{2})$-libre s'il ne contient aucun sous-poset isomorphe à $\mathbf{2}+\mathbf{2}$, l'union disjointe de deux chaînes à deux éléments. Dans un article récent, Bousquet-Mélou et al. ont trouvé, à l'aide de "suites de montées", la fonction génératrice des nombres de posets $(\mathbf{2}+\mathbf{2})$-libres: c'est $P(t) = \sum_{n\geq 0}\prod_{i=1}^{n}\left(1-(1-t)^i\right)$. Nous étendons ce résultat en trouvant la fonction génératrice des posets $(\mathbf{2}+\mathbf{2})$-libres rendant compte de quatre statistiques, dont le nombre d'éléments minimaux du poset. Nous montrons aussi que lorsqu'on ne s'intéresse qu'au nombre d'éléments minimaux, notre fonction génératrice assez compliquée peut être simplifiée en $P(t,z) = \sum_{n,k\geq 0} p_{n,k}t^n z^k = 1+\sum_{n\geq 0}\frac{zt}{(1-zt)^{n+1}}\prod_{i=1}^{n}(1-(1-t)^i)$, où $p_{n,k}$ est le nombre de posets $(\mathbf{2}+\mathbf{2})$-libres de taille n avec k éléments minimaux.

Keywords: (2+2)-free posets, minimal elements, generating function

1 Introduction

A poset is said to be $(\mathbf{2}+\mathbf{2})$-*free* if it does not contain an induced subposet that is isomorphic to $\mathbf{2}+\mathbf{2}$, the union of two disjoint 2-element chains. We let $\mathcal{P}$ denote the set of $(\mathbf{2}+\mathbf{2})$-free posets. Fishburn [7] showed that a poset is $(\mathbf{2}+\mathbf{2})$-free precisely when it is isomorphic to an *interval order*. Bousquet-Mélou et al. [1] showed that the generating function for the number p_n of $(\mathbf{2}+\mathbf{2})$-free posets on n elements is

$$P(t) = \sum_{n\geq 0} p_n\, t^n = \sum_{n\geq 0}\prod_{i=1}^{n}\left(1-(1-t)^i\right). \tag{1}$$

†The work presented here was supported by grant no. 090038011 from the Icelandic Research Fund.
‡Partially supported by NSF grant DMS 0654060.

In fact, El-Zahar [4] and Khamis [9] used a recursive description of $(\mathbf{2}+\mathbf{2})$-free posets, different from that of [1], to derive a pair of functional equations that define the series $P(t)$. However, they did not solve these equations. Haxell, McDonald and Thomasson [8] provided an algorithm, based on a complicated recurrence relation, to produce the first numbers p_n. Moreover, the above series was proved by Zagier [12] to count certain involutions introduced by Stoimenow [10]. Bousquet-Mélou et al. [1] gave a bijection between $(\mathbf{2}+\mathbf{2})$-free posets and the involutions, as well as a certain class of restricted permutations and so called *ascent sequences*. Given an integer sequence $(x_1, \ldots, x_i)$, the number of *ascents* of this sequence is

$$\mathrm{asc}(x_1, \ldots, x_i) = |\{\, 1 \le j < i \,:\, x_j < x_{j+1} \,\}|.$$

A sequence $(x_1, \ldots, x_n) \in \mathbb{N}^n$ an *ascent sequence of length* n if it satisfies $x_1 = 0$ and $x_i \in [0, 1 + \mathrm{asc}(x_1, \ldots, x_{i-1})]$ for all $2 \le i \le n$. For instance, (0, 1, 0, 2, 3, 1, 0, 0, 2) is an ascent sequence. We let $\mathcal{A}$ denote the set of all ascent sequences (we assume the empty word to be an ascent sequence).

Amongst other results concerning $(\mathbf{2}+\mathbf{2})$-free posets [5, 6], the following characterization plays an important role in [1]: a poset is $(\mathbf{2}+\mathbf{2})$-free if and only if the collection of strict principal down-sets (for an element, a down-set is the set of its predecessors) can be linearly ordered by inclusion [6]. Here for any poset $\mathcal{P} = (P, <_p)$ and $x \in P$, the strict principal down set of x, $D(x)$, in $\mathcal{P}$ is the set of all $y \in P$ such that $y <_p x$. The *trivial down-set* is the empty set. Thus if $\mathcal{P}$ is a $(\mathbf{2}+\mathbf{2})$-free poset, we can write $D(P) = \{D(x) : x \in P\}$ as

$$D(P) = \{D_0, D_1, \ldots, D_k\}$$

where $\emptyset = D_0 \subset D_1 \subset \cdots \subset D_k$. In such a situation, we say that $x \in P$ has level i if $D(x) = D_i$.

Bousquet-Mélou et al. [1] described a decomposition of a $(\mathbf{2}+\mathbf{2})$-free poset removing at each step a maximal element located on the lowest level, together with certain relations. Recording the levels from which we just removed a maximal element, and reading the obtained sequence backwards after removing all the elements, one obtains an ascent sequence. This gives a bijection between $(\mathbf{2}+\mathbf{2})$-free posets and ascent sequences. We note that in the process of decomposing a $(\mathbf{2}+\mathbf{2})$-free poset, element by element, at some point, the current poset will be a (possibly 1-element) antichain. The statistic lds is defined as the size of the (maximum) antichain in the last sentence, which is the size of the down-set of the last removed element that has a non-trivial down-set. By definition, the value of lds on an antichain is 0 (there are no non-trivial down-sets there). We refer to [1, Section 3] for the detailed description of the decomposition, as it is rather space-consuming to state here.

Bousquet-Mélou et al. [1] studied a more general generating function $F(t, u, v)$ of $(\mathbf{2}+\mathbf{2})$-free posets, which are counted by size="number of elements" (variable t), levels="*number of levels*" defined below (variable u), and minmax="level of minimum maximal element" (variable v). The first few terms of $F(t, u, v)$ are

$$F(t, u, v) = 1 + t + (1 + uv)t^2 + (1 + 2uv + u + u^2v^2)t^3 + O(t^4).$$

An explicit form of $F(t, u, v)$ can be obtained from [1, Lemma 13] and [1, Proposition 14]. The main result of this paper, Theorem 4, is an explicit form of the generating function $G(t, u, v, z, x)$ for a generalization of $F(t, u, v)$, when two more statistics are taken into account — min="number of minimal elements" in a poset (variable z) and lds="size of *non-trivial last down-set*" (variable x). That is, we shall study the following generating function:

$$G(t, u, v, z, x) = \sum_{p \in \mathcal{P}} t^{\mathrm{size}(p)} u^{\mathrm{levels}(p)} v^{\mathrm{minmax}(p)} z^{\mathrm{min}(p)} x^{\mathrm{lds}(p)}.$$

Reduction of the main problem to considering ascent sequences. The basic idea used by Bousquet-Mélou et al. [1] to find the generating function $F(t,u,v)$ was to reduce the problem to counting ascent sequences using their bijection between $(\mathbf{2}+\mathbf{2})$-free posets and ascent sequences. We follow a similar strategy to find $G(t,u,v,z,x)$. That is, we define the following statistics on an ascent sequence: length="the number of elements in the sequence," last="the rightmost element of the sequence," zeros="the number of 0's in the sequence," run="the number of elements in the leftmost run of 0's"="the number of 0's to the left of the leftmost non-zero element." By definition, if there are no non-zero elements in an ascent sequence, the value of run is 0.

Lemma 1 *The function $G(t,u,v,z,x)$ defined above can alternatively be defined on ascent sequences as*

$$G(t,u,v,z,x)=\sum_{w\in\mathcal{A}} t^{\mathrm{length}(w)}u^{\mathrm{asc}(w)}v^{\mathrm{last}(w)}z^{\mathrm{zeros}(w)}x^{\mathrm{run}(w)}=\sum_{n,a,\ell,m,r\geq 0} G_{n,a,\ell,m,r}t^n u^a v^\ell z^m x^r.$$

Proof: To prove the statement we need to show equidistribution of the statistics involved. All but one case follow from the results in [1]. More precisely, we can use the bijection from $(\mathbf{2}+\mathbf{2})$-free posets to ascent sequences presented in [1] which sends $\mathrm{size}\to\mathrm{length}$, $\mathrm{levels}\to\mathrm{asc}$, $\mathrm{minmax}\to\mathrm{last}$, and $\mathrm{min}\to\mathrm{zeros}$.

The fact that lds goes to run follows from the definition of the statistics and the idea of the bijection in [1] described above. Indeed, while recording levels of just removed elements, after we removed the element, say e, whose down-set gives lds, we will be left with incomparable elements located on level 0, which gives in the corresponding ascent sequence the initial run of 0's followed by 1 corresponding to e located on level 1. □

Note that $G(t,u,v,1,1)=F(t,u,v)$ as studied in [1].

Organization of the paper. In Section 2 we find explicitly the function $G=G(t,u,v,z,x)$ using ascent sequences (see Theorem 4). In Section 3 we show that in a special case when only minimal elements are of interest, a rather involved generating function $G(t,u,v,z,x)$ can be rewritten in the form

$$P(t,z)=\sum_{n,k\geq 0} p_{n,k}t^n z^k = 1+\sum_{n\geq 0}\frac{zt}{(1-zt)^{n+1}}\prod_{i=1}^{n}(1-(1-t)^i)$$

where $p_{n,k}$ equals the number of $(\mathbf{2}+\mathbf{2})$-free posets of size n with k minimal elements. We shall see that our expression for $P(t,z)$ cannot be directly derived from $G(t,u,v,z,x)$ by substituting 1 for the variables u, v, and x.

2 Main results

For $r\geq 1$, let $G_r(t,u,v,z)$ denote the coefficient of x^r in $G(t,u,v,z,x)$. Thus $G_r(t,u,v,z)$ is the generating function of those ascent sequences that begin with $r\geq 1$ 0's followed by 1. We let $G^r_{a,l,m,n}$ denote the number of ascent sequences of length n which begin with r 0's followed by 1, have a ascents, the last letter ℓ, and a total of m zeros. We then let

$$G_r := G_r(t,u,v,z)=\sum_{a,\ell,m\geq 0,n\geq r+1} G^r_{a,l,m,n}t^n u^a v^\ell z^m. \qquad (2)$$

Clearly, since the sequence $0\dots 0$ has no ascents and no initial run of 0's (by definition), we have that the generating function for such sequences is

$$1+tz+(tz)^2+\cdots=\frac{1}{1-tz}$$

where 1 corresponds to the empty word. Thus, we have the following relation between G and G_r:

$$G=\frac{1}{1-tz}+\sum_{r\geq 1}G_r\,x^r. \tag{3}$$

Lemma 2 *For $r\geq 1$, the generating function $G_r(t,u,v,z)$ satisfies*

$$(v-1-tv(1-u))G_r=(v-1)t^{r+1}uvz^r+t((v-1)z-v)G_r(t,u,1,z)+tuv^2G_r(t,uv,1,z). \tag{4}$$

Proof:

Our proof follows the same steps as in Lemma 13 in [1]. Fix $r\geq 1$. Let $x'=(x_1,\dots,x_{n-1})$ be an ascent sequence beginning with r 0's followed by 1, with a ascents and m zeros where $x_{n-1}=\ell$. Then $x=(x_1,\dots,x_{n-1},i)$ is an ascent sequence if and only if $i\in[0,a+1]$. Clearly x also begins with r 0's followed by 1. Now, if $i=0$, the sequence x has a ascents and $m+1$ zeros. If $1\leq i\leq\ell$, x has a ascents and m zeros. Finally if $i\in[\ell+1,a+1]$, then x has $a+1$ ascents and m zeros. Counting the sequence $0\dots01$ with r 0's separately, we have

$$\begin{aligned}
G_r &= t^{r+1}u^1v^1z^r+\sum_{\substack{a,\ell,m\geq 0\\ n\geq r+1}}G^r_{a,\ell,m,n}t^{n+1}\left(u^av^0z^{m+1}+\sum_{i=1}^{\ell}u^av^iz^m+\sum_{i=\ell+1}^{a+1}u^{a+1}v^iz^m\right)\\
&= t^{r+1}uvz^r+t\sum_{\substack{a,\ell,m\geq 0\\ n\geq r+1}}G_{a,\ell,m,n}t^nu^az^m\left(z+\frac{v^{\ell+1}-v}{v-1}+u\frac{v^{a+2}-v^{\ell+1}}{v-1}\right)\\
&= t^{r+1}uvz^r+tzG_r(t,u,1,z)+tv\frac{G_r-G_r(t,u,1,z)}{v-1}+tuv\frac{vG_r(t,uv,1,z)-G_r}{v-1}.
\end{aligned}$$

The result follows. □

Next just like in Subsection 6.2 of [1], we use the kernel method to proceed. Setting $(v-1-tv(1-u))=0$ and solving for v, we obtain that the substitution $v=1/(1+t(u-1))$ will kill the left-hand side of (4). We can then solve for $G_r(t,u,1,z)$ to obtain that

$$G_r(t,u,1,z)=\frac{(1-u)t^{r+1}uz^r+uG_r\left(t,\frac{u}{1+t(u-1)},1,z\right)}{(1+zt(u-1))(1+t(u-1))}. \tag{5}$$

Next we define

$$\delta_k = u-(1-t)^k(u-1)\text{ and} \tag{6}$$
$$\gamma_k = u-(1-zt)(1-t)^{k-1}(u-1) \tag{7}$$

for $k \geq 1$. We also set $\delta_0 = \gamma_0 = 1$. Observe that $\delta_1 = u - (1-t)(u-1) = 1 + t(u-1)$ and $\gamma_1 = u - (1-zt)(u-1) = 1 + zt(u-1)$. Thus we can rewrite (5) as

$$G_r(t,u,1,z) = \frac{t^{r+1}z^r u(1-u)}{\delta_1\gamma_1} + \frac{u}{\delta_1\gamma_1} G_r(t, \frac{u}{\delta_1}, 1, z). \tag{8}$$

For any function of $f(u)$, we shall write $f(u)|_{u=\frac{u}{\delta_k}}$ for $f(u/\delta_k)$. It is then easy to check that

1. $(u-1)|_{u=\frac{u}{\delta_k}} = \dfrac{(1-t)^k(u-1)}{\delta_k}$,
2. $\delta_s|_{u=\frac{u}{\delta_k}} = \dfrac{\delta_{s+k}}{\delta_k}$,
3. $\gamma_s|_{u=\frac{u}{\delta_k}} = \dfrac{\gamma_{s+k}}{\delta_k}$, and
4. $\dfrac{u}{\delta_s}|_{u=\frac{u}{\delta_k}} = \dfrac{u}{\delta_{s+k}}$.

Using these relations, one can iterate the recursion (8) to prove by induction that for all $n \geq 1$,

$$\begin{aligned} G_r(t,u,1,z) \quad &= \quad \frac{t^{r+1}z^r u(1-u)}{\delta_1\gamma_1} + \left(t^{r+1}z^r u(1-u) \sum_{s=2}^{2^n-1} \frac{u^s(1-t)^s}{\delta_s\delta_{s+1}\prod_{i=1}^{s+1}\gamma_i} \right) + \\ &\qquad \frac{u^{2^n}}{\delta_{2^n}\prod_{i=1}^{2^n}\gamma_i} G_r(t, \frac{u}{\delta_{2^n}}, 1, z). \end{aligned} \tag{9}$$

Since $\delta_0 = 1$, it follows that as a power series in u,

$$G_r(t,u,1,z) = t^{r+1}z^r u(1-u) \sum_{s\geq 0} \frac{u^s(1-t)^s}{\delta_s\delta_{s+1}\prod_{i=1}^{s+1}\gamma_i}. \tag{10}$$

We have used Mathematica to compute that

$$\begin{aligned} &G_1(t,u,1,z) = uzt^2 + \left(uz + u^2z + uz^2\right)t^3 \\ &+ \left(uz + 3u^2z + u^3z + uz^2 + 3u^2z^2 + uz^3\right)t^4 \\ &+ \left(uz + 6u^2z + 7u^3z + u^4z + uz^2 + 8u^2z^2 + 7u^3z^2 + uz^3 + 5u^2z^3 + uz^4\right)t^5 + O[t]^6. \end{aligned}$$

For example, the coefficient of t^4u^2, $3z + 3z^2$ makes sense as there are 3 ascent sequences of length 4 with 2 ascents and 1 zero, namely, 0112, 0121, and 0122, while there are 3 ascent sequences of length 4 with 2 ascents and 2 zeros, namely, 0101, 0102, and 0120 (there are no other ascents sequences of length 4 with 2 ascents).

Note that we can rewrite (4) as

$$G_r(t,u,v,z) = \frac{t^{r+1}z^r uv(1-v)}{v\delta_1 - 1} + \frac{t(z(v-1)-v)}{v\delta_1 - 1} G_r(t,u,1,z) + \frac{uv^2t}{v\delta_1 - 1} G_r(t,uv,1,z). \tag{11}$$

For $s \geq 1$, we let

$$\begin{aligned}\bar{\delta}_s &= \delta_s|_{u=uv} = uv - (1-t)^s(uv-1) \text{ and} \\ \bar{\gamma}_s &= \gamma_s|_{u=uv} = uv - (1-zt)(1-t)^{s-1}(uv-1)\end{aligned}$$

and set $\bar{\delta}_0 = \bar{\gamma}_0 = 1$. Then using (11) and (10), we have the following theorem.

Theorem 3 *For all $r \geq 1$,*

$$\begin{aligned}G_r(t,u,v,z) &= \frac{t^{r+1}z^r u}{v\delta_1 - 1}\left(v(v-1) + t(1-u)(z(v-1)-v)\sum_{s\geq 0}\frac{u^s(1-t)^s}{\delta_s\delta_{s+1}\prod_{i=1}^{s+1}\gamma_i}\right. \\ &\left. + uv^3t(1-uv)\sum_{s\geq 0}\frac{(uv)^s(1-t)^s}{\bar{\delta}_s\bar{\delta}_{s+1}\prod_{i=1}^{s+1}\bar{\gamma}_i}\right) \qquad (12)\end{aligned}$$

It is easy to see from Theorem 3 that

$$G_r(t,u,v,z) = t^{r-1}z^{r-1}G_1(t,u,v,z). \qquad (13)$$

This is also easy to see combinatorially since every ascent sequence counted by $G_r(t,u,v,z)$ is of the form $0^{r-1}a$ where a is an ascent sequence a counted by $G_1(t,u,v,z)$.

We have used Mathematica to compute that

$$\begin{aligned}&G_1(t,u,v,z) = uvzt^2 + \left(uvz + u^2v^2z + uz^2\right)t^3 \\ &+ \left(uvz + u^2vz + 2u^2v^2z + u^3v^3z + uz^2 + u^2z^2 + u^2vz^2 + u^2v^2z^2 + uz^3\right)t^4 \\ &+ \left(uvz + 3u^2vz + u^3vz + 3u^2v^2z + 2u^3v^2z + 4u^3v^3z + u^4v^4z + uz^2 + 3u^2z^2 + u^3z^2 + 3u^2vz^2\right. \\ &\left. + u^3vz^2 + 2u^2v^2z^2 + 2u^3v^2z^2 + 3u^3v^3z^2 + uz^3 + 3u^2z^3 + u^2vz^3 + u^2v^2z^3 + uz^4\right)t^5 + O[t]^6.\end{aligned}$$

For example, the coefficient of t^4u is $zv + z^2 + z^3$ which makes sense since the sequences counted by the terms are 0111, 0110, and 0100, respectively.

Note that

$$\begin{aligned}G(t,u,v,z,x) &= \frac{1}{(1-tz)} + \sum_{r\geq 1} G_r(t,u,v,z)x^r \\ &= \frac{1}{(1-tz)} + \sum_{r\geq 1} t^{r-1}z^{r-1}G_1(t,u,v,z)x^r \\ &= \frac{1}{(1-tz)} + \frac{1}{1-tzx}xG_1(t,u,v,z)\end{aligned}$$

Thus we have the following theorem.

Theorem 4

$$G(t,u,v,z,x) = \frac{1}{(1-tz)} + \frac{t^2zxu}{(1-tzx)(v\delta_1-1)}\Bigg(v(v-1)$$
$$+ t(1-u)(z(v-1)-v)\sum_{s\geq 0}\frac{u^s(1-t)^s}{\delta_s\delta_{s+1}\prod_{i=1}^{s+1}\gamma_i} + uv^3t(1-uv)\sum_{s\geq 0}\frac{(uv)^s(1-t)^s}{\bar{\delta}_s\bar{\delta}_{s+1}\prod_{i=1}^{s+1}\bar{\gamma}_i}\Bigg). \quad (14)$$

Again, we have used Mathematica to compute the first few terms of this series:

$$G(t,u,v,z,x) = 1 + zt + \left(uvxz + z^2\right)t^2 + \left(uvxz + u^2v^2xz + uxz^2 + uvx^2z^2 + z^3\right)t^3$$
$$+ \big(uvxz + u^2vxz + 2u^2v^2xz + u^3v^3xz + uxz^2 + u^2xz^2 + u^2vxz^2$$
$$+u^2v^2xz^2 + uvx^2z^2 + u^2v^2x^2z^2 + uxz^3 + ux^2z^3 + uvx^3z^3 + z^4\big)t^4$$
$$\big(uvxz + 3u^2vxz + u^3vxz + 3u^2v^2xz + 2u^3v^2xz + 4u^3v^3xz + u^4v^4xz$$
$$+uxz^2 + 3u^2xz^2 + u^3xz^2 + 3u^2vxz^2 + u^3vxz^2 + 2u^2v^2xz^2 + 2u^3v^2xz^2 + 3u^3v^3xz^2$$
$$+uvx^2z^2 + u^2vx^2z^2 + 2u^2v^2x^2z^2 + u^3v^3x^2z^2 + uxz^3 + 3u^2xz^3 + u^2vxz^3 + u^2v^2xz^3$$
$$+ux^2z^3 + u^2x^2z^3 + u^2vx^2z^3 + u^2v^2x^2z^3 + uvx^3z^3 + u^2v^2x^3z^3 + uxz^4$$
$$+ux^2z^4 + ux^3z^4 + uvx^4z^4 + z^5\big)t^5 + O[t]^6.$$

One can check that, for instance, the 3 sequences corresponding to the term $3u^2v^2xzt^5$ are 01112, 01122 and 01222.

3 Counting $(\mathbf{2}+\mathbf{2})$-free posets by size and number of minimal elements

In this section, we shall compute the generating function of $(\mathbf{2}+\mathbf{2})$-free posets by size and the number of minimal elements which is equivalent to finding the generating function for ascent sequences by length and the number of zeros.

For $n \geq 1$, let $H_{a,b,\ell,n}$ denote the number of ascent sequences of length n with a ascents and b zeros which have last letter ℓ. Then we first wish to compute

$$H(u,z,v,t) = \sum_{n\geq 1, a,b,\ell\geq 0} H_{a,b,\ell,n}u^az^bv^\ell t^n. \quad (15)$$

Using the same reasoning as in the previous section, we see that

$$\begin{aligned}
H(u,z,v,t) &= tz + \sum_{\substack{a,b,\ell\geq 0\\ n\geq 1}} H_{a,b,\ell,n}t^{n+1}\left(u^av^0z^{b+1} + \sum_{i=1}^{\ell}u^av^iz^b + \sum_{i=\ell+1}^{a+1}u^{a+1}v^iz^b\right)\\
&= tz + t\sum_{\substack{a,b,\ell\geq 0\\ n\geq r+1}} H_{a,b\ell,n}t^nu^az^b\left(z + \frac{v^{\ell+1}-v}{v-1} + u\frac{v^{a+2}-v^{\ell+1}}{v-1}\right)\\
&= tz + \frac{tv(1-u)}{v-1}H(u,v,z,t) + \frac{t(z(v-1)-v)}{v-1}H(u,1,z,t) + \frac{tuv^2}{v-1}H(uv,1,z,t).
\end{aligned}$$

Thus we have the following lemma.

Lemma 5

$$(v-1-tv(1-u))H(u,v,z,t) = tz(v-1) + t(z(v-1)-v)H(u,1,z,t) + tuv^2H(uv,1,z,t). \quad (16)$$

Setting $(v-1-t(1-u)) = 0$, we see that the substitution $v = 1+t(u-1) = \delta_1$ kills the left-hand side of (16). We can then solve for $H(u,1,z,t)$ to obtain the recursion

$$H(u,1,z,t) = \frac{zt(1-u)}{\gamma_1} + \frac{u}{\delta_1\gamma_1}H(uv,1,z,t). \quad (17)$$

By iterating (17), we can prove by induction that for all $n \geq 1$,

$$H(u,1,z,t) = \frac{zt(1-u)}{\gamma_1} + \left(\sum_{s=1}^{2^n-1} \frac{zt(1-u)u^s(1-t)^s}{\delta_s \prod_{i=1}^{s+1}\gamma_i}\right) + \frac{u^{2^n}}{\delta_{2^n}\prod_{i=1}^{2^n}\gamma_i}H(\frac{u}{\delta_{2^n}},1,z,t). \quad (18)$$

Since $\delta_0 = 1$, we can rewrite (18) as

$$H(u,1,z,t) = \left(\sum_{s=0}^{2^n-1} \frac{zt(1-u)u^s(1-t)^s}{\delta_s \prod_{i=1}^{s+1}\gamma_i}\right) + \frac{u^{2^n}}{\delta_{2^n}\prod_{i=1}^{2^n}\gamma_i}H(\frac{u}{\delta_{2^n}},1,z,t). \quad (19)$$

Thus as a power series in u, we can conclude the following.

Theorem 6

$$H(u,1,z,t) = \sum_{s=0}^{\infty} \frac{zt(1-u)u^s(1-t)^s}{\delta_s \prod_{i=1}^{s+1}\gamma_i}. \quad (20)$$

We would like to set $u=1$ in the power series $\sum_{s=0}^{\infty} \frac{zt(1-u)u^s(1-t)^s}{\delta_s \prod_{i=1}^{s+1}\gamma_i}$, but the factor $(1-u)$ in the series does not allow us to do that in this form. Thus our next step is to rewrite the series in a form where it is obvious that we can set $u=1$ in the series. To that end, observe that for $k \geq 1$,

$$\delta_k = u-(1-t)^k(u-1) = 1+u-1-(1-t)^k(u-1) = 1-(1-t)^k-1)(u-1)$$

so that

$$\frac{1}{\delta_k} = \sum_{n\geq 0}((1-t)^k-1)^n(u-1)^n \sum_{n\geq 0}(u-1)^n = \sum_{m=0}^{n}(-1)^{n-m}\binom{n}{m}(1-t)^{km}. \quad (21)$$

Substituting (21) into (20), we see that

$$\begin{aligned}
H(u,1,z,t) &= \frac{zt(1-u)}{\gamma_1} + \sum_{k\geq 1} \frac{zt(1-u)u^k(1-t)^k}{\prod_{i=1}^{k+1}\gamma_i} \sum_{n\geq 0}(u-1)^n \sum_{m=0}^{n}(-1)^{n-m}\binom{n}{m}(1-t)^{km} \\
&= \frac{zt(1-u)}{\gamma_1} + \sum_{n\geq 0}\sum_{m=0}^{n}(-1)^{n-m-1}\binom{n}{m}(u-1)^{n-m} zt \sum_{k\geq 1}\frac{(u-1)^{m+1}u^k(1-t)^{k(m+1)}}{\prod_{i=1}^{k+1}\gamma_i} \\
&= \frac{zt(1-u)}{\gamma_1} + \sum_{n\geq 0}\sum_{m=0}^{n}(-1)^{n-m-1}\binom{n}{m}(u-1)^{n-m}\frac{zt}{(1-zt)^{m+1}} \times \\
&\quad \sum_{k\geq 1}\frac{(u-1)^{m+1}(1-zt)^{m+1}u^k(1-t)^{k(m+1)}}{\prod_{i=1}^{k+1}\gamma_i}.
\end{aligned}$$

Next we need to study the series

$$\sum_{k\geq 1}\frac{(u-1)^{m+1}(1-zt)^{m+1}u^k(1-t)^{k(m+1)}}{\prod_{i=1}^{k+1}\gamma_i}$$

where $m \geq 0$. We can rewrite this series in the form

$$-\frac{(u-1)^{m+1}(1-zt)^{m+1}}{\gamma_1} + \sum_{k\geq 0}\frac{(u-1)^{m+1}(1-zt)^{m+1}u^k(1-t)^{k(m+1)}}{\prod_{i=1}^{k+1}\gamma_i}.$$

We let

$$\psi_{m+1}(u) = \sum_{k\geq 0}\frac{(u-1)^{m+1}(1-zt)^{m+1}u^k(1-t)^{k(m+1)}}{\prod_{i=1}^{k+1}\gamma_i}. \tag{22}$$

We shall show that $\psi_{m+1}(u)$ is in fact a polynomial for all $m \geq 0$. First, we claim that $\psi_{m+1}(u)$ salsifies the following recursion:

$$\psi_{m+1}(u) = \frac{(u-1)^{m+1}(1-zt)^{m+1}}{\gamma_1} + \frac{u\delta_1^m}{\gamma_1}\psi_{m+1}\left(\frac{u}{\delta_1}\right). \tag{23}$$

That is, one can easily iterate (23) to prove by induction that for all $n \geq 1$,

$$\psi_{m+1}(u) = \left(\sum_{s=0}^{2^n-1}\frac{(u-1)^{m+1}(1-zt)^{m+1}u^s(1-t)^{s(m+1)}}{\prod_{i=1}^{s+1}\gamma_i}\right) + \frac{u^{2n}(\delta_{2^n})^m}{\prod_{i=1}^{2^n}\gamma_i}\psi_{m+1}(\frac{u}{\delta_{2^n}}). \tag{24}$$

Hence it follows that if $\psi_{m+1}(u)$ satisfies the recursion (23), then $\psi_{m+1}(u)$ is given by the power series in (22). However, it is routine to check that the polynomial

$$\phi_{m+1}(u) = -\sum_{j=0}^{m}(u-1)^j(1-zt)^j u^{m-j}\prod_{i=j+1}^{m}(1-((1-t)^i) \tag{25}$$

satisfies the recursion that

$$\gamma_1 \phi_{m+1}(u) = (u-1)^{m+1}(1-zt)^{m+1} + u\delta_1^m \phi_{m+1}\left(\frac{u}{\delta_1}\right). \tag{26}$$

Thus we have proved the following lemma.

Lemma 7

$$\begin{aligned}\psi_{m+1}(u) &= \sum_{k\geq 0} \frac{(u-1)^{m+1}(1-zt)^{m+1}u^k(1-t)^{k(m+1)}}{\prod_{i=1}^{k+1}\gamma_i} \\ &= -\sum_{j=0}^{m}(u-1)^j(1-zt)^j u^{m-j} \prod_{i=j+1}^{m}(1-((1-t)^i).\end{aligned} \tag{27}$$

It thus follows that

$$\begin{aligned}H(u,1,z,t) &= \frac{zt(1-u)}{\gamma_1} + \sum_{n\geq 0}\sum_{m=0}^{n}(-1)^{n-m-1}\binom{n}{m}(u-1)^{n-m}\frac{zt}{(1-zt)^{m+1}} \times \\ &\quad -\frac{(u-1)^{m+1}(1-zt)^{m+1}}{\gamma_1} - \sum_{j=0}^{m}(u-1)^j(1-zt)^j u^{m-j} \prod_{i=j+1}^{m}(1-((1-t)^i).\end{aligned}$$

There is no problem in setting $u = 1$ in this expression to obtain that

$$H(1,1,z,t) = \sum_{n\geq 0}\frac{zt}{(1-zt)^{n+1}}\prod_{i=1}^{n}(1-(1-t)^i). \tag{28}$$

Clearly our definitions ensure that $1 + H(1,1,z,t) = P(t,z)$ as defined in the introduction so that we have the following theorem.

Theorem 8

$$P(t,z) = \sum_{n,k\geq 0} p_{n,k}t^n z^k = 1 + \sum_{n\geq 0}\frac{zt}{(1-zt)^{n+1}}\prod_{i=1}^{n}(1-(1-t)^i). \tag{29}$$

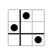
For example, we have used Mathematica to compute the first few terms of $P(t,z)$ as

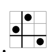
$$\begin{aligned}&P(t,z) = 1 + zt + \left(z+z^2\right)t^2 + \left(2z+2z^2+z^3\right)t^3 + \left(5z+6z^2+3z^3+z^4\right)t^4 \\ &+ \left(15z+21z^2+12z^3+4z^4+z^5\right)t^5 + \left(53z+84z^2+54z^3+20z^4+5z^5+z^6\right)t^6 + O[t]^7.\end{aligned}$$

Next we observe that one can easily derive the ordinary generating function for the number of $(\mathbf{2}+\mathbf{2})$-free posets or, equivalently, for the number of ascent sequences proved by Bousquet-Mélou et al. [1] from Theorem 8. That is, for any sequence of natural numbers $a = a_1 \ldots a_n$, let $a^+ = (a_1+1)\ldots(a_n+1)$ be the result of adding one from each element of the sequence. Moreover, if all the elements of $a = a_1 \ldots a_n$ are positive, then we let $a^- = (a_1-1)\ldots(a_n-1)$ be the result of subtracting one to each element of the sequence. It is easy to see that if $a = a_1 \dots a_n$ is an ascent sequence, then $0a^+$ is also an ascent sequence. Vice versa, if $b = 0a$ is an ascent sequence with only one zero where $a = a_1 \dots a_n$, then a^- is an ascent sequence. It follows that the number of ascent sequences of length n is equal to the number of ascent sequences of length $n+1$ which have only one zero. Hence

$$\begin{aligned} P(t) &= \sum_{n\geq 0} p_n t^n = \frac{1}{t}\frac{\partial P(t,z)}{\partial z}\Big|_{z=0} \\ &= \sum_{n\geq 0}\prod_{i=1}^{n}(1-(1-t)^i). \end{aligned}$$

Results in [1, 2, 3] show that $(\mathbf{2}+\mathbf{2})$-free posets of size n with k minimal elements are in bijection with the following objects. (See [1, 2, 3] for the precise definitions.)

- ascent sequences of length n with k zeros;
- permutations of length n *avoiding*  whose *leftmost-decreasing run* is of size k;
- *regular linearized chord diagrams* on $2n$ points with *initial run of openers* of size k;
- *upper triangular matrices* whose non-negative integer entries sum up to n, each row and column contains a non-zero element, and the sum of entries in the first row is k.

Thus (29) provides generating functions for  -avoiding permutations by the size of the leftmost-decreasing run, for regular linearized chord diagrams by the size of the initial run of openers, and for the upper triangular matrices by the sum of entries in the first row. Moreover, Theorem 4, together with bijections in [1, 2, 3] can be used to enumerate the permutations, diagrams, and matrices subject to 4 statistics.

References

[1] M. Bousquet-Mélou, A. Claesson, M. Dukes, S. Kitaev: Unlabeled (2+2)-free posets, ascent sequences and pattern avoiding permutations. *J. Combin. Theory Ser. A*, to appear.

[2] A. Claesson, M. Dukes, and S. Kitaev, A direct encoding of Stoimenow's matchings as ascent sequences, preprint.

[3] M. Dukes and R. Parviainen, Ascent sequences and upper triangular matrices containing non-negative integers, Elect. J. Combin. **17(1)** (2010), #R53 (16pp).

[4] M. H. El-Zahar, Enumeration of ordered sets, in: I. Rival (Ed.), *Algorithms and Order*, Kluwer Academic Publishers, Dordrecht, 1989, 327–352.

[5] P. C. Fishburn, *Interval Graphs and Interval Orders*, Wiley, New York, 1985.

[6] P. C. Fishburn, Intransitive indifference in preference theory: a survey, *Oper. Res.* **18** (1970) 207–208.

[7] P. C. Fishburn, Intransitive indifference with unequal indifference intervals, *J. Math. Psych.* **7** (1970) 144–149.

[8] P. E. Haxell, J. J. McDonald, and S. K. Thomasson, Counting interval orders, *Order* **4** (1987) 269–272.

[9] S. M. Khamis, Height counting of unlabeled interval and N-free posets, *Discrete Math.* **275** (2004) 165–175.

[10] A. Stoimenow, Enumeration of chord diagrams and an upper bound for Vassiliev invariants, *J. Knot Theory Ramifications* **7** no. 1 (1998) 93–114.

[11] J. Wimp and D. Zeilberger, Resurrecting the asymptotics of linear recurrences, *J. Math. Anal. Appl.*, **111** no. 1 (1985) 162–176.

[12] D. Zagier, Vassiliev invariants and a strange identity related to the Dedeking eta-function, *Topology*, **40** (2001) 945–960.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 702–713

A Closed Character Formula for Symmetric Powers of Irreducible Representations

Stavros Kousidis†

Max Planck Institute for Mathematics, Vivatsgasse 7, 53111 Bonn, Germany

Abstract. We prove a closed character formula for the symmetric powers $S^N V(\lambda)$ of a fixed irreducible representation $V(\lambda)$ of a complex semi-simple Lie algebra $\mathfrak{g}$ by means of partial fraction decomposition. The formula involves rational functions in rank of $\mathfrak{g}$ many variables which are easier to determine than the weight multiplicities of $S^N V(\lambda)$ themselves. We compute those rational functions in some interesting cases. Furthermore, we introduce a residue-type generating function for the weight multiplicities of $S^N V(\lambda)$ and explain the connections between our character formula, vector partition functions and iterated partial fraction decomposition.

Résumé. Nous établissons une formule fermée pour le caractère de la puissance symétrique $S^N V(\lambda)$ d'une représentation irréductible $V(\lambda)$ d'une algèbre de Lie semi-simple complexe $\mathfrak{g}$, en utilisant des décompositions en fractions partielles. Cette formule exprime ce caractère en termes de fractions rationnelles en r variables, où r est le rang de $\mathfrak{g}$. Ces fractions sont plus faciles à déterminer que les multiplicités de la décomposition de $S^N V(\lambda)$ elles-mêmes. Nous calculons ces fonctions rationnelles dans quelques cas intéressants. Nous introduisons par ailleurs une fonction génératrice de type résidu pour les multiplicités de $S^N V(\lambda)$ et relions notre formule aux fonctions de partitions vectorielles et aux décompositions itérées en fractions partielles.

Keywords: character, symmetric power, irreducible representation, generating function, residue, partial fraction decomposition, vector partition function

1 Notation

Let $\mathfrak{g}$ be a complex semi-simple Lie algebra of rank r. Fix a Borel $\mathfrak{b}$ and a Cartan subalgebra $\mathfrak{h}$ in $\mathfrak{g}$ and let $Q = \bigoplus_{i=1}^{r} \mathbb{Z}\alpha_i$ and $X = \bigoplus_{i=1}^{r} \mathbb{Z}\omega_i$ be the corresponding root and weight lattice spanned by the simple roots and fundamental weights respectively. Let $\alpha_1^\vee, \dots, \alpha_r^\vee$ be the simple coroots and W the Weyl group. An irreducible representation of $\mathfrak{g}$ of highest weight $\lambda \in X^+$, where X^+ stands for all dominant weights, is denoted by $V(\lambda)$. Its character will be written as $\operatorname{Char} V(\lambda)$ and it is well-known that it is an element of $\mathbb{Z}[X]$, the integral group ring associated to the weight lattice. Each generator $e^\mu \in \mathbb{Z}[X]$ yields a function on $\mathfrak{h}_\mathbb{R}$, the real span of the simple coroots, by $x \mapsto e^{\langle \mu, x\rangle}$. In this sense we have the associated Fourier series of the character of $V(\lambda)$ as a function of $\mathfrak{h}_\mathbb{R}$, i.e. $\operatorname{Char} V(\lambda)(ix) = \sum_{\mu \in X} m_\mu e^{i\langle \mu, x\rangle}$. To simplify

†Supported by the Max Planck Institute for Mathematics and the Deutsche Forschungsgemeinschaft SFB/TR12 Symmetries and Universality in Mesoscopic Systems.

notation in what follows we define $q = e^{i\langle \cdot ,x\rangle}$, i.e. $q^\mu = e^{i\langle \mu ,x\rangle}$. Then, with respect to the coordinate system $\{\alpha_1^\vee,\dots,\alpha_r^\vee\}$ of $\mathfrak{h}_\mathbb{R}$ we have $q = (q_1,\dots,q_r)$ with $q_i = e^{i\langle \cdot ,x_i\alpha_i^\vee\rangle}$ and for $\mu = c_1\omega_1+\ldots+c_r\omega_r$

$$q^\mu = (q_1,\dots,q_r)^{(c_1,\dots,c_r)} = q_1^{c_1}\cdots q_r^{c_r} \in \mathbb{Z}[q_1^{\pm 1},\dots,q_r^{\pm 1}]. \tag{1}$$

Note, whenever we write $\mathbb{N}$ we mean the non-negative integers $\{0,1,2,\dots\}$.

2 Introduction and the Main Theorem

Let $m_{\lambda,N} : X \to \mathbb{N}$ be the weight multiplicity function for the N-th symmetric power of a fixed irreducible representation $V(\lambda)$ of $\mathfrak{g}$, i.e. $\operatorname{Char} S^N V(\lambda) = \sum_{\nu\in X} m_{\lambda,N}(\nu)e^\nu \in \mathbb{Z}[X]$. Then, we have the combinatorial identity

$$m_{\lambda,N}(\nu) = \sum_{\substack{\{\nu_1,\dots,\nu_N\}\subset X\\ \nu_1+\ldots+\nu_N=\nu}} m_{\lambda,1}(\nu_1)\cdots m_{\lambda,1}(\nu_N). \tag{2}$$

In general it is a non-trivial problem to determine $m_{\lambda,N}$. That is, to establish a formula depending on N that counts the unordered pairs $\{\nu_1,\dots,\nu_N\}$ subject to the restriction $\nu_1+\ldots+\nu_N = \nu$.

We will instead identify the Fourier series associated to the character of $S^N V(\lambda)$ as an element of $\mathbb{C}(q_1,\dots,q_r)[X]$ (see section 1 for the notation). The key point is that this identification involves data (apart from terms in N) which is easier to determine than the function $m_{\lambda,N}$ and depends only on the fixed representation $V(\lambda)$. Starting point will be Molien's formula (see (Procesi, 2007, Chapter 9, §4.3)) which identifies the graded character of the symmetric algebra of $V(\lambda)$ as a product of geometric series. We will state this result here for a quick reference.

Lemma 2.1 (compare (Procesi, 2007, Chapter 9, §4.3))**.**

$$\operatorname{Char} SV(\lambda) = \sum_{N=0}^{\infty} z^N \operatorname{Char} S^N V(\lambda) = \prod_{\nu\in X} \frac{1}{(1-e^\nu z)^{\dim V(\lambda)_\nu}} \tag{3}$$

Our main result will be Theorem 3.4 in Section 3. That is,

Theorem (Character formula)**.** *Let $\mathfrak{g}$ be a semi-simple complex Lie algebra of rank r and $V(\lambda)$ a fixed irreducible representation of $\mathfrak{g}$ with weight space decomposition $V(\lambda) = \bigoplus_{\nu\in X} V(\lambda)_\nu$ and weight multiplicity function $m_\lambda : X \to \mathbb{N}$. Then, with $q = e^{i\langle \cdot ,x\rangle} = (q_1,\dots,q_r)$ as above, we have*

$$\operatorname{Char} S^N V(\lambda)(ix) = \sum_{\nu\in X} q^{N\nu} \sum_{k=1}^{m_\lambda(\nu)} A_{\nu,k}(q)\cdot p_k(N) \in \mathbb{C}(q_1,\dots,q_r)[X] \tag{4}$$

with rational functions $A_{\nu,k}(q) \in \mathbb{C}(q_1,\dots,q_r)$ and polynomials $p_k(N) \in \mathbb{Q}[N]$ of degree $k-1$ given by

$$p_k(N) = \binom{N+k-1}{N}. \tag{5}$$

Furthermore, for a weight $\mu \in X$ and $l = 0,\dots,m_\lambda(\mu)-1$ we have

$$A_{\mu,m_\lambda(\mu)-l}(q) = \frac{(-1)^l}{l!\,q^{l\mu}} \cdot \frac{d^l}{(dz)^l}\left[\prod_{\nu\in X\setminus\mu} \frac{1}{(1-q^\nu z)^{m_\lambda(\nu)}}\right]_{z=q^{-\mu}}. \tag{6}$$

We will apply this theorem to prove character formulas in some interesting cases, involving in particular concrete expressions for the rational functions. To the authors' knowledge there is no formula of such type known so far although the derivation of the Main Theorem is based on simple observations[(i)].

In Section 4, Proposition 4.1 we will prove an integral expression for the generating function associated to the weight multiplicity functions $m_{N,\lambda}$ (evaluated at a fixed weight $\mu \in X$) of the sequence of representations $S^N V(\lambda)$. Based on this identity and our Main Theorem above we will explain the nature of this generating function and in particular why it is of residue-type.

Section 5 will be a short sketch of the connections between the results of section 3, 4 and vector partition functions and iterated partial fraction decomposition (see e.g. Beck (2004) and Bliem (2009)).

Section 6 comments on an important continuation of the present discussion. That is, the character formula established in the Main Theorem can be split into individual parts belonging to the Weyl group orbits of dominant weights. The question is what can be expected from the iterated partial fraction decomposition of those individual terms. We illustrate a possible answer by an example. A detailed treatment will appear in the full version of this extended abstract.

3 A closed character formula for symmetric powers

We will derive a closed character formula for the representation $S^N V(\lambda)$ in terms of a basis of weight vectors of the irreducible representation $V(\lambda)$ with weight multiplicity function m_λ and the parameter N. The term "closed" will be explained in detail in Note 3.5 once we have proven our main result, Theorem 3.4. The method we use is the partial fraction decomposition. That is, consider the identity of Lemma 2.1 for $q = e^{i\langle \cdot, x\rangle}$, $x \in \mathfrak{h}_{\mathbb{R}}$,

$$\sum_{N=0}^{\infty} z^N \operatorname{Char} S^N V(\lambda)(ix) = \prod_{\nu \in X} \frac{1}{(1-q^\nu z)^{m_\lambda(\nu)}}. \tag{7}$$

Partial fraction decomposition with respect to the variable z (abbreviated by PFD_z) of the right-hand side of Equation (7) gives

Proposition 3.1. *Let $\mathfrak{g}$ be a semi-simple complex Lie algebra of rank r and $V(\lambda)$ a fixed irreducible representation of $\mathfrak{g}$ with weight space decomposition $V(\lambda) = \bigoplus_{\nu \in X} V(\lambda)_\nu$ and weight multiplicity function $m_\lambda : X \to \mathbb{N}$. With $q = e^{i\langle \cdot, x\rangle} = (q_1, \dots, q_r)$ as above,*

$$\mathrm{PFD}_z\left(\prod_{\nu \in X} \frac{1}{(1-q^\nu z)^{m_\lambda(\nu)}}\right) = \sum_{\nu \in X} \sum_{k=1}^{m_\lambda(\nu)} A_{\nu,k}(q) \frac{1}{(1-q^\nu z)^k} \tag{8}$$

where for each $\nu \in X$ and $k \in \mathbb{N}$ we have $A_{\nu,k}(q) \in \mathbb{C}(q_1, \dots, q_r)$.

Proof. See e.g. Eustice and Klamkin (1979), Lang (2002), (Bliem, 2009, Lemma 1). □

Note 3.2. The right-hand side of Equation (8) is a finite sum since the second summation gives zero if a weight ν does not appear in $V(\lambda)$, i.e. $m_\lambda(\nu) = 0$.

We aim at a power series expansion of the right-hand side of Equation (8) with respect to the variable z. The following proposition will make life easier.

(i) and on "Mickey Mouse"-analysis as Alan Huckleberry has put it to me in private communication

Proposition 3.3. *For $\nu \in X$ and $q = e^{i\langle \cdot, x\rangle}$ as above, we have*

$$\frac{1}{(1-q^{\nu}z)^k} = \sum_{N=0}^{\infty} z^N q^{N\nu} p_k(N) \tag{9}$$

where $p_k(N)$ is a polynomial in N of degree $k-1$ given by $p_k(N) = \binom{N+k-1}{N}$.

Proof. Write down the Cauchy product of the k-th power of the geometric series $(1-q^{\nu}z)^{-1}$. Then, you see that $p_k(N)$ is given by

$$p_k(N) = \sum_{j_{k-1}=0}^{N} \sum_{j_{k-2}=0}^{j_{k-1}} \cdots \sum_{j_1=0}^{j_2} 1 = \binom{N+k-1}{N}. \tag{10}$$

□

As a direct consequence of Equation (8) and Proposition 3.3 we obtain our main result.

Theorem 3.4 (Character formula)**.** *Let $\mathfrak{g}$ be a semi-simple complex Lie algebra of rank r and $V(\lambda)$ a fixed irreducible representation of $\mathfrak{g}$ with weight space decomposition $V(\lambda) = \bigoplus_{\nu \in X} V(\lambda)_\nu$ and weight multiplicity function $m_\lambda : X \to \mathbb{N}$. Then, with $q = e^{i\langle \cdot, x\rangle} = (q_1, \ldots, q_r)$ as above, we have*

$$\operatorname{Char} S^N V(\lambda)(ix) = \sum_{\nu \in X} q^{N\nu} \sum_{k=1}^{m_\lambda(\nu)} A_{\nu,k}(q) \cdot p_k(N) \in \mathbb{C}(q_1, \ldots, q_r)[X] \tag{11}$$

with rational functions $A_{\nu,k}(q) \in \mathbb{C}(q_1, \ldots, q_r)$ and polynomials $p_k(N) \in \mathbb{Q}[N]$ of degree $k-1$ given by

$$p_k(N) = \binom{N+k-1}{N}. \tag{12}$$

Furthermore, for a weight $\mu \in X$ and $l = 0, \ldots, m_\lambda(\mu) - 1$ we have

$$A_{\mu, m_\lambda(\mu)-l}(q) = \frac{(-1)^l}{l!q^{l\mu}} \cdot \frac{d^l}{(dz)^l} \left[\prod_{\nu \in X \setminus \mu} \frac{1}{(1-q^{\nu}z)^{m_\lambda(\nu)}} \right]_{z=q^{-\mu}}. \tag{13}$$

Proof. From Equation (8) we see that

$$\operatorname{Char} S^N V(\lambda)(ix) = \operatorname{Res}_{z=0} \left[\frac{1}{z^{N+1}} \sum_{\nu \in X} \sum_{k=1}^{m_\lambda(\nu)} A_{\nu,k}(q) \frac{1}{(1-q^{\nu}z)^k} \right]. \tag{14}$$

Then, Proposition 3.3 finishes the first part of the proof. For the second part multiply the right-hand side of Equation (8) by $(1-q^{\mu}z)^{m_\lambda(\mu)}$ which is equivalent to take the product over $X \setminus \mu$ in Equation (13). By the product rule of differentiation we see that all summands except the μ-th one give zero after differentiation and evaluation at $q^{-\mu}$. Therefore, the remaining part is

$$\frac{d^l}{(dz)^l} \left[\sum_{k=1}^{m_\lambda(\mu)} A_{\mu,k}(q)(1-q^{\nu}z)^{m_\lambda(\mu)-k} \right]_{z=q^{-\mu}} \tag{15}$$

$$= \left[\sum_{k=m_\lambda(\mu)-l}^{m_\lambda(\mu)} A_{\mu,k}(q)(-1)^l q^{l\mu} \prod_{i=0}^{l} (m_\lambda(\mu)-k-i)(1-q^\nu z)^{m_\lambda(\mu)-k-l} \right]_{z=q^{-\mu}}$$
$$= A_{\mu,m_\lambda(\mu)-l}(q)(-1)^l q^{l\mu} l!.$$

□

Note 3.5. Now we are able to explain the term "closed" used in the beginning of this section. Namely, the identity stated in Theorem 3.4 shows that all relevant data needed to describe the character of $S^N V(\lambda)$, in particular the rational functions $A_{\nu,k}(q)$, depends on the weight space decomposition and weight multiplicity function m_λ of the fixed representation $V(\lambda)$.

Note 3.6. Equation (13) might be a simple observation but it is a very effective method to compute the rational functions associated to weights of multiplicity 1. Then, we have no differentiation but just simple evaluation. In particular, one can immediately compute the character of the symmetric powers of a multiplicity free irreducible representation V. Note that in this case one could also obtain the character of $S^N V$ by plugging the k-many weights of the representation V into the complete homogeneous symmetric polynomial identity

$$h_N(x_1,\ldots,x_k) = \sum_{i=1}^{k} \frac{x_i^N}{\prod_{j\neq i}(1-x_j x_i^{-1})}. \tag{16}$$

As a consequence of Note 3.6 we can prove concrete character formulas for the symmetric powers of the irreducible representations $V(m)$ of $\mathfrak{g}$ being of type A_1 and furthermore for the symmetric powers of the fundamental representation $V(\omega_1)$ of $\mathfrak{g}$ of type A_r.

Corollary 3.7. *For $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$ and its irreducible representation $V(m)$, $m \in \mathbb{N}$, the Fourier series of the character of $S^N V(m)$ is given by*

$$\operatorname{Char} S^N V(m)(ix) = \sum_{i=0}^{m} q^{N(m-2i)} A_{m-2i,1}(q) \in \mathbb{C}(q)[X] \tag{17}$$

where $q = e^{ix}$ as above and with rational functions

$$A_{m-2i,1}(q) = (-1)^i q^{(m-i)(m-i+1)} \prod_{\substack{j=0\\ j\neq i}}^{m} \frac{1}{q^{2|i-j|}-1}. \tag{18}$$

Proof. The weights of $V(m)$ are given by $(m-2i)\omega_1$ where $i = 0,\ldots,m$. By Theorem 3.4 we immediately obtain the claimed character formula and

$$A_{m-2i,1}(q) = \frac{(-1)^0}{0!q^{0(m-2i)\omega_1}} \frac{d^0}{(dz)^0} \left[\prod_{\substack{j=0\\ j\neq i}}^{m} \frac{1}{1-q^{(m-2j)}z} \right]_{z=q^{-(m-2i)}} = \prod_{\substack{j=0\\ j\neq i}}^{m} \frac{1}{1-q^{(m-2j)}q^{-(m-2i)}} \tag{19}$$
$$= \prod_{0\leq j<i} \frac{1}{1-q^{2(i-j)}} \prod_{i<j\leq m} \frac{q^{2(j-i)}}{q^{2(j-i)}-1} = (-1)^i q^{(m-i)(m-i+1)} \prod_{\substack{j=0\\ j\neq i}}^{m} \frac{1}{q^{2|i-j|}-1}.$$

□

Example 3.8. Let $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$. Since $S^N V(0) = V(0)$ and $S^N V(1) = V(N)$, the first non-trivial example is given by the adjoint representation $V(2)$ and its symmetric powers $S^N V(2)$. By Corollary 3.7 we have

$$\operatorname{Char} S^N V(2)(ix) = \frac{q^6}{(q^4-1)(q^2-1)} \cdot q^{2N} + \frac{-q^2}{(q^2-1)^2} \cdot q^0 + \frac{1}{(q^4-1)(q^2-1)} \cdot q^{-2N}. \tag{20}$$

Corollary 3.9. *Let $\mathfrak{g} = \mathfrak{sl}(r+1,\mathbb{C})$ and consider its fundamental representation $V(\omega_1)$. Set $\omega_0 = \omega_{r+1} = 0$, i.e. extend $q = e^{i\langle \cdot, x\rangle} = (q_1,\dots,q_r)$ by $q_0 = q_{r+1} = 1$. Then,*

$$\operatorname{Char} S^N V(\omega_1)(ix) = \sum_{i=0}^{r} q_i^{-N} q_{i+1}^{N} A_{-\omega_i+\omega_{i+1},1}(q) \in \mathbb{C}(q_1,\dots,q_r)[X] \tag{21}$$

with rational functions

$$A_{-\omega_i+\omega_{i+1},1}(q) = q_{i+1}^r \prod_{\substack{j=0\\ j\neq i}}^{r} \frac{q_j}{q_j q_{i+1} - q_{j+1} q_i}. \tag{22}$$

Proof. The weights of the fundamental representation $V(\omega_1)$ are $\omega_1, -\omega_1+\omega_2, \dots, -\omega_{n-1}+\omega_n, -\omega_n$ all of multiplicity 1. Again, by Theorem 3.4 the claimed character formula follows and

$$\begin{aligned} A_{-\omega_i+\omega_{i+1},1}(q) &= \frac{(-1)^0}{0! q^{0(-\omega_i+\omega_{i+1})}} \frac{d^0}{(dz)^0} \left[\prod_{\substack{j=0\\ j\neq i}}^{r} \frac{1}{1 - q_j^{-1} q_{j+1} z} \right]_{z = q_i q_{i+1}^{-1}} = \prod_{\substack{j=0\\ j\neq i}}^{r} \frac{1}{1 - q_j^{-1} q_{j+1} q_i q_{i+1}^{-1}} \\ &= \prod_{\substack{j=0\\ j\neq i}}^{r} \frac{q_j q_{i+1}}{q_j q_{i+1} - q_{j+1} q_i} = q_{i+1}^r \prod_{\substack{j=0\\ j\neq i}}^{r} \frac{q_j}{q_j q_{i+1} - q_{j+1} q_i}. \end{aligned} \tag{23}$$

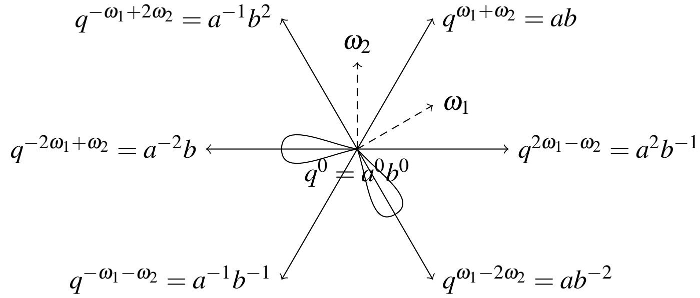
□

Remark 3.10. For $\mathfrak{g} = \mathfrak{sl}(r+1,\mathbb{C})$ we have an interesting aspect coming up. Since $S^N V(\omega_1) = V(N\omega_1)$, it is interesting to ask how the formulas obtained in Corollary 3.9 compare to the asymptotic theory of the Duistermaat-Heckman measure with respect to the sequence of representations $V(N\omega_1)$.

Note 3.11. Similarly to Corollary 3.9 one can compute the characters of the symmetric powers of the representations $V(\omega_i)$ for $i = 2,\dots,r$. Note that although the number of weights contributing to $S^N V(\omega_1) = V(N\omega_1)$ grows with N, their multiplicities always remain equal to 1. Nevertheless, the rational functions associated to $V(\omega_1)$ do not carry only trivial information, the number 1, but also encode which weights appear in $S^N V(\omega_1)$. In contrast, the weight multiplicities in $S^N V(\omega_i)$ for $i = 2,\dots,r$ are non-trivial and consequently their associated rational functions encode much more information. It is part of the full version of this extended abstract to compute the characters of the $S^N V(\omega_i)$ and compare those formulas.

For representations with higher dimensional weight spaces (dim ≥ 2) the computations become more difficult. We will demonstrate this by an example.

Example 3.12. Let $\mathfrak{g} = \mathfrak{sl}(3,\mathbb{C})$ and $V(\omega_1+\omega_2)$ be its adjoint representation which decomposes as shown in the following picture, where $q = e^{i\langle \cdot, x\rangle} = (q_1,q_2) = (a,b)$ with respect to the fundamental weights ω_1,ω_2 and the simple coroots $\alpha_1^\vee, \alpha_2^\vee$.



The picture shows the Littelmann paths $\mathscr{P}_{\omega_1+\omega_2}$ of shape $\omega_1+\omega_2$ (see e.g. Littelmann (1994)) and the elements of $\mathbb{Z}[a^{\pm 1}, b^{\pm 1}]$ corresponding to the weights of $V(\omega_1+\omega_2)$. Here the difficulty lies in computing the rational function associated to the zero weight which has multiplicity 2. This is a first example of a non-trivial polynomial $p_k(N)$ coming up, namely $p_2(N) = N+1$. We have

$$\begin{aligned}\operatorname{Char} V(N\omega_1)(ix) =&(A_{0,1}(q)+A_{0,2}(q)p_2(N))\cdot q^{N0} \\ &+A_{2\omega_1-\omega_2,1}(q)\cdot q^{N(2\omega_1-\omega_2)}+A_{-2\omega_1+\omega_2,1}(q)\cdot q^{N(-2\omega_1+\omega_2)} \\ &+A_{\omega_1-2\omega_2,1}(q)\cdot q^{N(\omega_1-2\omega_2)}+A_{-\omega_1+2\omega_2,1}(q)\cdot q^{N(-\omega_1+2\omega_2)} \\ &+A_{-\omega_1-\omega_2,1}(q)\cdot q^{N(-\omega_1-\omega_2)}+A_{\omega_1+\omega_2,1}(q)\cdot q^{N(\omega_1+\omega_2)}.\end{aligned} \tag{24}$$

The difficult part is

$$A_{0,1}(q) = \frac{d}{dz}\left[\prod_{\nu\in X\setminus 0}\frac{1}{(1-q^{\nu}z)^{m_{\omega_1+\omega_2}(\nu)}}\right]_{z=q^0=a^0b^0=1} = \frac{-3a^4b^4}{(ab-1)^2(a-b^2)^2(a^2-b)^2} \tag{25}$$

and we obtain

$$\begin{aligned}\operatorname{Char} V(N\omega_1)(ix) =&\frac{-(3a^4b^4+a^4b^4p_2(N))}{(ab-1)^2(a-b^2)^2(a^2-b)^2}\cdot a^0b^0 \\ &+\frac{a^{16}b}{(ab-1)(a-b^2)(a^2-b)^2(a^3-1)(a^3-b^3)(a^4-b^2)}\cdot a^{2N}b^{-N} \\ &+\frac{-b^9}{(ab-1)(a-b^2)(a^2-b)^2(a^3-1)(a^3-b^3)(a^4-b^2)}\cdot a^{-2N}b^{N} \\ &+\frac{a^9}{(ab-1)(a-b^2)^2(a^2-b)(b^3-1)(a^3-b^3)(a^2-b^4)}\cdot a^{N}b^{-2N} \\ &+\frac{-ab^{16}}{(ab-1)(a-b^2)^2(a^2-b)(b^3-1)(a^3-b^3)(a^2-b^4)}\cdot a^{-N}b^{2N} \\ &+\frac{ab}{(ab-1)^2(a-b^2)(a^2-b)(b^3-1)(a^3-1)(a^2b^2-1)}\cdot a^{-N}b^{-N} \\ &+\frac{-a^9b^9}{(ab-1)^2(a-b^2)(a^2-b)(b^3-1)(a^3-1)(a^2b^2-1)}\cdot a^{N}b^{N}.\end{aligned} \tag{26}$$

Let us end this section with an important note.

Note 3.13. In the notation of Theorem 3.4 note that iterated partial fraction decomposition with respect to the variables $q_1,\dots,q_r$ gives the Fourier series associated to the character of $S^N V(\lambda)$. Thus, decomposing the character formula in Theorem 3.4 further with respect to $q_1,\dots,q_r$ yields the weight multiplicity functions $m_{\lambda,N}$. We illustrate this by elaborating on Example 3.8 where $r=1$. That is, let us decompose the character

$$\operatorname{Char} S^N V(2)(ix) = \frac{q^6}{(q^4-1)(q^2-1)}\cdot q^{2N} + \frac{-q^2}{(q^2-1)^2}\cdot q^0 + \frac{1}{(q^4-1)(q^2-1)}\cdot q^{-2N} \tag{27}$$

further with respect to q. For e.g. $N=0,\dots,5$ this gives

N	$\operatorname{Char} S^N V(2)(ix)$
1	q^2+1+q^{-2}
2	$q^4+q^2+2+q^{-2}+q^{-4}$
3	$q^6+q^4+2q^2+2+2q^{-2}+q^{-4}+q^{-6}$
4	$q^8+q^6+2q^4+2q^2+3+2q^{-2}+2q^{-4}+q^{-6}+q^{-8}$
5	$q^{10}+q^8+2q^6+2q^4+3q^2+3+3q^{-2}+2q^{-4}+2q^{-6}+q^{-8}+q^{-10}.$

4 A residue-type generating function for the weight multiplicities

Consider the Fourier series associated to the character of the representation $S^N V(\lambda)$ of our Lie algebra $\mathfrak{g}$, i.e. $\operatorname{Char} S^N V(\lambda)(ix) = \sum_{\nu\in X} m_{\lambda,N}(\nu) e^{i\langle\nu,x\rangle}$. Here $m_{\lambda,N}$ denotes the weight multiplicity function of $S^N V(\lambda)$. Then, by inverse Fourier transform we can recover the Fourier coefficients $m_{\lambda,N}(\nu)$ as

$$m_{\lambda,N}(\nu) = \frac{1}{(2\pi)^r}\int\limits_{\mathfrak{h}_{\mathbb{R}}/2\pi X^*} e^{-i\langle\nu,x\rangle}\operatorname{Char} S^N V(\lambda)(ix)dx. \tag{28}$$

Here dx is Lebesgue measure on $\mathfrak{h}_{\mathbb{R}}$ normalized such that the volume of the torus $T^r = \mathfrak{h}_{\mathbb{R}}/2\pi X^*$ is $(2\pi)^r$. Note that r is the rank of $\mathfrak{g}$. This yields the generating function for the weight multiplicity functions $m_{\lambda,N}$ evaluated at a specific weight. That is,

Proposition 4.1. *Let $\mathfrak{g}$ be a semi-simple complex Lie algebra of rank r and $V(\lambda)$ a fixed irreducible representation of $\mathfrak{g}$. Let $m_{\lambda,N}$ be the weight multiplicity function of the N-th symmetric power $S^N V(\lambda)$. Let $\mu\in X$ be a fixed weight. Then, the formal power series $\sum_{N=0}^\infty z^N m_{\lambda,N}(\mu)$ is a holomorphic function in the variable z on $|z|\le R<1$. Moreover, we have the identity*

$$\sum_{N=0}^{\infty} z^N m_{\lambda,N}(\mu) = \frac{1}{(2\pi)^r}\int\limits_{T^r} e^{-i\langle\mu,x\rangle}\prod_{\nu\in X}\frac{1}{(1-e^{i\langle\nu,x\rangle}z)^{m_{\lambda,1}(\nu)}}dx. \tag{29}$$

Proof. The assertion follows from the fact that the dimension of the symmetric power of a representation grows sub-exponentially in N as

$$\dim S^N V(\lambda) = \binom{\dim V(\lambda)-1+N}{N}. \tag{30}$$

This amounts to say that for the fixed weight $\mu \in X$ the power series $\sum_{N=0}^{\infty} z^N e^{-i\langle \mu,x\rangle} \operatorname{Char} S^N V(\lambda)(ix)$ converges absolutely on $|z| \leq R < 1$ uniformely in $x \in T^r$. Namely, for arbitrary such x we have

$$\begin{aligned} |e^{-i\langle \mu,x\rangle} \operatorname{Char} S^N V(\lambda)(ix)| &= |\sum_{\nu \in X} m_{\lambda,N}(\nu) e^{i\langle \nu,x\rangle}| \\ {\scriptstyle\text{(triangle inequality)}} &\leq \sum_{\nu\in X} m_{\lambda,N}(\nu) \\ &= \binom{\dim V(\lambda) - 1 + N}{N} \\ {\scriptstyle (C\text{ some constant})} &= CN^{\dim V(\lambda)-1} + \text{lower terms}. \end{aligned} \tag{31}$$

Therefore the radius of convergence is given by

$$r = \frac{1}{\limsup\limits_{N\to\infty} \sqrt[N]{|e^{-i\langle \mu,x\rangle} \operatorname{Char} S^N V(\lambda)(ix)|}} = \frac{1}{\limsup\limits_{N\to\infty} \sqrt[N]{|CN^{\dim V(\lambda)-1} + {\scriptstyle\text{lower terms}}|}} = 1. \tag{32}$$

By Lemma 2.1 the right-hand side of Equation (29) equals

$$\frac{1}{(2\pi)^r} \int_{T^r} e^{-i\langle \mu,x\rangle} \sum_{N=0}^{\infty} z^N \operatorname{Char} S^N V(\lambda)(ix) dx \tag{33}$$

and since the previous convergence arguments allow us to integrate term by term, this finishes the proof. □

Now we are able to explain why the generating function in Equation (29) is of residue-type.

Corollary 4.2 (Residue-type)**.** *Let $\mathfrak{g}$ be a semi-simple complex Lie algebra of rank r and $V(\lambda)$ a fixed irreducible representation of $\mathfrak{g}$. Let $m_{\lambda,N}$ be the weight multiplicity function of the N-th symmetric power $S^N V(\lambda)$. Let $\mu \in X$ be a fixed weight and denote $q^\mu = e^{i\langle \mu,x\rangle}$ as above. Then,*

$$m_{\lambda,N}(\mu) = \frac{1}{(2\pi)^r} \int_{T^r} q^{-\mu} \sum_{\nu\in X} q^{N\nu} \sum_{k=1}^{m_\lambda(\nu)} A_{\nu,k}(q) \cdot p_k(N) dx. \tag{34}$$

In particular, the multiplicity $m_{\lambda,N}(\mu)$ equals the constant term of the function $\operatorname{Char} S^N V(\lambda)(ix)$ *shifted by $q^{-\mu}$.*

Proof. This is a direct consequence of Proposition 4.1 and Theorem 3.4. □

Example 4.3. In the case $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$ and the symmetric powers $S^N V(2)$ of the adjoint representation we have described in Note 3.13 that, e.g. for $N = 4$,

$$\operatorname{Char} S^4 V(2)(ix) = q^8 + q^6 + 2q^4 + 2q^2 + 3 + 2q^{-2} + 2q^{-4} + q^{-6} + q^{-8}. \tag{35}$$

Note that $q = e^{ix}$. Now, in view of Corollary 4.2, the multiplicity of the weight $\mu = 2\omega_1$ in $S^4 V(2)$ is given by

$$m_{\lambda,N}(\mu) = m_{2,4}(2) = \frac{1}{2\pi} \int_{S^1} q^{-2}(q^8 + q^6 + 2q^4 + 2q^2 + 3 + 2q^{-2} + 2q^{-4} + q^{-6} + q^{-8}) dx \tag{36}$$

$$= \frac{1}{2\pi}\int\limits_{S^1} q^6+q^4+2q^2+2q^0+3q^{-2}+2q^{-4}+2q^{-6}+q^{-8}+q^{-10}dx = \frac{1}{2\pi}\int\limits_{S^1} 2q^0 dx = 2$$

Remark 4.4. Similar to Proposition 4.1 we have a generating function for the weight multiplicity functions $m^{\Lambda}_{\lambda,N}$ of the exterior powers $\Lambda^N V(\lambda)$ of an irreducible representation $V(\lambda)$. First, realize (see (Procesi, 2007, Chapter 9, §4.3)) that the graded character of the exterior algebra of $V(\lambda)$ is given by

$$\operatorname{Char}\Lambda V(\lambda) = \sum_{N=0}^{\infty} z^N \operatorname{Char}\Lambda^N V(\lambda) = \prod_{\nu\in X}(1+e^{\nu}z)^{m^{\Lambda}_{\lambda,1}(\nu)}. \tag{37}$$

Then, again by inverse Fourier transform and the same convergence arguments we obtain a generating function with radius of convergence equal to 1, satisfying the identity

$$\sum_{N=0}^{\infty} z^N m^{\Lambda}_{\lambda,N}(\mu) = \frac{1}{(2\pi)^r}\int\limits_{T^r} e^{-i\langle\mu,x\rangle}\prod_{\nu\in X}(1+e^{i\langle\nu,x\rangle}z)^{m^{\Lambda}_{\lambda,1}(\nu)}dx. \tag{38}$$

Remark 4.5. For the tensor powers $T^N V(\lambda)$ of a fixed irreducible representation $V(\lambda)$ with weight multiplicity functions $m^T_{\lambda,N}$ we have the identity $\operatorname{Char}T^N V(\lambda) = (\operatorname{Char}V(\lambda))^N$ and consequently

$$(2\pi)^r \sum_{N=0}^{\infty} z^N m^T_N(\mu) = \int\limits_{T^r} e^{-i\langle\mu,x\rangle}\sum_{N=0}^{\infty} z^N \operatorname{Char}T^N V(\lambda)(ix)dx = \int\limits_{T^r} e^{-i\langle\mu,x\rangle}\frac{1}{1-\operatorname{Char}V(\lambda)(ix)z}dx. \tag{39}$$

This constitutes a holomorphic function with radius of convergence equal to $\frac{1}{\dim V(\lambda)}$.

5 Connection to vector partition functions

For an integral matrix $A\in\mathbb{Z}^{(m,d)}$ with $\ker(A)\cap\mathbb{R}^d_+ = \{0\}$ we define the vector partition function $\phi_A : \mathbb{Z}^m \to \mathbb{N}$ by

$$\phi_A(b) = \#\{x\in\mathbb{N}^d : Ax = b\}. \tag{40}$$

Let $c_1,\dots,c_d$ denote the columns of A and use multiexponent notation $z^b = z_1^{b_1}\cdots z_m^{b_m}$, $b\in\mathbb{Z}^m$. Then, as stated in (Bliem, 2009, Equation (1)), on $\{z\in\mathbb{C}^m : |z^{c_k}|<1 \text{ for } k=1,\dots,d\}$ we have the identity

$$f_A(z) := \sum_{b\in\mathbb{Z}^m}\phi_A(b)z^b = \prod_{k=1}^{d}\frac{1}{1-z^{c_k}} \tag{41}$$

and

$$\phi_A(b) = \operatorname{const}\left[f_A(z)\cdot z^{-b}\right]. \tag{42}$$

Now, there is an obvious connection between the graded character of the symmetric algebra $SV(\lambda)$ of an irreducible representation $V(\lambda)$ of a complex semi-simple Lie algebra $\mathfrak{g}$ and the theory of vector partition functions, which is given by Lemma 2.1. Namely, if $\mathfrak{g}$ is of rank r, then one has a matrix $A\in\mathbb{Z}^{(r+1,\dim V(\lambda))}$ encoding the weights of $V(\lambda)$ in terms of the coordinate system given by the fundamental weights $\omega_1,\dots,\omega_r$. This information corresponds to the first r rows of each column of A. In addition to that, we have the $(r+1)$-th row which associates to the grading given by z in Lemma 2.1. That is, our particular matrix A has the following properties

1. The last row of A equals $(1,\dots,1)$. (grading)
2. The columns of A reflect the Weyl group action. (symmetry)
3. The columns of A appear with multiplicities. (multiplicity)

In contrast to the computational and algorithmic aspects of iterated partial fraction decomposition as proposed in Beck (2004) and continued e.g. in Bliem (2009) for "arbitrary" matrices A, our interests are different. They lie in investigating further the closed character formulas for the symmetric powers and the impact of the grading, symmetry and multiplicity properties of our matrix A on the iterated partial fraction decomposition. One aspect is described in detail in Section 6.

6 Weyl group orbits and the Main Theorem

In the notation of Theorem 3.4 write the character of $S^N V(\lambda)$ as the sum over the dominant weights and their Weyl group orbits, i.e.

$$\operatorname{Char} S^N V(\lambda)(ix) = \sum_{\nu \in X^+} \sum_{w \in W/W_\nu} q^{Nw.\nu} \sum_{k=1}^{m_\lambda(\nu)} A_{w.\nu,k}(q) \cdot p_k(N). \tag{43}$$

Here W_ν denotes the stabilizer of the weight ν. Note that the multiplicity of a weight is invariant under the operation of the Weyl group (see e.g. (Carter, 2005, Proposition 10.22)). Now, for a fixed dominant weight $\nu \in X^+$ let

$$f_{\nu,N}(q) = \sum_{w \in W/W_\nu} q^{Nw.\nu} \sum_{k=1}^{m_\lambda(\nu)} A_{w.\nu,k}(q) \cdot p_k(N) \in \mathbb{C}(q_1,\dots,q_r)[X] \tag{44}$$

so that $\operatorname{Char} S^N V(\lambda)(ix) = \sum_{\nu \in X^+} f_{\nu,N}(q)$. It is interesting to ask how the iterated partial fraction decomposition with respect to the variables $q_1,\dots,q_r$ of a single summand $f_{\nu,N}(q)$ looks like. Examples indicate that this decomposition of $f_{\nu,N}(q)$ does not yield information about the weights outside the convex hull of the Weyl group orbit $W.(N\nu)$. Furthermore, some additional terms appear which sum up to zero when taken over all dominant weights X^+. We will illustrate this by an example in the case of $\mathfrak{g}$ being of rank 1 to avoid confusing computations.

Example 6.1. Consider the sequence of representations $S^N V(3)$ of $\mathfrak{g} = \mathfrak{sl}(2,\mathbb{C})$. Then, by Corollary 3.7 we have

$$\begin{aligned}\operatorname{Char} S^N V(3)(ix) =& \frac{q^{12}}{(q^6-1)(q^4-1)(q^2-1)} \cdot q^{3N} + \frac{-q^6}{(q^4-1)(q^2-1)^2} \cdot q^N \\ &+ \frac{q^2}{(q^4-1)(q^2-1)^2} \cdot q^{-N} + \frac{-1}{(q^6-1)(q^4-1)(q^2-1)} \cdot q^{-3N}\end{aligned} \tag{45}$$

where $q = e^{ix}$. Hence, following the notation introduced in Equation (44) we set

$$f_{1,N}(q) = \frac{-q^6}{(q^4-1)(q^2-1)^2} \cdot q^N + \frac{q^2}{(q^4-1)(q^2-1)^2} \cdot q^{-N} \tag{46}$$

$$f_{3,N}(q) = \frac{q^{12}}{(q^6-1)(q^4-1)(q^2-1)} \cdot q^{3N} + \frac{-1}{(q^6-1)(q^4-1)(q^2-1)} \cdot q^{-3N}. \tag{47}$$

Now, e.g. for $N = 4$, we obtain

$$\mathrm{PFD}_q(f_{1,4}(q)) = -q^2 - 2 - q^{-2} - \frac{3}{4(q-1)^2} + \frac{3}{4(q+1)} - \frac{3}{4(q-1)} - \frac{3}{4(q+1)^2} \tag{48}$$

and

$$\begin{aligned}\mathrm{PFD}_q(f_{3,4}(q)) =& q^{12} + q^{10} + 2q^8 + 3q^6 + 4q^4 + 5q^2 + 7 + 5q^{-2} + 4q^{-4} + 3^{-6} + 2q^{-8} \\ & + q^{-10} + q^{-12} + \frac{3}{4(q-1)^2} - \frac{3}{4(q+1)} + \frac{3}{4(q-1)} + \frac{3}{4(q+1)^2},\end{aligned} \tag{49}$$

where in each individual decomposition the last four summands are the additional terms which sum up to zero. Unfortunately this example indicates that we cannot expect a positive formula for the weight multiplicities of the symmetric powers.

7 Acknowledgements

The author would like to thank the referees for their careful reviews and comments, and for pointing me to the identity in Note 3.6.

References

M. Beck. The partial-fractions method for counting solutions to integral linear systems. *Discrete Comput. Geom.*, 32(4):437–446, 2004. ISSN 0179-5376. doi: 10.1007/s00454-004-1131-5. URL `http://dx.doi.org/10.1007/s00454-004-1131-5`.

T. Bliem. Towards computing vector partition functions by iterated partial fraction decomposition. *Preprint*, arXiv:0912.1131v1, 2009.

R. W. Carter. *Lie algebras of finite and affine type*, volume 96 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2005. ISBN 978-0-521-85138-1; 0-521-85138-6.

D. Eustice and M. S. Klamkin. On the coefficients of a partial fraction decomposition. *Amer. Math. Monthly*, 86(6):478–480, 1979. ISSN 0002-9890. doi: 10.2307/2320421. URL `http://dx.doi.org/10.2307/2320421`.

S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002. ISBN 0-387-95385-X.

P. Littelmann. A Littlewood-Richardson rule for symmetrizable Kac-Moody algebras. *Invent. Math.*, 116 (1-3):329–346, 1994. ISSN 0020-9910.

C. Procesi. *Lie groups*. Universitext. Springer, New York, 2007. ISBN 978-0-387-26040-2; 0-387-26040-4. An approach through invariants and representations.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 714–725

On extensions of the Newton-Raphson iterative scheme to arbitrary orders

Gilbert Labelle[†]

LaCIM et Département de mathématiques, Université du Québec à Montréal (UQAM) case postale 8888, succursale Centre-Ville, Montréal (Québec) Canada H3C 3P8

Abstract. The classical quadratically convergent Newton-Raphson iterative scheme for successive approximations of a root of an equation $f(t) = 0$ has been extended in various ways by different authors, going from cubical convergence to convergence of arbitrary orders. We introduce two such extensions, using appropriate differential operators as well as combinatorial arguments. We conclude with some applications including special series expansions for functions of the root and enumeration of classes of tree-like structures according to their number of leaves.

Résumé. Le schéma itératif classique à convergence quadratique de Newton-Raphson pour engendrer des approximations successives d'une racine d'une équation $f(t) = 0$ a été étendu de plusieurs façons par divers auteurs, allant de la convergence cubique à des convergences d'ordres arbitraires. Nous introduisons deux telles extensions en utilisant des opérateurs différentiels appropriés ainsi que des arguments combinatoires. Nous terminons avec quelques applications incluant des développements en séries exprimant des fonctions de la racine et l'énumération de classes de structures arborescentes selon leur nombre de feuilles.

Keywords: Newton-Raphson iteration, order of convergence, combinatorial species, tree-like structures

1 Introduction

Let $(t_n)_{n\geq 0}$ be a sequence of real numbers converging to a. The convergence is said to be *of order* p if

$$t_{n+1} - a = O\left((t_n - a)^p\right), \text{ as } n \to \infty. \tag{1}$$

This means that the convergence is very rapid, when $p \geq 2$, since the number of correct decimal digits in the approximation of a is essentially multiplied by p at each step. Now, let $U \subseteq \mathbb{R}$ be an open set and $f : U \to \mathbb{R}$ be twice differentiable. If the equation $f(t) = 0$ has a simple root $a \in U$, then the classical Newton-Raphson iterative scheme[(i)],

$$t_{n+1} = \mathcal{N}(t_n), \quad n = 0, 1, 2, ..., \quad \text{with} \quad \mathcal{N}(t) = t - \frac{f(t)}{f'(t)}, \tag{2}$$

[†]With the support of NSERC (Canada)

[(i)] Also called the tangent method

produces a quadratically convergent $(p = 2)$ sequence of approximations $t_n \to a$, as $n \to \infty$, whenever the first approximation, t_0, is sufficiently near to a. For suitably regular functions, cubical convergence $(p = 3)$, can be achieved using Householder's method (see Householder (1970)),

$$\mathcal{N}(t) = t - \frac{f(t)}{f'(t)}\left(1 + \frac{f(t)\,f''(t)}{2f'(t)^2}\right), \tag{3}$$

or the method of the astronomer Halley (1656-1743),

$$\mathcal{N}(t) = t - \frac{2f(t)\,f'(t)}{2f'(t)^2 - f(t)\,f''(t)}. \tag{4}$$

More generally, to achieve convergence of order $k+1$, $k \geq 3$, one can use the general Householder's method (see Householder (1970)),

$$\mathcal{N}(t) = t + k\frac{(1/f)^{(k-1)}(t)}{(1/f)^{(k)}(t)}. \tag{5}$$

Another way to achieve arbitrary order convergence is to make use of the method of indeterminate coefficients together with a Taylor expansion around the root (see Sebah and Gourdon (2001)). In Section 2, we use the inverse function theorem and differential operators to replace (5) by two explicit finite sums which also provide arbitrary order convergence. Section 3 is devoted to a combinatorial approach to these finite sums. We conclude, in Section 4, with some applications including special series expansions for functions of the root and enumeration of classes of tree-like structures according to their number of leaves.

2 Differential operators and higher order convergence

Assume that f is of class C^{k+1} around the simple root a. Then, by the inverse function theorem, there exists an open interval V containing the root a which is mapped bijectively by f onto an open interval W containing 0. Moreover, the inverse function $f^{<-1>} : W \to V$ is also of class C^{k+1}. Using these facts, we can express the root a in the following way,

$$a = f^{<-1>}(0) = f^{<-1>}(f(t) - f(t)) = f^{<-1>}(f(t) + u)|_{u:=-f(t)}, \tag{6}$$

whenever t is sufficiently near of a. Now, fix such a t and consider the function

$$\Phi_t(u) = f^{<-1>}(f(t) + u). \tag{7}$$

Then, by Taylor's expansion with remainder, we have, for every small value of u,

$$\Phi_t(u) = \sum_{\nu=0}^{k} \Phi_t^{(\nu)}(0)\frac{u^k}{k!} + \Phi_t^{(k+1)}(\theta u)\frac{u^{k+1}}{(k+1)!}, \quad 0 \leq \theta \leq 1. \tag{8}$$

The Taylor's coefficients can be computed very easily in terms of f as follows :

Lemma 2.1 *Let $D = d/dt$, then for $\nu \leq k+1$,*

$$\Phi_t^{(\nu)}(0) = \left(\frac{1}{f'(t)}D\right)^{\nu} t. \tag{9}$$

Proof: Of course, $\Phi_t^{(\nu)}(0) = (f^{<-1>})^{(\nu)}(f(t))$. Applying the operator D on both sides, we get by the chain-rule,

$$D\Phi_t^{(\nu)}(0) = D(f^{<-1>})^{(\nu)}(f(t)) = (f^{<-1>})^{(\nu+1)}(f(t))f'(t) = f'(t)\Phi_t^{(\nu+1)}(0). \tag{10}$$

Hence, $\Phi_t^{(\nu+1)}(0) = \left(\frac{1}{f'(t)}D\right)\Phi_t^{(\nu)}(0)$ and we conclude using the fact that, for $\nu = 0$, $\Phi_t(0) = (f^{<-1>})(f(t)) = t$. □

Since $a = \Phi_t(u)|_{u=-f(t)}$ and $f(t)^{k+1} \sim f'(a)^{k+1} \cdot (t-a)^{k+1}$, as $t \to a$, we finally obtain,

Proposition 2.2 *Let f be of class C^{k+1} around the simple root a and let*

$$\mathcal{N}(t) = \sum_{\nu=0}^{k} (-1)^{\nu} \frac{f(t)^{\nu}}{\nu!} \left(\frac{1}{f'(t)}D\right)^{\nu} t. \tag{11}$$

Then $\mathcal{N}(t) - a = O((t-a)^{k+1})$ and, for every t_0 sufficiently near to a, the sequence $(t_n)_{n\geq 0}$, defined by $t_{n+1} = \mathcal{N}(t_n)$, converges to a to the order $k+1$. More precisely,

$$t_{n+1} - a \sim C \cdot (t_n - a)^{k+1}, \quad n \to \infty, \tag{12}$$

where

$$C = (-1)^{k+1} \left[\frac{f'(t)^{k+1}}{(k+1)!} \left(\frac{1}{f'(t)}D\right)^{k+1} t\right]_{t:=a}. \tag{13}$$

□

Note that for $k = 1$ (resp. $k = 2$), (11) corresponds to the Newton-Raphson (resp. Householder) iterative step. However, for $k \geq 3$, (11) and the general Houseolder's iterative step (5) are completely different. For example, for $k = 3$, (5) is given by

$$\mathcal{N}(t) = t - f(t)\left(\frac{f'(t)^2 - f(t)f''(t)/2}{f'(t)^3 - f(t)f'(t)f''(t) + f'''(t)f(t)^2/6}\right), \tag{14}$$

while (11) has the form

$$\mathcal{N}(t) = t - \frac{f(t)}{f'(t)}\left(1 + \frac{f(t)f''(t)}{2!f'(t)^2} + \frac{f(t)^2\left(3f''(t)^2 - f'(t)f'''(t)\right)}{3!f'(t)^4}\right). \tag{15}$$

The iteration step (11) can also be written the following equivalent way,

$$\mathcal{N}(t) = \sum_{\nu=0}^{k} (-1)^{\nu} \binom{\frac{f(t)}{f'(t)}D}{\nu} t, \tag{16}$$

where the binomial coefficient is interpreted as the polynomial $\binom{z}{\nu} = \frac{z(z-1)(z-2)\cdots(z-\nu+1)}{\nu!}$. This can be seen as follows : let

$$\theta_\nu(t) = (-1)^\nu \frac{f(t)^\nu}{\nu!} \left(\frac{1}{f'(t)} D\right)^\nu t, \tag{17}$$

and apply the operator $\frac{1}{f'(t)} D$ on both sides of the equality,

$$\nu!(-f(t))^{-\nu} \theta_\nu(t) = \left(\frac{1}{f'(t)} D\right)^\nu t. \tag{18}$$

Multiplying both sides of the resulting equality by $(-f(t))^{\nu+1}$, one gets after some computation,

$$\theta_{\nu+1}(t) = -\frac{1}{\nu+1} \left(\frac{f(t)}{f'(t)} D - \nu\right) \theta_\nu(t), \tag{19}$$

from which (16) follows immediately.

Corollary 2.3 *Let f be analytic around the simple root a. Then, for every g, analytic around a, the following identities hold,*

$$g(a) = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{f(t)^\nu}{\nu!} \left(\frac{1}{f'(t)} D\right)^\nu g(t) \tag{20}$$

$$= \sum_{\nu=0}^{\infty} (-1)^\nu \binom{\frac{f(t)}{f'(t)} D}{\nu} g(t), \tag{21}$$

whenever t belongs to a suitably small neigborhood of a.

Proof: (Sketch) Use the analytical version of the inverse function theorem, apply Taylor expansion to the function $\Psi_t(u) = g \circ f^{<-1>}(f(t) + u)$, in ascending powers of u, and use the fact that $g(a) = \Psi_t(u)|_{u:=-f(t)}$ for every t sufficiently near to a. □

It is interesting to note that the function defined by the right members of (20-21) is a *constant function* in a neigborhood of a. Examples of this phenomenon and of Proposition 2.2 will be given in Section 4.

3 A combinatorial approach

Recall that a combinatorial species in the sense of Joyal is essentially a class of combinatorial structures which is closed under arbitrary relabellings of their underlying sets [ii]. Given a combinatorial species, R, the species, $A = A(X)$, of *R-enriched rooted trees* is recursively defined by the combinatorial equation,

$$A = XR(A), \tag{22}$$

where X denotes the species of *singletons.* Figure 1 describes an A-structure, where the black dots denote X-singletons and each circular arc, centered at a black dot, denotes an R-structure put on the (possibly empty) set of its children.

[ii] Technically speaking, a species is an endofunctor of the category of finite sets with bijections. See Joyal (1981) and Bergeron et al. (1998)

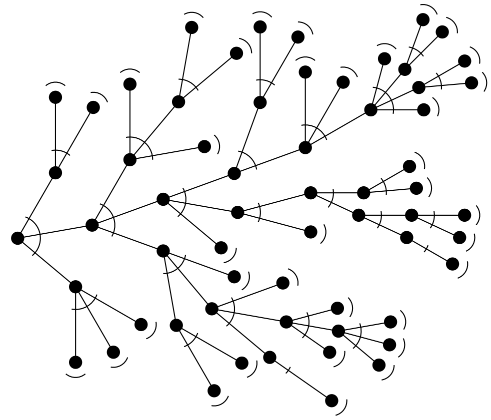


Fig. 1: An R-enriched rooted tree

The species of R-enriched rooted trees was introduced by the author (see Labelle (1981)) in order to give a combinatorial proof of Lagrange inversion. Equivalently, considering X as a combinatorial parameter, the species, $A = A(X)$, can be seen as the solution of the combinatorial equation

$$F(T) = 0, \quad \text{where} \quad F(T) = T - XR(T). \tag{23}$$

Later, Décoste, Labelle, and Leroux (see Décoste et al. (1982)) used R-enriched rooted trees to give a combinatorial proof of the classical (order 2) Newton-Raphson iterative scheme (2). They also sketched a combinatorial model for higher order iterations but their approach was non explicit and computationally unsatisfactory. The purpose of the present section is to give a combinatorial proof of the general (arbitrary order) Newton-Raphson iterative scheme (11) and of the full expansion (20).

Let $D = d/dT$ denote the usual combinatorial differentiation operator with respect to singletons of sort T. Since $-F(T) = XR(T) - T$ and $DF(T) = F'(T) = 1 - XR'(T)$, the analytical iterative scheme (11) suggests the following combinatorial iterative scheme, $T \mapsto \mathcal{N}(T)$, where

$$\mathcal{N}(T) = \sum_{\nu=0}^{k} \frac{1}{\mathfrak{S}_\nu} (XR(T) - T)^\nu \left(\frac{1}{1 - XR'(T)} D \right)^\nu T, \tag{24}$$

and where $\mathfrak{S}_\nu$ denotes the symmetric group of order ν. In other words, the right-hand side of (24) should be a finite sum of quotient species under some actions of the symmetric groups $\mathfrak{S}_\nu, \nu = 0, 1, \dots, k$, into which the successive approximations of A should be substituted for T.

The first approximation T_0 should be of the form $T_0 = \alpha$, where $\alpha = A|_{\leq m}$ denotes the species of R-enriched rooted trees restricted to *small sets,* say of cardinalities $\leq m$, the value of m being fixed. Following the terminology of Décoste et al. (1982), α-structures will be called *light R-enriched rooted trees*. An A-structure which is not light is called *heavy.* Convergence of order $k+1$ is interpreted combinatorially as follows:

$$\alpha = A|_{\leq m} \Rightarrow \mathcal{N}(\alpha)|_{\leq (k+1)(m+1)} = A|_{\leq (k+1)(m+1)}. \tag{25}$$

In other words, if α coincides with the species A on sets of cardinality $\leq m$ then $\mathcal{N}(\alpha)$ will coincide with the species A on sets of cardinality $\leq (k+1)(m+1)$. In order to make this statement more precise, we introduce the auxiliary concepts.

Given m, the species, $B = B(X)$, of *m-broccolis,* is defined by $B = XR(\alpha) - \alpha$. In other words, a m-broccoli is an heavy A-structure consisting of a root (of sort X) followed by an R-assembly of light R-enriched rooted trees, see Figure 2, where $m = 6$.

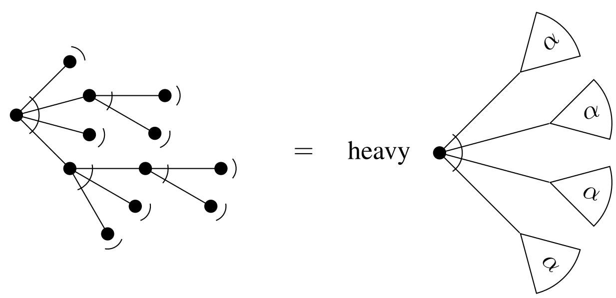


Fig. 2: A m-broccoli for $m = 6$

Using the terminology of Labelle (1985), the differential operator ,

$$\mathcal{D} = \frac{1}{1 - XR'(T)} D, \tag{26}$$

can be called an *eclosion* operator. Since $1/(1 - XR'(T))$ is the species of lists of $XR'(T)$-structures, the eclosion operator $\mathcal{D}$ transforms any species $K = K(X,T)$ to another species $\mathcal{D}K = \mathcal{D}K(X,T)$ as shown in Figure 3, where the T-singletons are represented by black triangles.

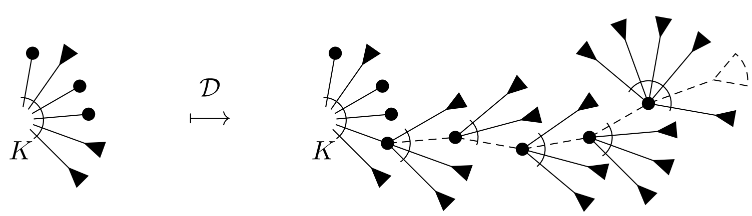


Fig. 3: The eclosion operator $\mathcal{D}$ applied to a species K

Now consider an R-enriched rooted tree τ on a set (of X-singletons) of cardinality $\leq (k+1)(m+1)$. Let ν be the number of broccolis contained in τ. Then $0 \leq \nu \leq k$, since each broccoli contains at least $m+1$ points. Number arbitrarily these broccolis from 1 to ν as in Figure 4 (where $m = 6$, $\nu = 3$).

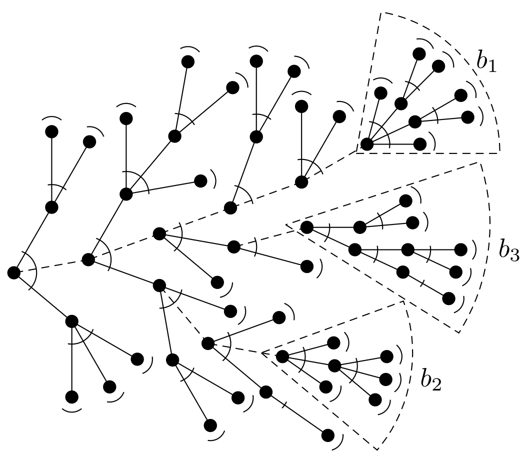


Fig. 4: Numbering the broccolis

Detach next the numbered broccolis and put them as a list $b_1, b_2, \ldots, b_\nu$ as in Figure 5. This shows that for $0 \leq \nu \leq k$, the species $A^{[\nu]}$ of R-enriched rooted trees having exactly ν broccolis numbered 1 to ν coincides with the species

$$\left[(XR(T)-T)^\nu \left(\frac{1}{1-XR'(T)}D\right)^\nu T\right]_{T:=\alpha}, \tag{27}$$

on sets of cardinality $\leq (k+1)(m+1)$.

Finally, since the symmetric group $\mathfrak{S}_\nu$ acts faithfully on the $A^{[\nu]}$-structures by simply renumbering the broccolis, we can erase the numbering by considering the quotient species of $A^{[\nu]}/\mathfrak{S}_\nu$. Summarizing, we can state the following combinatorial version of the Newton-Raphson scheme of arbitrary orders.

Proposition 3.1 *Let $m \geq 0$ be a fixed integer, $D = d/dT$, and consider a species α coinciding with the species $A = XR(A)$ of R-enriched rooted trees on sets up to cardinality m. Then, the species*

$$\mathcal{N}(\alpha) = \sum_{\nu=0}^{k} \frac{1}{\mathfrak{S}_\nu}(XR(\alpha)-\alpha)^\nu \left[\left(\frac{1}{1-XR'(T)}D\right)^\nu T\right]_{T:=\alpha}, \tag{28}$$

coincides with the species A on sets up to cardinality $(k+1)(m+1)$.

□

The following combinatorial analogue of (20) also holds. The proof is similar and left to the reader.

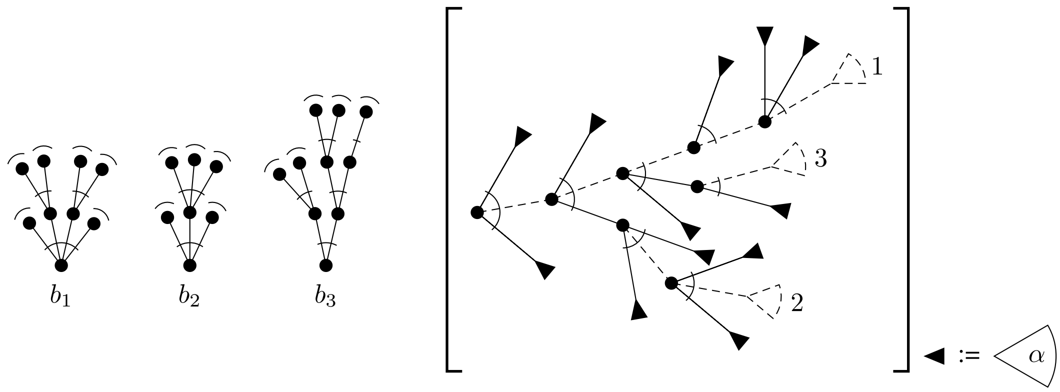


Fig. 5: $\left[(XR(\alpha)-\alpha)^3\mathcal{D}^3T\right]_{T:=\alpha}$-structure

Corollary 3.2 *Let $m \geq 0$ be a fixed integer, $D = d/dT$, and consider a species α coinciding with the species $A = XR(A)$ of R-enriched rooted trees on sets up to cardinality m. Then, for any species G we have the following expansion*

$$G(A) = \sum_{\nu=0}^{\infty} \frac{1}{\mathfrak{S}_\nu}(XR(\alpha)-\alpha)^\nu \left[\left(\frac{1}{1-XR'(T)}D\right)^\nu G(T)\right]_{T:=\alpha}. \tag{29}$$

□

Note that since there is no A-structure on the empty set, we can take, in particular, $m = 0$ and $\alpha = 0$ (the empty species), in Corollary 3.2. In this case, broccolis are (R-enriched) leaves in expansion (29). Special instances of Proposition 3.1 and of Corollary 3.2 will be given in the next section.

4 Examples and applications

4.1 Analytical examples

The analytical iterative scheme (11) and the corresponding full expansions (20-21) can be illustrated in a variety of ways. We now give three typical such illustrations.

• *Root extraction.* Let $0 \neq n \in \mathbb{Z}$, $0 < c \in \mathbb{R}$, and consider the equation $t^n - c = 0$, whose solution is $a = c^{1/n}$. Then, for t near $c^{1/n}$, the $(k+1)^{\text{th}}$ order iteration step (11) takes the form,

$$\mathcal{N}(t) = \sum_{\nu=0}^{k}(-1)^\nu \binom{\frac{1}{n}}{\nu} t\left(1-\frac{c}{t^n}\right)^\nu. \tag{30}$$

Of course, when $k=1$, this reduces to the classical $\mathcal{N}(t)=(t+c/t)/2$, if $n=2$, and to $\mathcal{N}(t)=t(2-ct)$, if $n=-1$. It is worthwhile to note that, taking $g(t)=t^m$ and t near $c^{1/n}$, (20) becomes

$$c^{m/n}=\sum_{\nu=0}^{\infty}(-1)^{\nu}\binom{\frac{m}{n}}{\nu}t^m\left(1-\frac{c}{t^n}\right)^{\nu}, \tag{31}$$

in which the right-hand-side corresponds to the full expansion, in ascending powers of $(1-c/t^n)$, of the trivial identity $c^{m/n}=t^m\left(1-(1-c/t^n)\right)^{\frac{m}{n}}$.

• *Computing logarithms.* Let $0<c\in\mathbb{R}$ and consider the equation $e^t-c=0$, whose solution is $a=\ln(c)$. This time, for t near $\ln(c)$, the $(k+1)^{\text{th}}$ order iteration step (11) takes the form,

$$\mathcal{N}(t)=t-\sum_{\nu=1}^{k}\frac{(1-ce^{-t})^{\nu}}{\nu}, \tag{32}$$

and for any analytic function g around $\ln(c)$, (20) becomes,

$$g(\ln(c))=\sum_{\nu=0}^{\infty}(ce^{-t}-1)^{\nu}\binom{D}{\nu}g(t). \tag{33}$$

In particular, if $g(t)=t$, this last equality corresponds to the full expansion, in ascending powers of $(1-ce^{-t})$, of the trivial identity, $\ln(c)=t+\ln(1-(1-ce^{-t}))$.

• *Approximating* π. Consider the simple root $a=\pi$ of the equation $\sin(t)=0$. It turns out that the form (16) of the iteration step is easier to manipulate in this case. Note first that $(f(t)/f'(t))D=\tan(t)D$. It is then easy to check by induction that, for $\nu>0$,

$$\binom{\tan(t)D}{\nu}t=\text{a polynomial function of }\tan(t)\text{ with constant coefficients.} \tag{34}$$

Collecting alike powers of $\tan(t)$ in (16), massive cancellation produces the following $(2p+1)^{\text{th}}$ order iterative scheme, $t_{n+1}=\mathcal{N}(t_n)$, for successive approximations of π, where,

$$\mathcal{N}(t)=t-\tan(t)+\frac{\tan(t)^3}{3}-\frac{\tan(t)^5}{5}+\cdots+(-1)^{2p-1}\frac{\tan(t)^{2p-1}}{2p-1}, \tag{35}$$

and $\frac{3}{4}\pi<t_0<\frac{5}{4}\pi$ (one can take $t_0=3$, for example). Moreover, for analytic g near π, (20) becomes,

$$g(\pi)=\sum_{\nu=0}^{\infty}(-1)^{\nu}\binom{\tan(t)D}{\nu}g(t), \tag{36}$$

whenever t is sufficiently near to π. In particular, if $g(t)=t$, this last equality corresponds to the full expansion, in ascending powers of $\tan(t)$, of the trivial identity, $\pi=t-\arctan(\tan(t))$, $\frac{3}{4}\pi<t\leq\frac{5}{4}\pi$.

4.2 Combinatorial applications

It is well-known that the classical classes of rooted-trees are special instances of R-enriched rooted trees. For example, if $1, X, L, E, E_n, C$, respectively denote the species of *the empty set*, of *singletons*, of *lists*, of *sets*, of *n-sets*, of *oriented cycles*, then,

- the class of *ordinary* (Cayley) rooted trees corresponds to the species $R = E$,
- the class of *topological* rooted trees corresponds to the species $R = E - X$,
- the class of *binary* rooted trees corresponds to the species $R = 1 + X^2$,
- the class of *unary-binary* rooted trees corresponds to the species $R = 1 + X + X^2$,
- the class of *unoriented binary* rooted trees corresponds to the species $R = 1 + E_2$,
- the class of *unoriented unary-binary* rooted trees corresponds to the species $R = 1 + X + E_2$,
- the class of *ordered* rooted trees corresponds to the species $R = L$,
- the class of *mobiles* (see Bergeron et al. (1998)) corresponds to the species $R = 1 + C$,

etc.

For each choice of the species R, the iteration step (28) produces a combinatorial computational scheme of order $k + 1$ for successive approximations of the species $A = XR(A)$ of R-enriched rooted trees. For example, taking $R = 1 + X + E_2$, then, $R'(X) = 1 + X$, and (28) becomes,

$$\mathcal{N}(\alpha) = \sum_{\nu=0}^{k} \frac{1}{\mathfrak{S}_\nu} (X(1 + \alpha + E_2(\alpha)) - \alpha)^\nu \left[\left(\frac{1}{1 - X(1+T)} D \right)^\nu T \right]_{T:=\alpha}, \tag{37}$$

where $D = d/dT$ and α is the species of light A-structures, namely, the class of unoriented unary-binary rooted trees that are living on sets of cardinality $\leq m$. Moreover, for arbitrary R, taking $m = 0$ and $\alpha = 0$ in (29), the 0-broccolis become $XR(0)$-structures (enriched singletons) and we get the following two new full expansions involving the species A of R-enriched rooted trees according to their number ν of leaves:

$$A = \sum_{\nu=0}^{\infty} \frac{1}{\mathfrak{S}_\nu} X^\nu \left[\left(\frac{R(0)}{1 - XR'(T)} D \right)^\nu T \right]_{T:=0}, \tag{38}$$

$$G(A) = \sum_{\nu=0}^{\infty} \frac{1}{\mathfrak{S}_\nu} X^\nu \left[\left(\frac{R(0)}{1 - XR'(T)} D \right)^\nu G(T) \right]_{T:=0}. \tag{39}$$

For example, taking $R = E$, the species or ordinary (Cayley) rooted trees can be written in the form

$$A = \sum_{\nu=0}^{\infty} \frac{1}{\mathfrak{S}_\nu} X^\nu \left[\left(\frac{1}{1 - XE(T)} D \right)^\nu T \right]_{T:=0}, \tag{40}$$

and taking, for G, the species of permutations, S, the species of endofunctions, $\mathrm{End} = S(A)$, can be written in the form

$$\mathrm{End} = \sum_{\nu=0}^{\infty} \frac{1}{\mathfrak{S}_\nu} X^\nu \left[\left(\frac{1}{1 - XE(T)} D \right)^\nu S(T) \right]_{T:=0}. \tag{41}$$

We have the following result concerning the enumeration of $G(A)$-structures according to their number of leaves ($=$ 0-broccolis).

Corollary 4.1 *Let $R(x) = \sum_{n=0}^{\infty} r_n x^n/n!$ and $G(x) = \sum_{n=0}^{\infty} g_n x^n/n!$ be the exponential generating series of the species R and G. Let $\gamma_{n,\nu}$ be the number of G-assemblies of R-enriched rooted trees on $[n]$ having exactly ν leaves. Then, for $\nu \geq 1$,*

$$\sum_{n=0}^{\infty} \gamma_{n,\nu} x^n/n! = \frac{r_0^{\nu} x^{\nu}}{\nu!(1-r_1 x)^{2\nu-1}} p_{\nu}(x), \tag{42}$$

where $p_{\nu}(x) = \omega_{\nu}(x,0)$ is a polynomial, with $\omega_1(x,t) = G'(t)$, and for $\nu > 1$,

$$\omega_{\nu}(x,t) = \left((1 - xR'(t))\frac{\partial}{\partial t} + (2\nu-3)xR''(t)\right)\omega_{\nu-1}(x,t). \tag{43}$$

Proof: Take the underlying power series of (39), and use induction on ν. □

This provides an uniform algorithm for the computation of the sequences $(\gamma_{n,\nu})_{n\geq 0}$ which is easily implementable in computer algebra systems. If r_0 and r_1 are formal variables, they can be interpreted as ”leafs” and ”nodes having one child” counters, respectively. Moreover, since (42) is always a rational function of x, exact and asymptotic expressions for $\gamma_{n,\nu}$ are easily obtained using partial fractions expansions. For special choices of R and G, reccurence (43) can often be simplified, as the following examples show.

• *Counting ordinary rooted trees having ν leaves.* Taking $R = E$ and $G = X$, one finds that $\omega_{\nu}(x,t) = p_{\nu}(xe^t)$. Consequently, the exponential generating series of the species of ordinary rooted trees having exactly ν leaves is given by the rational function, $x^{\nu}(1-x)^{-2\nu+1}p_{\nu}(x)/\nu!$, where,

$$p_1(x) = 1, \quad p_{\nu}(x) = x\left((1-x)p'_{\nu-1}(x) + (2\nu-3)p_{\nu-1}(x)\right), \quad \nu > 1. \tag{44}$$

For small values of ν, the resulting sequences $(\gamma_{n,\nu})_{n\geq 0}$ can be found in Sloane (2009).

• *Counting mobiles having ν leaves.* Taking $R = 1 + C$ and $G = X$, one finds that $\omega_{\nu}(x,t) = Q_{\nu}(x,t)/(1-t)^{2(\nu-1)}$, where $Q_{\nu}(x,t)$ is a polynomial in x and t. Consequently, the exponential generating series of the species of mobiles having exactly ν leaves is given by the rational function, $x^{\nu}(1-x)^{-2\nu+1}q_{\nu}(x)/\nu!$, where, $q_{\nu}(x) = Q_{\nu}(x,0)$, $Q_1(x,t) = 1$, and for $\nu > 1$,

$$Q_{\nu}(x,t) = \left((1-t)(1-t-x)\frac{\partial}{\partial t} + x + (2\nu-4)(1-t)\right)Q_{\nu-1}(x,t). \tag{45}$$

Again, for small values of ν, the resulting sequences $(\gamma_{n,\nu})_{n\geq 0}$ can be found in Sloane (2009).

• *Counting topological rooted trees having ν leaves.* Taking $R = E - X$ and $G = X$, one finds that $\omega_{\nu}(x,t) = P_{\nu}(x, xe^t)$, where $P_{\nu}(x,y)$ is a polynomial in x and y. Since, in this case, $r_1 = 0$, the exponential generating series of the species of topological rooted trees having exactly ν leaves is a *polynomial function* of the form, $x^\nu p_\nu(x)/\nu!$, where, $p_\nu(x) = P_\nu(x,x)$, $P_1(x,y) = 1$, and for $\nu > 1$,

$$P_\nu(x,y) = \left((1+x-y)y\frac{\partial}{\partial y} + (2\nu-3)y\right) P_{\nu-1}(x,y). \tag{46}$$

• *Counting endofunctions according to their number of leaves.* As mentioned before, taking $R = E$ and $G = S$, the species $G(A) = S(A)$ coincides with the species, End, of endofunctions. In this case, being a leaf of an endofunction ϕ is the same as being a leave of a rooted tree in the S-assembly of rooted trees corresponding to ϕ. Equivalently, a leave of an endofunction ϕ is a periodic element having a 1-element fiber or an element having an empty fiber. In this case, it turns out that $\omega_\nu(x,t) = K_\nu(xe^t,t)/(1-t)^{\nu+1}$, where $K_\nu(x,t)$ is a polynomial in x and t. The corresponding exponential generating series of the species of endofunctions having exactly ν leaves is given by the rational function, $x^\nu(1-x)^{-2\nu+1}\epsilon_\nu(x)/\nu!$, where, $\epsilon_\nu(x) = K_\nu(x,0)$, $K_1(x,t) = 1$, and for $\nu > 1$,

$$K_\nu(x,t) = \left((1-x)(1-t)\left(x\frac{\partial}{\partial x} + \frac{\partial}{\partial t}\right) + \nu + (\nu-3)x - (2\nu-3)xt\right) K_{\nu-1}(x,t). \tag{47}$$

Acknowledgements

The author would like to thank Jérôme Tremblay who programmed all the figures directly within the LaTeX language and for his general technical help during the preparation of this document.

References

F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial Species and Tree-Like Structures*. Cambridge University Press, 1998.

H. Décoste, G. Labelle, and P. Leroux. Une approche combinatoire pour l'itération de Newton-Raphson. *Advances in Applied Mathematics*, 3:407 – 416, 1982.

A. S. Householder. *The Numerical Treatment of a Single Nonlinear Equation*. McGraw-Hill, New-York, 1970.

A. Joyal. Une théorie combinatoire des séries formelles. *Advances in Mathematics*, 42:1 – 82, 1981.

G. Labelle. Une nouvelle démonstration combinatoire des formules d'inversion de Lagrange. *Advances in Mathematics*, 42:217 – 247, 1981.

G. Labelle. Eclosions combinatoires appliquées à l'inversion multidimensionnelle des séries formelles. *Journal of Combinatorial Theory, Series A*, 39:52 – 82, 1985.

P. Sebah and X. Gourdon. Newton's method and high order iterations, 2001. URL `http://numbers.computation.free.fr/Constants/constants.html`.

N. Sloane. The On-Line Encyclopedia of Integer Sequences, 2009. URL `http://www.research.att.com/~njas/sequences`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 726–737

Combinatorial formulas for double parabolic R-polynomials

Justin Lambright and Mark Skandera

Lehigh University, Bethlehem, PA, USA

Abstract. The well-known R-polynomials in $\mathbb{Z}[q]$, which appear in Hecke algebra computations, are closely related to certain *modified R-polynomials* in $\mathbb{N}[q]$ whose coefficients have simple combinatorial interpretations. We generalize this second family of polynomials, providing combinatorial interpretations for expressions arising in a much broader class of computations. In particular, we extend results of Brenti, Deodhar, and Dyer to new settings which include parabolic Hecke algebra modules and the quantum polynomial ring.

Résumé. Les bien connues polynômes-R en $\mathbb{Z}[q]$, qui apparaissent dans les calcules d'algébre de Hecke, sont relationés à certaines *polynômes-R modifiés* en $\mathbb{N}[q]$, dont les coefficients ont simples interprétations combinatoires. Nous généralisons cette deuxième famille de polynômes, fournissant des interprétations combinatoires pour les expressions qui se posent dans une catégorie beaucoup plus vaste de calculs. En particulier, nous étendons des résultats de Brenti, Deodhar, et Dyer à des situations nouvelles, qui comprennent modules paraboliques pour l'algébre de Hecke, et l'anneau des polynômes quantiques.

Resumen. Los ilustres polinomios-R en $\mathbb{Z}[q]$, que aparecen en los cálculos del álgebra de Hecke, están relacionados con ciertos *polinomios-R modificados* en $\mathbb{N}[q]$, cuyos coeficientes tienen interpretaciones combinatorias sencillas. Generalizamos esta segunda familia de polinomios, proporcionando interpretaciones combinatorias para las expresiones que surgen en una clase de cálculos más amplia. En particular, se amplian unos resultados de Brenti, Deodhar, y Dyer a nuevas situaciones que incluyen los módulos parabólicos del álgebra de Hecke, y el anillo de polinomios cuánticos.

Keywords: Immanants, Kazhdan-Lusztig polynomials, quantum groups

1 Introduction

An important ingredient in the definition of Kazhdan and Lusztig's basis [KL79] for the Hecke algebra $H_n(q)$ of a Coxeter group W is a map now known as the *bar involution*. Applying this involution to a natural basis of the algebra, one obtains a second basis, related to the first by polynomials $\{R_{u,v}(q)\,|\,u, v \in W\}$ in $\mathbb{Z}[q]$ now known as *R-polynomials*. Alternatively, one may relate this second basis to the first by polynomials $\{\widetilde{R}_{u,v}(q)\,|\,u, v \in W\}$ in $\mathbb{N}[q]$ which we call *modified R-polynomials*. Coefficients of the modified R-polynomials and their combinatorial interpretations were studied by Brenti [Bre94], [Bre97a], [Bre97b], [Bre98], [Bre02], Deodhar [Deo85], and Dyer [Dye93].

Certain $\mathbb{C}[q^{\frac{1}{2}}, q^{-\frac{1}{2}}]$-submodules of $H_n(q)$ called *double parabolic modules* inherit a bar involution from $H_n(q)$, and therefore inherit analogs of R-polynomials called *parabolic R-polynomials*. Also belonging to $\mathbb{Z}[q]$, these parabolic R-polynomials appear in numerous papers, yet somehow have not received the modification and combinatorial interpretation granted to their nonparabolic syblings.

Related to the bar involutions on the type-A Hecke algebra and its parabolic modules is another involution on a certain noncommutative polynomial ring $\mathcal{A}(n;q)$ which we call the *quantum polynomial ring*. This last involution, also called the *bar involution*, is an important ingredient in the definition of a certain *dual canonical basis* of the quantum polynomial ring, related by Hopf algebra duality to Kashiwara's [Kas91] and Lusztig's [Lus90] canonical basis of $\mathfrak{sl}(n,\mathbb{C})$. Again, applying this involution to a natural basis of $\mathcal{A}(n;q)$, one obtains a second basis, related to the first by *inverse R-polynomials* and *inverse parabolic R-polynomials* (equivalently, by modifications of these).

To summarize, we have several algebras with the property that a natural basis and its bar image are related by a transition matrix whose entries are variations of R-polynomials. Using an elementary family of bases of $\mathcal{A}(n;q)$, we show that in all cases, the above entries have simple combinatorial interpretations in terms of walks in the Bruhat order. These interpretations enable us to express all double parabolic analogs of R-polynomials as sums of the nonparabolic polynomials. In all sections, we work specifically in type A, but many of our results carry over to Hecke algebras of other types.

In Section 2 we review definitions concerning the symmetric group $\mathfrak{S}_n$ and Hecke algebra $H_n(q)$ of type A. We define the bar involution on $H_n(q)$, *R*-polynomials, and modified *R*-polynomials. We also define double parabolic analogs of these, thus extending one of the two parabolic conventions appearing in the literature. These polynomials are easily seen to be sums of nonparabolic R-polynomials and modified R-polynomials. In Section 3, we define the quantum polynomial ring $\mathcal{A}(n;q)$, its bar involution, inverse R-polynomials, and modified inverse R-polynomials. We also define double parabolic analogs of these, thus extending the second of the two parabolic conventions appearing in the literature. These polynomials are *not* easily seen to be sums of nonparabolic inverse R-polynomials and modified inverse R-polynomials. In Section 4, we consider various bases of the so-called *immanant* subspace of $\mathcal{A}(n;q)$ and provide combinatorial interpretations for the transition matrices relating these to the natural basis of the subspace. These lead to interpretations in Section 5 for all variations of the R-polyomials mentioned above, and to our main result which expresses double parabolic inverse R-polynomials as sums of nonparabolic inverse R-polynomials.

2 The symmetric group and Hecke algebra

Let $\mathfrak{S}_n$ be the Coxeter group of type A_{n-1}, i.e., the symmetric group on n letters. $\mathfrak{S}_n$ is generated by the standard adjacent transpositions $s_1, \ldots, s_{n-1}$, subject to the relations

$$\begin{aligned} {s_i}^2 &= e && \text{for } i = 1, \ldots, n-1, \\ s_i s_j s_i &= s_j s_i s_j && \text{for } |i-j| = 1, \\ s_i s_j &= s_j s_i && \text{for } |i-j| \geq 2. \end{aligned} \tag{2.1}$$

A standard action of $\mathfrak{S}_n$ on rearrangements of the word $1 \cdots n$ is defined by letting s_i swap the letters in positions i and $i+1$,

$$s_i \circ a_1 \cdots a_n = a_1 \cdots a_{i+1} a_i \cdots a_n. \tag{2.2}$$

For each element $v = s_{i_1} \cdots s_{i_\ell} \in \mathfrak{S}_n$, we define the *one-line* notation of v to be the word $v_1 \cdots v_n = v \circ 1 \cdots n$. Thus the one-line notation of the identity permutation e is $12 \cdots n$. Using this convention, the one-line notation of vw is

$$(vw)_1 \cdots (vw)_n = v \circ (w \circ 1 \cdots n) = w_{v_1} \cdots w_{v_n}. \tag{2.3}$$

Let $\ell(w)$ be the minimum length of an expression for w in terms of generators, and let w_0 denote the longest word in $\mathfrak{S}_n$. Let $\leq$ denote the Bruhat order on $\mathfrak{S}_n$, i.e., $v \leq w$ if every reduced expression for w contains a reduced expression for v as a subword.

The *(Iwahori-)Hecke algebra* $H_n(q)$ of $\mathfrak{S}_n$ is the $\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-algebra generated by the set of *(modified) natural generators*, $\widetilde{T}_{s_1}, \ldots, \widetilde{T}_{s_{n-1}}$, subject to the relations

$$\begin{aligned} \widetilde{T}_{s_i}^2 &= (q^{\frac{1}{2}} - q^{\frac{-1}{2}})\widetilde{T}_{s_i} + 1 && \text{for } i = 1, \ldots, n-1, \\ \widetilde{T}_{s_i}\widetilde{T}_{s_j}\widetilde{T}_{s_i} &= \widetilde{T}_{s_j}\widetilde{T}_{s_i}\widetilde{T}_{s_j} && \text{for } |i-j| = 1, \\ \widetilde{T}_{s_i}\widetilde{T}_{s_j} &= \widetilde{T}_{s_j}\widetilde{T}_{s_i} && \text{for } |i-j| \geq 2. \end{aligned} \tag{2.4}$$

(We follow the notation of [Lus85], using modified generators $\widetilde{T}_{s_i}$ instead of the more common generators $T_{s_i} = q^{\frac{1}{2}}\widetilde{T}_{s_i}$.) If $s_{i_1} \cdots s_{i_\ell}$ is a reduced expression for $w \in \mathfrak{S}_n$ we define $\widetilde{T}_w = \widetilde{T}_{s_{i_1}} \cdots \widetilde{T}_{s_{i_\ell}}$, where $\widetilde{T}_e = 1$. It is known that the definition of $\widetilde{T}_w$ does not depend upon the choice of a reduced expression for w. We shall call the elements $\{\widetilde{T}_w \mid w \in \mathfrak{S}_n\}$ the *(modified) natural basis* of $H_n(q)$. For $u, v \in \mathfrak{S}_n$, we define $\epsilon_{u,v} = (-1)^{\ell(v)-\ell(u)}$ and $q_{u,v} = (q^{\frac{1}{2}})^{\ell(v)-\ell(u)}$.

An involutive automorphism on $H_n(q)$ commonly known as the *bar involution* is defined by

$$\overline{q^{\frac{1}{2}}} = q^{\frac{-1}{2}}, \quad \overline{\widetilde{T}_w} = (\widetilde{T}_{w^{-1}})^{-1}. \tag{2.5}$$

Taking the bar involution of an element of $H_n(q)$ and expanding in terms of the natural basis [KL79], we have

$$\overline{\widetilde{T}_w} = \sum_{v \leq w} \epsilon_{v,w} q_{v,w}^{-1} R_{v,w}(q)\widetilde{T}_v, \tag{2.6}$$

where $\{R_{v,w}(q) \mid v, w \in \mathfrak{S}_n\}$ are polynomials in $\mathbb{Z}[q]$, which are commonly called *R-polynomials*. Modifying the *R*-polynomials by

$$q_{v,w}^{-1} R_{v,w}(q) = \widetilde{R}_{v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}) \tag{2.7}$$

gives us the *modified R-polynomials* $\{\widetilde{R}_{v,w}(q) \mid v, w \in \mathfrak{S}_n\}$, which belong to $\mathbb{N}[q]$. Thus we may rewrite (2.6) as

$$\overline{\widetilde{T}_w} = \sum_{v \leq w} \epsilon_{v,w} \widetilde{R}_{v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}})\widetilde{T}_v. \tag{2.8}$$

Often appearing in the literature are $\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-submodules of $H_n(q)$ spanned by sums of natural basis elements corresponding to cosets of $\mathfrak{S}_n$. For a subset I of generators of $\mathfrak{S}_n$, the subgroup W_I of $\mathfrak{S}_n$ generated by I is said to be *parabolic*. Note that we have $W_\emptyset = \{e\}$ and $W_{\{s_1,\ldots,s_{n-1}\}} = \mathfrak{S}_n$.

Two parabolic subgroups W_I and W_J partition $\mathfrak{S}_n$ into *double cosets* of the form $W_I w W_J$. If $J = \emptyset$ or $I = \emptyset$, these cosets are denoted $W_I w$ and $w W_J$, respectively. Thus, ordinary single cosets are special cases of double cosets. It is known that each double coset is an interval in the Bruhat order, containing a unique maximal element and a unique minimal element. Denote the collections of maximal and minimal coset representatives by $W_+^{I,J}$ and $W_-^{I,J}$, respectively. Denote the longest element of a subgroup W_I by w_0^I.

The $\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-submodule of $H_n(q)$ corresponding to parabolic subgroups W_I, W_J and their double cosets $W_I w W_J$ is the span of certain double coset sums. For each permutation $w \in W_+^{I,J}$, define the element

$$\widetilde{T}'_{W_I w W_J} = \sum_{v \in W_I w W_J} (-q^{\frac{1}{2}})^{\ell(w)-\ell(v)} \widetilde{T}_v. \tag{2.9}$$

Let $H'_{I,J}$ denote the submodule of $H_n(q)$ spanned by these elements,

$$H'_{I,J} = \operatorname{span}_{\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]} \{\widetilde{T}'_{W_I w W_J} \mid w \in W_+^{I,J}\}. \tag{2.10}$$

The bar involution on $H_n(q)$ induces a bar involution on $H'_{I,J}$. Curtis [Cur85] and Du [Du94] showed that the elements $\{\overline{\widetilde{T}'_{W_I w W_J}} \mid w \in W_+^{I,J}\}$ form a basis of $H'_{I,J}$. Expanding this basis in terms of the natural basis, Du showed that we have

$$\overline{\widetilde{T}'_{W_I w W_J}} = \sum_{\substack{v \in W_+^{I,J} \\ v \leq w}} (q^{\frac{1}{2}})^{\ell(w)-\ell(v)} R_{v,w}^{I,J}(q^{-1}) \widetilde{T}'_{W_I v W_J}, \tag{2.11}$$

where $\{R_{v,w}^{I,J} \mid v, w \in W_+^{I,J}\}$ are polynomials belonging to $\mathbb{Z}[q]$. When $I = \emptyset$ or $J = \emptyset$, we call these *single parabolic R-polynomials* and if neither are empty we call them *double parabolic R-polynomials*. Douglass [Dou90] and Deodhar [Deo87] looked at the single parabolic *R*-polynomials, while Du was probably the first to mention the double parabolic versions. Applying the bar involution to both sides of (2.9) and comparing terms, one sees that double parabolic *R*-polynomials are related to ordinary *R*-polynomials by

$$R_{u,w}^{I,J}(q) = \sum_{v \in W_I w W_J} R_{v,w}(q). \tag{2.12}$$

It is often necessary to factor a permutation w in $\mathfrak{S}_n$ in terms of elements of W_I, W_J, and a minimal or maximal representative of the coset $W_I w W_J$. For instance, each element v of a single coset $W_I v$ has a unique factorization $v = uw$ with $u \in W_I$ and $v \in W_-^{I,\emptyset}$. Similarly, each element v of a single coset vW_J has a unique factorization $v = wu$ with $u \in W_J$ and $v \in W_-^{\emptyset,J}$. Factorization in double cosets is a bit more complicated. For $v \in W_I w W_J$, there is not always a unique $h \in W_I$ and $k \in W_J$ such that $v = hwk$. On the other hand, we can define a canonical factorization in terms of a third parabolic subgroup of $\mathfrak{S}_n$. For every $u \in W_-^{I,J}$ we define a set of generators

$$K' = K'(u) = \{s_i \in I \mid s_i u = u s_j \text{for some } s_j \in J\}, \tag{2.13}$$

By definition $W_{K'}$ is contained in W_I and thus we can construct single cosets of the form $wW_{K'}$ within W_I. As before, there are unique maximal and minimal representatives for each coset $wW_{K'} \subset W_I$. Denote the sets of maximal and minimal coset representatives by $(W_I)_+^{\emptyset,K'}$ and $(W_I)_-^{\emptyset,K'}$, respectively. It follows that for $u \in W_-^{I,J}$ each double coset factors as

$$W_I u W_J = (W_I)_-^{\emptyset,K'} W_{K'} u W_J. \tag{2.14}$$

In other words each element v of the double coset $W_I u W_J$ has a unique factorization $v = v_-^I u v^J$ satisfying

$$v_-^I \in (W_I)_-^{\emptyset,K'}, \quad v^J \in W_J. \tag{2.15}$$

Furthermore the length of the word is the sum of the lengths of the factors.

3 The quantum polynomial ring

For each $n > 0$, let the *quantum polynomial ring* $\mathcal{A}(n;q)$ be the noncommutative $\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-algebra generated by n^2 variables $x = (x_{1,1}, \dots, x_{n,n})$ representing matrix entries, subject to the relations

$$\begin{aligned} x_{i,\ell} x_{i,k} &= q^{\frac{1}{2}} x_{i,k} x_{i,\ell}, \\ x_{j,k} x_{i,k} &= q^{\frac{1}{2}} x_{i,k} x_{j,k}, \\ x_{j,k} x_{i,\ell} &= x_{i,\ell} x_{j,k}, \\ x_{j,\ell} x_{i,k} &= x_{i,k} x_{j,\ell} + (q^{\frac{1}{2}} - q^{\frac{-1}{2}}) x_{i,\ell} x_{j,k}, \end{aligned} \tag{3.1}$$

for all indices $1 \leq i < j \leq n$ and $1 \leq k < \ell \leq n$. $\mathcal{A}(n;q)$ often arises in conjunction with the the quantum group $\mathcal{O}_q(SL(n,\mathbb{C}))$. In particular, we have

$$\mathcal{O}_q(SL(n,\mathbb{C})) \cong \mathcal{A}(n;q)/(\det_q(x) - 1), \tag{3.2}$$

where

$$\det_q(x) \underset{\text{def}}{=} \sum_{w \in \mathfrak{S}_n} (-q^{\frac{-1}{2}})^{\ell(w)} x_{1,w_1} \cdots x_{n,w_n} \tag{3.3}$$

is the *quantum determinant*. Notice that $\mathcal{A}(n;1)$ is the commutative polynomial ring $\mathbb{C}[x_{1,1}, \dots, x_{n,n}]$.

We can use the relations above to convert any monomial into a linear combination of monomials in lexicographic order. Thus as a $\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-module, $\mathcal{A}(n;q)$ is spanned by monomials in lexicographic order. $\mathcal{A}(n;q)$ has a natural grading by degree,

$$\mathcal{A}(n;q) = \bigoplus_{r \geq 0} \mathcal{A}_r(n;q), \tag{3.4}$$

where $\mathcal{A}_r(n;q)$ consists of the homogeneous degree r polynomials within $\mathcal{A}(n;q)$. Furthermore, we may decompose each homogeneous component $\mathcal{A}_r(n;q)$ by considering pairs (L,M) of multisets of r integers, written as weakly increasing sequences $1 \leq \ell_1 \leq \cdots \leq \ell_r \leq n$, and $1 \leq m_1 \leq \cdots \leq m_r \leq n$. Let $\mathcal{A}_{L,M}(n;q)$ be the $\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-span of monomials whose row indices and column indices (with multiplicity) are equal to the multisets L and M, respectively. This leads to the multigrading

$$\mathcal{A}(n;q) = \bigoplus_{r \geq 0} \bigoplus_{L,M} \mathcal{A}_{L,M}(n;q). \tag{3.5}$$

The graded component $\mathcal{A}_{[n],[n]}(n;q)$ is spanned by the monomials

$$\{x_{1,w_1} \cdots x_{n,w_n} \mid w \in \mathfrak{S}_n\}. \tag{3.6}$$

Defining $x^{u,v} = x_{u_1,v_1} \cdots x_{u_n,v_n}$ for any $u, v \in \mathfrak{S}_n$, we may express the above basis as $\{x^{e,w} \mid w \in \mathfrak{S}_n\}$. We will call elements of this submodule *(quantum) immanants* and we will call the module itself the *(immanant space)* of $\mathcal{A}(n;q)$.

In general, $\mathcal{A}_{L,M}(n;q)$ is the $\mathbb{C}[q^{\frac{1}{2}}, q^{\frac{-1}{2}}]$-submodule of $\mathcal{A}(n;q)$ spanned by the monomials

$$\{x_{\ell_1,m_{w_1}} \cdots x_{\ell_r,m_{w_r}} \mid w \in \mathfrak{S}_r\} = \{(x_{L,M})^{e,w} \mid w \in \mathfrak{S}_r\}, \tag{3.7}$$

where the generalized submatrix $x_{L,M}$ of x is defined by

$$x_{L,M} = \begin{bmatrix} x_{\ell_1,m_1} & x_{\ell_1,m_2} & \cdots & x_{\ell_1,m_r} \\ x_{\ell_2,m_1} & x_{\ell_2,m_2} & \cdots & x_{\ell_2,m_r} \\ \vdots & \vdots & & \vdots \\ x_{\ell_r,m_1} & x_{\ell_r,m_2} & \cdots & x_{\ell_n,m_r} \end{bmatrix}. \tag{3.8}$$

An involutive automorphism on $\mathcal{A}(n;q)$ commonly known as the *bar involution* is defined by by $\overline{q^{\frac{1}{2}}} = q^{\frac{-1}{2}}$, $\overline{x_{i,j}} = x_{i,j}$ and

$$\overline{x_{a_1,b_1} \cdots x_{a_r,b_r}} = (q^{\frac{1}{2}})^{\alpha(a)-\alpha(b)} x_{a_r,b_r} \cdots x_{a_1,b_1}, \tag{3.9}$$

where $\alpha(a)$ is the number of pairs $i < j$ for which $a_i = a_j$. Equivalently, for $x_{a_1,b_1} \cdots x_{a_r,b_r} \in \mathcal{A}_{L,M}(n;q)$, we have

$$\overline{x_{a_1,b_1} \cdots x_{a_r,b_r}} = q_{w_0^J,w_0^I} x_{a_r,b_r} \cdots x_{a_1,b_1}, \tag{3.10}$$

where

$$\begin{aligned} I = I(L) &= \{s_i \in \mathfrak{S}_r \mid \ell_i = \ell_{i+1}\} \\ J = J(M) &= \{s_j \in \mathfrak{S}_r \mid m_j = m_{j+1}\}. \end{aligned} \tag{3.11}$$

In the immanant space, the bar involution reduces to

$$\overline{x^{e,v}} = x_{n,v_n} \cdots x_{1,v_1} = x^{w_0,w_0 v}. \tag{3.12}$$

Taking the bar involution of an element of $\mathcal{A}_{[n],[n]}(n;q)$ and expanding in terms of the natural basis, we have

$$\overline{x^{e,v}} = \sum_{w \geq v} q_{v,w}^{-1} S_{v,w}(q) x^{e,w}, \tag{3.13}$$

where $\{S_{v,w}(q) \mid v, w \in \mathfrak{S}_n\}$ are polynomials in $\mathbb{Z}[q]$, which we call *inverse R-polynomials*. Modifying these polynomials by

$$q_{v,w}^{-1} S_{v,w}(q) = \widetilde{S}_{v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}) \tag{3.14}$$

gives us the *modified inverse R-polynomials* $\{\widetilde{S}_{v,w}(q) \mid v, w \in \mathfrak{S}_n\}$, which belong to $\mathbb{N}[q]$. Thus we may rewrite (3.13) as

$$\overline{x^{e,v}} = \sum_{w \geq v} \widetilde{S}_{v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}) x^{e,w}. \tag{3.15}$$

In an arbitrary multigraded component $\mathcal{A}_{L,M}(n;q)$ of $\mathcal{A}(n;q)$, for $v \in W_+^{I,J}$ we have

$$\overline{(x_{L,M})^{e,v}} = \sum_{\substack{w \in W_+^{I,J} \\ w \geq v}} \epsilon_{v,w} q_{v,w} S_{v,w}^{I,J}(q^{-1}) x^{e,w}, \tag{3.16}$$

where $\{S^{I,J}_{v,w}(q) \mid v,w \in W^{I,J}_+\}$ are polynomials in $\mathbb{Z}[q]$, which we call *parabolic inverse R-polynomials*. Modifying these polynomials in an analogous way as in the immanant space creates some difficulty, since they must be expressed as functions of two variables, $q^{\frac{1}{2}} - q^{\frac{-1}{2}}$ and $q^{\frac{-1}{2}}$. The algebraic relationship between these two variables causes problems when we try to define *modified parabolic inverse R-polynomials* in a manner analogous to (3.14). Instead for all $v,w \in W^{I,J}_+$, given any reduced expression $s_{i_1}\cdots s_{i_k}$ for u, let us define the polynomials $\{\widetilde{S}^{I,J}_{v,w}(q_1,q_2) \in \mathbb{N}[q_1,q_2] \mid v,w \in W^{I,J}_+\}$ to be the polynomials whose coefficient of $q_1^a q_2^b$ is equal to the number of sequences $(\pi^{(0)},\dots,\pi^{(k)})$ of permutations satisfying

1. $\pi^{(0)} = w_0 v, \pi^{(k)} \in W_I w W_J$,
2. $\pi^{(j)} \in \{\pi^{(j-1)}, s_{i_j}\pi^{(j-1)}\}$ for $j = 1,\dots,k$,
3. $\pi^{(j)} = s_{i_j}\pi^{(j-1)}$ if $s_{i_j}\pi^{(j-1)} > \pi^{(j-1)}$,
4. $\pi^{(j)} = \pi^{(j-1)}$ for exactly a indices j,
5. $\ell(w) - \ell(\pi^{(k)}) = b$.

It will be shown later that this definition will lead to

$$\overline{(x_{L,M})^{e,v}} = \sum_{\substack{w\in W^{I,J}_+ \\ w\geq v}} \widetilde{S}^{I,J}_{v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}})x^{e,w}, \tag{3.17}$$

and therefore

$$\epsilon_{v,w}q_{v,w}S^{I,J}_{v,w}(q^{-1}) = \widetilde{S}^{I,J}_{v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}}), \tag{3.18}$$

as desired.

Unlike the double parabolic R-polynomials $\{R^{I,J}_{v,w}(q) \mid v,w \in W^{I,J}_+\}$, the double parabolic inverse R-polynomials $\{S^{I,J}_{v,w}(q) \mid v,w \in W^{I,J}_+\}$ can not readily be written as sums of nonparabolic polynomials $\{S_{v,w}(q) \mid v,w \in \mathfrak{S}_n\}$. That is, we know of no identity in $\mathcal{A}_{L,M}(n;q)$ analogous Equation (2.9) which might lead to an analog of Equation (2.12) for inverse parabolic R-polynomials. While actions of $H_n(q)$ on submodules of $\mathcal{A}(n;q)$ corresponding to $L = [n]$ or $M = [n]$ can help produce identities for polynomials of the forms $S^{I,\emptyset}_{u,v}(q)$ and $S^{\emptyset,J}_{u,v}(q)$, this method fails in the general double parabolic setting.

Nevertheless, we will succeed in expressing a polynomial $\widetilde{S}^{I,J}_{v,w}(q)$ in terms of nonparabolic polynomials. To do so we will consider various bases of the immanant space $\mathcal{A}_{[n],[n]}(n;q)$.

4 A family of bases for the quantum immanant space

Working in the quantum immanant space $\mathcal{A}_{[n],[n]}(n;q)$, one often obtains a monomial of the form $x^{u,v}$ and wishes to express it in terms of the natural basis. The relations (3.1) imply that we have

$$x^{u,v} \in x^{e,u^{-1}v} + \sum_{w>u^{-1}v} \mathbb{N}[q^{\frac{1}{2}} - q^{\frac{-1}{2}}]x^{e,w}. \tag{4.1}$$

It follows that for each permutation $u \in \mathfrak{S}_n$, the set $\{x^{u,v} \mid v \in \mathfrak{S}_n\}$ is a basis for $\mathcal{A}_{[n],[n]}(n;q)$. Indeed the natural basis and barred natural basis are special cases corresponding to $u = e$ and $u = w_0$, respectively.

To state transition matrices relating all of these bases to the natural basis, let us define polynomials $p_{u,v,w}(q)$ in $\mathbb{N}[q]$ by the equations

$$x^{u,v} = \sum_{w \geq u^{-1}v} p_{u,v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}) x^{e,w}. \tag{4.2}$$

Apparently we have the special cases

$$p_{e,v,w}(q) = \begin{cases} 1 & \text{if } v = w \\ 0 & \text{otherwise,} \end{cases} \qquad p_{w_0,w_0v,w}(q) = \widetilde{S}_{v,w}(q). \tag{4.3}$$

The polynomials $\{p_{u,v,w}(q) \mid u, v, w \in \mathfrak{S}_n\}$ have an elementary combinatorial interpretation. The following result generalizes those of Deodhar [Deo85] and Dyer [Dye93], for the special case $u = w_0$.

Theorem 4.1 *Given any reduced expression $s_{i_1} \cdots s_{i_\ell}$ for u, the coefficient of q^k in $p_{u,v,w}(q)$ is equal to the number of sequences $(\pi^{(0)}, \ldots, \pi^{(\ell)})$ of permutations satisfying*

1. $\pi^{(0)} = v$, $\pi^{(\ell)} = w$.
2. $\pi^{(j)} \in \{s_{i_j}\pi^{(j-1)}, \pi^{(j-1)}\}$ *for* $j = 1, \ldots, \ell$.
3. $\pi^{(j)} = s_{i_j}\pi^{(j-1)}$ *if* $s_{i_j}\pi^{(j-1)} > \pi^{(j-1)}$.
4. $\pi^{(j)} = \pi^{(j-1)}$ *for exactly k values of j.*

Proof: Omitted. □

These sequences of permutations can be thought of as walks in the Bruhat order from v to w, with steps up, steps down, and repeated vertices constrained by the fixed reduced expression for u. We remark that since the definition (4.2) does not depend on the chosen reduced expression for u, Theorem 4.1 implies several sets of walks in the Bruhat order are equinumerous.

Problem 4.2 *Find bijections between the sets of walks in Theorem 4.1 which correspond to different reduced expressions for u.*

An alternate basis for the immanant space consists of the monomials

$$\{x^{w^{-1},e} \mid w \in \mathfrak{S}_n\}. \tag{4.4}$$

Using this fact, we obtain the following identity.

Proposition 4.3 *For all $u, v, w \in \mathfrak{S}_n$, we have*

$$p_{v,u,w^{-1}}(q) = p_{u,v,w}(q). \tag{4.5}$$

Proof: Omitted. □

Theorem 4.1 then implies that two sets of walks in the Bruhat order are equinumerous.

Problem 4.4 *Find a bijective proof of the identity in Proposition 4.3.*

A straightforward argument shows that the polynomials $\{p_{u,v,w}(q) \mid u, v, w \in \mathfrak{S}_n\}$ also describe the expansions of certain products of natural basis elements of $H_n(q)$.

Corollary 4.5 *For u, v in $\mathfrak{S}_n$, we have*

$$\widetilde{T}_{u^{-1}}\widetilde{T}_v = \sum_{w \geq u^{-1}v} p_{u,v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}})\widetilde{T}_w. \tag{4.6}$$

Proof: Omitted. □

5 Main results

The double parabolic inverse R-polynomials $\{S^{I,J}_{v,w}(q) \mid u, v \in W^{I,J}_+\}$ and the modified double parabolic inverse R-polynomials $\{\widetilde{S}^{I,J}_{v,w}(q_1, q_2) \mid u, v \in W^{I,J}_+\}$ satisfy

$$\begin{aligned} S^{I,J}_{u,w}(q) &= \sum_{v \in W_I w W_J} \epsilon_{v,w} q^2_{v,w} S_{u,v}(q), \\ \widetilde{S}^{I,J}_{u,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}}) &= \sum_{v \in W_I w W_J} q^{-1}_{v,w} \widetilde{S}_{u,v}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}). \end{aligned} \tag{5.1}$$

As we have already mentioned, these identies are not easily seen unless $I = \emptyset$ or $J = \emptyset$. Following the results in Section 4, we will obtain these identities by considering various bases of $\mathcal{A}_{L,M}(n; q)$.

The relations (3.1) imply that we have

$$(x_{L,M})^{u,v} \in q^{\frac{k}{2}}(x_{L,M})^{e,u^{-1}v} + \sum_{\substack{w \in W^{I,J}_+ \\ w > u^{-1}v}} \mathbb{N}[q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}}](x_{L,M})^{e,w}. \tag{5.2}$$

It follows that for each permutation $u \in \mathfrak{S}_n$, the set $\{(x_{L,M})^{u,v} \mid u^{-1}v \in W^{I,J}_+\}$ is a basis for $\mathcal{A}_{L,M}(n; q)$. Indeed the natural basis and barred natural basis are special cases corresponding to $u = e$ and $u = w_0$, respectively.

To state transition matrices relating all of these bases to the natural basis, for all $u \in \mathfrak{S}_n$, $v \in W^{\emptyset,J}_-$ and $w \in W^{I,J}_+$, given any reduced expression $s_{i_1} \cdots s_{i_k}$ for u, let us define the Laurent polynomials $p^{I,J}_{u,v,w}(q_1, q_2)$ to be the polynomials whose coefficient of $q_1^a q_2^b$ is equal to the number of sequences $(\pi^{(0)}, \ldots, \pi^{(k)})$ of permutations satisfying

1. $\pi^{(0)} = v$, $\pi^{(k)} \in W_I w W_J$,
2. $\pi^{(j)} \in \{s_{i_j}\pi^{(j-1)}, \pi^{(j-1)}\}$ for $j = 1, \ldots, k$,
3. $\pi^{(j)} = s_{i_j}\pi^{(j-1)}$ if $s_{i_j}\pi^{(j-1)} > \pi^{(j-1)}$,
4. $\pi^{(j)} = \pi^{(j-1)}$ for exactly a values of j,
5. $\ell(w^I_-) - \ell\left((\pi^{(k)})^I_-\right) - \ell\left((\pi^{(k)})^J\right) + \ell(u^I) = b$.

We remark that this definition of $p^{I,J}_{u,v,w}(q_1,q_2)$ depends upon the chosen reduced expression for u, unlike our definition (4.2) of $p_{u,v,w}(q)$. Nevertheless, we will suppress this dependence from the notation.

These sequences of permutations can be thought of as walks in the Bruhat order from v to any permutation in $W_I w W_J$, with steps up, steps down, and repeated vertices constrained by the fixed reduced expression for u. Apparently we have the special cases

$$p^{I,J}_{e,v,w}(q_1,q_2) = \begin{cases} 1 & \text{if } v = w \\ 0 & \text{otherwise,} \end{cases} \qquad p^{I,J}_{w_0,w_0v,w}(q_1,q_2) = q_2^{\ell(w_0^I)-\ell(w_0^J)}\widetilde{S}^{I,J}_{v,w}(q_1,q_2). \tag{5.3}$$

Using the relations (3.1) and facts about double cosets of $\mathfrak{S}_n$, we can show that certan transition matrices consist of the polynomials $\{p^{I,J}_{u,v,w}(q_1,q_2) \mid u \in \mathfrak{S}_n, v \in W^{\emptyset,J}_-, w \in W^{I,J}_+\}$ evaluated at $q_1 = q^{\frac{1}{2}} - q^{\frac{-1}{2}}$ and $q_2 = q^{\frac{-1}{2}}$.

Theorem 5.1 *For $u \in \mathfrak{S}_n$, any reduced expression for u, and $v \in W^{\emptyset,J}_-$,*

$$(x_{L,M})^{u,v} = \sum_{w \in W^{I,J}_+} p^{I,J}_{u,v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}})(x_{L,M})^{e,w}. \tag{5.4}$$

Proof: Omitted. □

Problem 5.2 *Modify the definition of $p^{I,J}_{u,v,w}(q_1,q_2)$ to include all $v \in \mathfrak{S}_n$ in such a way that Theorem 5.1 holds for all $u, v \in \mathfrak{S}_n$.*

The combinatorial defintion above leads to the following identity connecting the parabolic and nonparabolic polynomials.

Corollary 5.3 *For all $u, v \in \mathfrak{S}_n$ and $w \in W^{I,J}_+$,*

$$p^{I,J}_{u,v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}}) = \sum_{z \in W_I w W_J} (q^{\frac{-1}{2}})^{\ell(w_0^I w_0^{K'})-\ell(z^I_-)-\ell(z^J)+\ell(u^I)} p_{u,v,z}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}). \tag{5.5}$$

Proof: Omitted. □

Now, using Equation (5.3) and Theorem 5.1 we can derive the desired identity (3.17) relating double parabolic inverse R-polynomials and modified double parabolic inverse R-polynomials.

Theorem 5.4 *For all $v \in W^{I,J}_+$*

$$\overline{(x_{L,M})^{e,v}} = \sum_{\substack{w \in W^{I,J}_+ \\ w \geq v}} \widetilde{S}^{I,J}_{v,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}})(x_{L,M})^{e,w}. \tag{5.6}$$

Proof: Omitted. □

Finally, using the previous theorem along with Theorem 4.1 and Equations (4.3) and (5.3) we can derive the desired identity (5.1) relating double parabolic and nonparabolic inverse R-polynomials.

Theorem 5.5 *The double parabolic inverse R-polynomials $\{S^{I,J}_{v,w}(q) \mid u, v \in W^{I,J}_+\}$ and their modifications $\{\widetilde{S}^{I,J}_{v,w}(q_1, q_2) \mid u, v \in W^{I,J}_+\}$ satisfy*

$$\begin{aligned} S^{I,J}_{u,w}(q) &= \sum_{v \in W_I w W_J} \epsilon_{v,w} q^2_{v,w} S_{u,v}(q), \\ \widetilde{S}^{I,J}_{u,w}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}, q^{\frac{-1}{2}}) &= \sum_{v \in W_I w W_J} q^{-1}_{v,w} \widetilde{S}_{u,v}(q^{\frac{1}{2}} - q^{\frac{-1}{2}}). \end{aligned} \tag{5.7}$$

Proof: Omitted. □

References

[Bre94] Francesco Brenti. A combinatorial formula for Kazhdan-Lusztig polynomials. *Invent. Math.*, 118(2):371–394, 1994.

[Bre97a] Francesco Brenti. Combinatorial expansions of Kazhdan-Lusztig polynomials. *J. London Math. Soc.*, 55(2):448–472, 1997.

[Bre97b] Francesco Brenti. Combinatorial properties of the Kazhdan-Lusztig R-polynomials for S_n. *Adv. Math.*, 126(1):21–51, 1997.

[Bre98] Francesco Brenti. Kazhdan-Lusztig and R-polynomials from a combinatorial point of view. *Discrete Math.*, 193(1-3):93–116, 1998. Selected papers in honor of Adriano Garsia (Taormina, 1994).

[Bre02] Francesco Brenti. Kazhdan-Lusztig and R-polynomials, Young's lattice, and Dyck partitions. *Pacific J. Math.*, 207(2):257–286, 2002.

[Cur85] Charles W. Curtis. On Lusztig's isomorphism theorem for Hecke algebras. *J. Algebra*, 92(2):348–365, 1985.

[Deo85] Vinay V. Deodhar. On some geometric aspects of Bruhat orderings. I. A finer decomposition of Bruhat cells. *Invent. Math.*, 79(3):499–511, 1985.

[Deo87] Vinay Deodhar. On some geometric aspects of Bruhat orderings. II. The parabolic analogue of Kazhdan-Lusztig polynomials. *J. Algebra*, 111(2):483–506, 1987.

[Dou90] J. M. Douglass. An inversion formula for relative Kazhdan-Luszig polynomials. *Comm. Algebra*, 18(2):371–387, 1990.

[Du94] Jie Du. IC bases and quantum linear groups. In *Algebraic groups and their generalizations: quantum and infinite-dimensional methods (University Park, PA, 1991)*, volume 56 of *Proc. Sympos. Pure Math*, pages 135–148. Amer. Math. Soc., Providence, RI, 1994.

[Dye93] M. J. Dyer. Hecke algebras and shellings of Bruhat intervals. *Compositio Math.*, 89:91–115, 1993.

[Kas91] M. Kashiwara. On crystal bases of the Q-analog of universal enveloping algebras. *Duke Math. J.*, 63:465–516, 1991.

[KL79] D. Kazhdan and G. Lusztig. Representations of Coxeter groups and Hecke algebras. *Invent. Math.*, 53:165–184, 1979.

[Lus85] G. Lusztig. Cells in affine Weyl groups. In *Algebraic Groups and Related Topics (Kyoto/Nagoya 1983)*, volume 6 of *Adv. Stud. Pure Math*, pages 255–287. North-Holland, Amsterdam, 1985.

[Lus90] G. Lusztig. Canonical bases arising from quantized enveloping algebras. *J. Amer. Math. Soc.*, 3:447–498, 1990.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 738–749

Crystals from categorified quantum groups

Aaron D. Lauda[1 †] and Monica Vazirani[2 ‡]

[1] *Department of Mathematics, Columbia University, New York, NY 10027*
[2] *Department of Mathematics, University of California, Davis, Davis, California 95616-8633*

Abstract. We study the crystal structure on categories of graded modules over algebras which categorify the negative half of the quantum Kac-Moody algebra associated to a symmetrizable Cartan data. We identify this crystal with Kashiwara's crystal for the corresponding negative half of the quantum Kac-Moody algebra. As a consequence, we show the simple graded modules for certain cyclotomic quotients carry the structure of highest weight crystals, and hence compute the rank of the corresponding Grothendieck group.

Résumé. Nous étudions la structure cristalline sur les catégories de modules gradués sur algèbres qui categorify la moitié négative du quantum de Kac-Moody algèbre associée à un ensemble de data symétrisables Cartan. Nous identifions ce cristal avec des cristaux de Kashiwara pour le négatif correspondant la moitié de l'algèbre de Kac-Moody quantum. En conséquence, nous montrons la simples modules classés pour certains quotients cyclotomique porter le structure des cristaux de poids le plus élevé, et donc de calculer le rang de le groupe correspondant Grothendieck.

Keywords: Khovanov-Lauda-Rouquier algebras, quiver Hecke algebras, categorification

1 Introduction

In [KL09, KL08a] a family R of graded algebras was introduced that categorifies the integral form ${}_{\mathcal{A}}\mathbf{U}_q^- := {}_{\mathcal{A}}\mathbf{U}_q^-(\mathfrak{g})$ of the negative half of the quantum enveloping algebra $\mathbf{U}_q(\mathfrak{g})$ associated to a symmetrizable Kac-Moody algebra $\mathfrak{g}$. Similar algebras were also independently introduced by Rouquier [Rou08]. The grading on these algebras equips the Grothendieck group $K_0(R{-}\mathrm{pmod})$ of the category of finitely-generated graded projective R-modules with the structure of a $\mathbb{Z}[q,q^{-1}]$-module. Natural parabolic induction and restriction functors give $K_0(R{-}\mathrm{pmod})$ the structure of a (twisted) $\mathbb{Z}[q,q^{-1}]$-bialgebra. In [KL09, KL08a] an explicit isomorphism of twisted bialgebras was given between ${}_{\mathcal{A}}\mathbf{U}_q^-$ and $K_0(R{-}\mathrm{pmod})$.

Several conjectures were also made in [KL09, KL08a]. One conjecture that remains unsolved is the so called cyclotomic quotient conjecture which suggests a close connection between certain finite dimensional quotients of the algebras R and the integrable representation theory of quantum Kac-Moody algebras. While this conjecture has been proven in finite and affine type A by Brundan and Kleshchev [BK09], very little is known in the case of an arbitrary symmetrizable Cartan datum. Here we show that simple graded modules for these cyclotomic quotients carry the structure of highest weight crystals. Hence we

† partially supported by the NSF grants DMS-0739392 and DMS-0855713
‡ partially supported by the NSA grant #H982300910076

identify the rank of the corresponding Grothendieck group with the rank of the integral highest weight representation, thereby laying to rest a major component of the cyclotomic conjecture.

This submission is an extended abstract of the preprint [LV09]. We refer the reader there the complete proofs and details, as well as more background and context on these results.

2 The algebra $R(\nu)$

We are given a Cartan data: the weight lattice P, simple roots α_i indexed by $i \in I$, simple coroots $h_i \in P^\vee = \mathrm{Hom}_{\mathbb{Z}}(P,\mathbb{Z})$, a bilinear form $(\,,\,)\colon P \times P \to \mathbb{Z}$, and the canonical pairing $\langle \cdot,\cdot \rangle\colon P^\vee \times P \to \mathbb{Z}$, such that $[\langle h_i, \alpha_j\rangle]_{i,j\in I}$ is a symmetrizable generalized Cartan matrix. In what follows we write $a_{ij} = -\langle i,j\rangle := -\langle h_i,\alpha_j\rangle$ for $i,j \in I$. Let $\Lambda_i \in P^+$ be the fundamental weights defined by $\langle h_j, \Lambda_i\rangle = \delta_{ij}$. Let $q_i = q^{\frac{(\alpha_i,\alpha_i)}{2}}$, $[a]_i = \frac{q_i^a - q_i^{-a}}{q_i - q_i^{-1}}$, $[a]_i! = [a]_i[a-1]_i \dots [1]_i$.

Recall the definition from [KL09, KL08a] of the algebra R associated to a Cartan datum. Let $\Bbbk$ be an algebraically closed field (of arbitrary characteristic).

For $\nu = \sum_{i\in I} \nu_i \cdot i \in \mathbb{N}[I]$ let $\mathrm{Seq}(\nu)$ be the set of all sequences of vertices $\boldsymbol{i} = i_1 \dots i_m$ where $i_r \in I$ for each r and vertex i appears ν_i times in the sequence. The length m of the sequence is equal to $|\nu| = \sum_{i\in I}\nu_i$. It is sometimes convenient to identify $\nu = \sum_{i\in I}\nu_i \cdot i \in \mathbb{N}[I]$ as $\nu \in \sum_{i\in I}\nu_i\alpha_i \in Q_+ = \oplus_{i\in I}\mathbb{Z}_{\geq 0}\alpha_i$.

For $\nu \in \mathbb{N}[I]$ with $|\nu| = m$, let $R(\nu)$ denote the associative, $\Bbbk$-algebra on generators $1_{\boldsymbol{i}}$ for $\boldsymbol{i} \in \mathrm{Seq}(\nu)$, x_r for $1 \leq r \leq m$, and ψ_r for $1 \leq r \leq m-1$ subject to the following relations for $\boldsymbol{i},\boldsymbol{j} \in \mathrm{Seq}(\nu)$:

$$1_{\boldsymbol{i}}1_{\boldsymbol{j}} = \delta_{\boldsymbol{i},\boldsymbol{j}}1_{\boldsymbol{i}}, \qquad x_r 1_{\boldsymbol{i}} = 1_{\boldsymbol{i}} x_r, \qquad \psi_r 1_{\boldsymbol{i}} = 1_{s_r(\boldsymbol{i})}\psi_r,$$
$$x_r x_t = x_t x_r, \qquad \psi_r\psi_t = \psi_t\psi_r \quad \text{if } |r-t|>1,$$

$$\psi_r\psi_r 1_{\boldsymbol{i}} = \begin{cases} 0 & \text{if } i_r = i_{r+1} \\ 1_{\boldsymbol{i}} & \text{if } (\alpha_{i_r},\alpha_{i_{r+1}}) = 0 \\ \left(x_r^{-\langle i_r,i_{r+1}\rangle} + x_{r+1}^{-\langle i_{r+1},i_r\rangle}\right)1_{\boldsymbol{i}} & \text{if } (\alpha_{i_r},\alpha_{i_{r+1}}) \neq 0 \text{ and } i_r \neq i_{r+1}, \end{cases}$$

$$\left(\psi_r\psi_{r+1}\psi_r - \psi_{r+1}\psi_r\psi_{r+1}\right)1_{\boldsymbol{i}} = \\ = \begin{cases} \displaystyle\sum_{t=0}^{-\langle i_r,i_{r+1}\rangle-1} x_r^t x_{r+2}^{-\langle i_r,i_{r+1}\rangle-1-t}1_{\boldsymbol{i}} & \text{if } i_r = i_{r+2} \text{ and } (\alpha_{i_r},\alpha_{i_{r+1}}) \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

$$\left(\psi_r x_t - x_{s_r(t)}\psi_r\right)1_{\boldsymbol{i}} = \begin{cases} 1_{\boldsymbol{i}} & \text{if } t = r \text{ and } i_r = i_{r+1} \\ -1_{\boldsymbol{i}} & \text{if } t = r+1 \text{ and } i_r = i_{r+1} \\ 0 & \text{otherwise.} \end{cases}$$

The algebra is graded with generators defined to have degrees $\deg(1_{\boldsymbol{i}}) = 0$, $\deg(x_r 1_{\boldsymbol{i}}) = (\alpha_{i_r},\alpha_{i_r})$, and $\deg(\psi_r 1_{\boldsymbol{i}}) = -(\alpha_{i_r},\alpha_{i_{r+1}})$.

We let the (nonunital) algebra R be defined by $R = \bigoplus_{\nu\in\mathbb{N}[I]} R(\nu)$.

Rouquier has independently defined a generalization of the algebras R, where the relations depend on Hermitian matrices [Rou08].

Remark 2.1 *For $\boldsymbol{i},\boldsymbol{j} \in \mathrm{Seq}(\nu)$ let ${}_{\boldsymbol{j}}S_{\boldsymbol{i}}$ be the subset of S_m consisting of permutations w that take $\boldsymbol{i}$ to $\boldsymbol{j}$ via the standard action of permutations on sequences. Denote the subset $\{\widehat{w}\}_{w\in{}_jS_i}$ of $1_{\boldsymbol{j}}R1_{\boldsymbol{i}}$ by ${}_{\boldsymbol{j}}\widehat{S}_{\boldsymbol{i}}$. It was shown in [KL09, KL08a] that the vector space $1_{\boldsymbol{j}}R(\nu)1_{\boldsymbol{i}}$ has a basis consisting of elements of the form $\{\psi_{\widehat{w}} \cdot x_1^{a_1}\cdots x_m^{a_m}1_{\boldsymbol{i}} \mid \widehat{w} \in {}_{\boldsymbol{j}}\widehat{S}_{\boldsymbol{i}},\;\; a_r \in \mathbb{Z}_{\geq 0}\}$.*

Let w_0 denote the longest element of $S_{|\nu|}$. We define an involution $\sigma = \sigma_\nu \colon R(\nu) \to R(\nu)$ by $\sigma(1_{\boldsymbol{i}}) = 1_{w_0(\boldsymbol{i})}$, $\sigma(x_r) = x_{|\nu|+1-r}$, $\sigma(\psi_r) = \psi_{|\nu|-r}$. Given an $R(\nu)$-module M, we let σ^*M denote the $R(\nu)$-module whose underlying set is M but with twisted action $r \cdot u = \sigma(r)u$.

Define the character $\mathrm{ch}(M)$ of an $R(\nu)$-module M as $\mathrm{ch}(M) = \sum_{\boldsymbol{i}\in\mathrm{Seq}(\nu)} \mathrm{gdim}(1_{\boldsymbol{i}}M)\cdot\boldsymbol{i}$, where gdim denotes the graded dimension. When M is finite dimensional, $\mathrm{ch}(M)$ is an element of the free $\mathbb{Z}[q,q^{-1}]$-module with basis $\mathrm{Seq}(\nu)$.

3 Functors on the modular category

Let $R(\nu)-\mathrm{fmod}$ be the category of finite dimensional graded $R(\nu)$-modules. The morphisms are grading-preserving module homomorphisms. Note that this category contains all of the simples. Henceforth, by an $R(\nu)$-module we will mean a finite dimensional graded $R(\nu)$-module, unless we say otherwise. We will denote the zero module by $\mathbf{0}$. Let $R-\mathrm{fmod} \overset{\text{def}}{=} \bigoplus_{\nu\in\mathbb{N}[I]} R(\nu)-\mathrm{fmod}$.

For any two $R(\nu)$-modules M, N denote by $\mathrm{Hom}(M,N)$ or $\mathrm{Hom}_{R(\nu)}(M,N)$ the $\Bbbk$-vector space of degree preserving homomorphisms, and by $\mathrm{Hom}(M\{r\},N) = \mathrm{Hom}(M,N\{-r\})$ the space of homogeneous homomorphisms of degree r. Here $N\{r\}$ denotes N with the grading shifted up by r, so that $\mathrm{ch}(N\{r\}) = q^r\mathrm{ch}(N)$. Then we write $\mathrm{HOM}(M,N) := \bigoplus_{r\in\mathbb{Z}} \mathrm{Hom}(M,N\{r\})$, for the $\mathbb{Z}$-graded $\Bbbk$-vector space of all $R(\nu)$-module morphisms.

Though it is essential to work with the degree preserving morphisms to get the $\mathbb{Z}[q,q^{-1}]$-module structure for the categorification theorems in [KL09, KL08a], for our purposes it will often be convenient to work with degree homogenous morphisms, but not necessarily degree preserving. Since any homogenous morphism can be interpreted as a degree preserving morphism by shifting the grading on the source or target, all results stated using homogeneous morphisms can be recast as degree zero morphisms for an appropriate shift on the source or target.

3.1 Induction and Restriction functors

There is an inclusion of graded algebras

$$\iota_{\nu,\nu'} \;:\; R(\nu)\otimes R(\nu') \hookrightarrow R(\nu+\nu')$$

taking the idempotent $1_{\boldsymbol{i}}\otimes 1_{\boldsymbol{j}}$ to $1_{\boldsymbol{ij}}$ and the unit element $1_\nu \otimes 1_{\nu'}$ to an idempotent of $R(\nu+\nu')$ denoted $1_{\nu,\nu'}$. This inclusion gives rise to restriction and induction functors denoted by $\mathrm{Res}_{\nu,\nu'}$ and $\mathrm{Ind}_{\nu,\nu'}$, respectively. When it is clear from the context, or when no confusion is likely to arise, we often simplify notation and write Res and Ind.

We can also consider these notions for any tuple $\underline{\nu} = (\nu^{(1)},\nu^{(2)},\ldots,\nu^{(k)})$ and sometimes refer to the image $R(\underline{\nu}) \overset{\text{def}}{=} \mathrm{Im}\,\iota_{\underline{\nu}} \subseteq R(\nu^{(1)}+\cdots+\nu^{(k)})$ as a parabolic subalgebra. This subalgebra has identity $1_{\underline{\nu}}$. Let $\mu = \nu^{(1)}+\cdots+\nu^{(k)}$, $m = \sum_r |\nu^{(r)}|$, and $P = P_{\underline{\nu}}$ be the composition $(|\nu^{(1)}|,\ldots,|\nu^{(k)}|)$ of m so that S_P is the corresponding parabolic subgroup of S_m. It follows from Remark 2.1 that $R(\mu)1_{\underline{\nu}}$ is a free right $R(\underline{\nu})$-module with basis $\{\psi_{\widehat{w}}1_{\underline{\nu}} \mid w \in S_m/S_P\}$ and $1_{\underline{\nu}}R(\mu)$ is a free left $R(\underline{\nu})$-module with basis $\{1_{\underline{\nu}}\psi_{\widehat{w}} \mid w \in S_P\backslash S_m\}$. By abuse of notation we will write S_m/S_P to denote the minimal length left coset representatives, i.e. $\{w \in S_m \mid \ell(wv) = \ell(w) + \ell(v)\,\forall\, v \in S_P\}$, and $S_P\backslash S_m$ for the minimal length right coset representatives.

Remark 3.1 *It is easy to see that if M is an $R(\underline{\nu})$-module with basis $\mathcal{U}$ consisting of weight vectors, then $\{\psi_{\widehat{w}} \otimes u \mid u \in \mathcal{U}, w \in S_m/S_P\}$ is a weight basis of $\mathrm{Ind}_{\underline{\nu}} M \overset{def}{=} R(\mu) \otimes_{R(\underline{\nu})} M$ (where for each w we fix just one reduced expression $\widehat{w}$). Note $R(\mu) \otimes_{R(\underline{\nu})} M = R(\mu)1_{\underline{\nu}} \otimes_{R(\underline{\nu})} M$ since $\psi_{\widehat{w}}1_{\underline{\nu}} \otimes u = \psi_{\widehat{w}} \otimes 1_{\underline{\nu}}u = \psi_{\widehat{w}} \otimes u$.*

One extremely important property of the functor $\mathrm{Ind}_{\underline{\nu}} - \overset{\text{def}}{=} R(\mu) \otimes_{R(\underline{\nu})} -$ is that it is left adjoint to restriction, i.e., there is a functorial isomorphism $\mathrm{HOM}_{R(\mu)}(\mathrm{Ind}_{\underline{\nu}} A, B) \cong \mathrm{HOM}_{R(\underline{\nu})}(A, \mathrm{Res}_{\underline{\nu}} B)$ where A, B are finite dimensional $R(\underline{\nu})$- and $R(\mu)$-modules, respectively. We refer to this property as Frobenius reciprocity and use it repeatedly, often for deducing information about characters.

3.2 Refining the restriction functor

For M in $R(\nu)-\mathrm{mod}$ and $i \in I$ let $\Delta_i M = (1_{\nu-i} \otimes 1_i)M = \mathrm{Res}_{\nu-i,i} M$, and, more generally, $\Delta_{i^n} M = (1_{\nu-ni} \otimes 1_{ni})M = \mathrm{Res}_{\nu-ni,ni} M$. We view Δ_{i^n} as a functor into the category $R(\nu-ni) \otimes R(ni)-\mathrm{mod}$. By Frobenius reciprocity, there are functorial isomorphisms

$$\mathrm{HOM}_{R(\nu)}(\mathrm{Ind}_{\nu-ni,ni} N \boxtimes L(i^n), M) \cong \mathrm{HOM}_{R(\nu-ni)\otimes R(ni)}(N \boxtimes L(i^n), \Delta_{i^n} M),$$

for M as above and $N \in R(\nu - ni)-\mathrm{mod}$.

Define

$$e_i := \mathrm{Res}^{\nu-i,i}_{\nu-i} \circ \Delta_i \colon R(\nu)-\mathrm{mod} \to R(\nu - i)-\mathrm{mod} \tag{1}$$

and for M an irreducible $R(\nu)$-module, set

$$\widetilde{e}_i M := \mathrm{soc}\, e_i M, \qquad \widetilde{f}_i M := \mathrm{cosoc}\ \mathrm{Ind}^{\nu+i}_{\nu,i} M \boxtimes L(i), \qquad \varepsilon_i(M) := \max\{n \geq 0 \mid \widetilde{e}_i^{\,n} M \neq \mathbf{0}\}.$$

We also define their σ-symmetric versions, which are indicated with a $\vee$. Note that $\sigma^*(\Delta_i(\sigma^* M)) = \mathrm{Res}_{i,\nu-i} M$. Set $e_i^\vee := \sigma^*(e_i(\sigma^* M))$, $\widetilde{e}_i^{\,\vee} M := \sigma^*(\widetilde{e}_i(\sigma^* M))$, $\widetilde{f}_i^{\,\vee} M := \sigma^*(\widetilde{f}_i(\sigma^* M))$, $\varepsilon_i^\vee(M) := \varepsilon_i(\sigma^* M)$. Observe that the functors e_i and $e_i^\vee$ are exact. It is a theorem of [KL09] that if M is irreducible, so are $\widetilde{f}_i M$ and $\widetilde{e}_i M$ (so long as the latter is nonzero), and likewise for $\widetilde{f}_i^{\,\vee} M$ and $\widetilde{e}_i^{\,\vee} M$. For other key properties of the functors $\widetilde{e}_i$ and $\widetilde{f}_i$ on simple modules, see [KL09] or [LV09].

3.3 The cyclotomic algebras $R^\Lambda(\nu)$

For $\Lambda = \sum_{i\in I} \lambda_i \Lambda_i \in P^+$ consider the two-sided ideal $\mathcal{J}^\Lambda_\nu$ of $R(\nu)$ generated by elements $(x_1 1_{\boldsymbol{i}})^{\lambda_{i_1}}$ over all sequences $\boldsymbol{i} \in \mathrm{Seq}(\nu)$. We sometimes write $\mathcal{J}^\Lambda_\nu = \mathcal{J}^\Lambda$ when no confusion is likely to arise. Define

$$R^\Lambda(\nu) := R(\nu)/\mathcal{J}^\Lambda_\nu$$

By analogy with the Ariki-Koike cyclotomic quotient of the affine Hecke algebra [AK94] (see also [Ari02]) this algebra is called the cyclotomic quotient at weight Λ of $R(\nu)$. As above we form the non-unital ring

$$R^\Lambda = \bigoplus_{\nu\in\mathbb{N}[I]} R^\Lambda(\nu).$$

For bookkeeping purposes we will denote $R^{\Lambda}(\nu)$ modules by $\mathcal{M}$ but $R(\nu)$-modules by M.

We introduce functors

$$\mathrm{infl}_{\Lambda}\colon R^{\Lambda}(\nu)\mathrm{-mod} \to R(\nu)\mathrm{-fmod} \qquad \mathrm{pr}_{\Lambda}\colon R(\nu)\mathrm{-fmod} \to R^{\Lambda}(\nu)\mathrm{-mod}$$

where infl_{Λ} is the inflation along the epimorphism $R(\nu) \to R^{\Lambda}(\nu)$, so that $\mathcal{M} = \mathrm{infl}_{\Lambda}\mathcal{M}$ on the level of sets. If $\mathcal{M},\mathcal{N}$ are $R^{\Lambda}(\nu)$-modules, then $\mathrm{Hom}_{R^{\Lambda}(\nu)}(\mathcal{M},\mathcal{N}) \cong \mathrm{Hom}_{R(\nu)}(\mathrm{infl}_{\Lambda}\mathcal{M}, \mathrm{infl}_{\Lambda}\mathcal{N})$. Note $\mathcal{M}$ is irreducible if and only if $\mathrm{infl}_{\Lambda}\mathcal{M}$ is. We define $\mathrm{pr}_{\Lambda} M = M/\mathcal{J}^{\Lambda}M$. If M is irreducible then $\mathrm{pr}_{\Lambda} M$ is either irreducible or zero. Observe infl_{Λ} is an exact functor and its left adjoint is pr_{Λ} which is only right exact.

A careful study of the modules $L(i^m)$ yields that for simple modules M, the algebraic statement $\mathcal{J}^{\Lambda}M = \mathbf{0}$ is equivalent to the measurement that $\varepsilon_i^{\vee}(M) \leq \lambda_i$ for all $i \in I$, see [Lau09, Proposition 2.8]. Likewise $\mathcal{J}^{\Lambda}M = M$ if and only if there exists some $i \in I$ such that $\varepsilon_i^{\vee}(M) > \lambda_i$. Hence, given a finite dimensional $R(\nu)$-module M, there exists a $\Lambda \in P^+$ such that $\mathcal{J}^{\Lambda}M = \mathbf{0}$, so that we can identify M with the $R^{\Lambda}(\nu)$-module $\mathrm{pr}_{\Lambda} M$. For instance, take any $\Lambda = \sum_{i\in I} m_i\lambda_i$ with $m_i > \dim_{\Bbbk} M$. We deduce the following remark.

Remark 3.2 *Let M be a simple $R(\nu)$-module. Then $\mathrm{pr}_{\Lambda} M \neq \mathbf{0}$ iff $\varepsilon_i^{\vee}(M) \leq \lambda_i$ for all $i \in I$.*

Let $\mathcal{M}$ be an irreducible $R^{\Lambda}(\nu)$-module. As in Section 3.2 define $e_i^{\Lambda}\mathcal{M} := \mathrm{pr}_{\Lambda} \circ e_i \circ \mathrm{infl}_{\Lambda}\mathcal{M}$ which is a functor $R^{\Lambda}(\nu)\mathrm{-mod} \to R^{\Lambda}(\nu - i)\mathrm{-mod}$, as well as $\widetilde{e_i}^{\Lambda}\mathcal{M} = \mathrm{pr}_{\Lambda} \circ \widetilde{e_i} \circ \mathrm{infl}_{\Lambda}\mathcal{M}$, $\widetilde{f_i}^{\Lambda}\mathcal{M} = \mathrm{pr}_{\Lambda} \circ \widetilde{f_i} \circ \mathrm{infl}_{\Lambda}\mathcal{M}$, $\varepsilon_i^{\Lambda}(\mathcal{M}) = \varepsilon_i(\mathrm{infl}_{\Lambda}\mathcal{M})$. Let $\mathcal{M} \in R^{\Lambda}(\nu)\mathrm{-mod}$ and $M = \mathrm{infl}_{\Lambda}\mathcal{M}$. Then $\mathrm{pr}_{\Lambda} M = \mathcal{M}$. Since $\mathcal{J}^{\Lambda}M = \mathbf{0}$ then $\mathcal{J}^{\Lambda}e_iM = \mathbf{0}$ too, so that $e_i^{\Lambda}\mathcal{M}$ is an $R(\nu - i)^{\Lambda}$-module with $\mathrm{infl}_{\Lambda}(e_i^{\Lambda}\mathcal{M}) = e_iM$. In particular, $\dim_{\Bbbk} e_i^{\Lambda}\mathcal{M} = \dim_{\Bbbk} e_iM$. If furthermore $\mathcal{M}$ is irreducible, then $\widetilde{e_i}^{\Lambda}\mathcal{M} = \mathrm{soc}\, e_i^{\Lambda}\mathcal{M}$.

3.4 Operators on the Grothendieck group

Let the Grothendieck groups $G_0(R) = \bigoplus_{\nu\in\mathbb{N}[I]} G_0(R(\nu)\mathrm{-fmod}),\ G_0(R^{\Lambda}) = \bigoplus_{\nu\in\mathbb{N}[I]} G_0(R^{\Lambda}(\nu)\mathrm{-fmod}).$ They have the structure of a $\mathbb{Z}[q,q^{-1}]$-module given by shifting the grading, $q[M] = [M\{1\}]$.

The functor e_i defined in (1) is clearly exact so descends to an operator on the Grothendieck group $G_0(R(\nu)\mathrm{-mod}) \longrightarrow G_0(R(\nu - i)\mathrm{-mod})$ and hence $e_i\colon G_0(R) \longrightarrow G_0(R)$. By abuse of notation, we will also call this operator e_i. Likewise $e_i^{\Lambda}\colon G_0(R^{\Lambda}) \longrightarrow G_0(R^{\Lambda})$. We also define divided powers $e_i^{(r)}\colon G_0(R) \longrightarrow G_0(R)$ given by $e_i^{(r)}[M] = \frac{1}{[r]_i^!}[e_i^rM]$, which are well-defined.

For irreducible M, we define $\widetilde{e_i}[M] = [\widetilde{e_i}M]$, $\widetilde{f_i}[M] = [\widetilde{f_i}M]$, and extend the action linearly.

The quantum Serre relations (2) are certain (minimal) relations that hold among the operators e_i on $G_0(R)$. The operator

$$\sum_{r=0}^{a+1}(-1)^r e_i^{(a+1-r)}e_je_i^{(r)} = 0. \tag{2}$$

In Section 6.2 below, we give an alternate proof to that of Khovanov-Lauda that the quantum Serre relation (2) holds by examining the structure of all simple $R((a+1)i+j)$-modules. We further construct simple $R(ci+j)$-modules that are witness to the nonvanishing of the analogous relation taking $c \leq a$.

4 Reminders on crystals

A main result of this paper is the realization of a crystal graph structure on $G_0(R)$ which we identify as the crystal $B(\infty)$. We assume the reader is familiar with the language and notation of crystals.

Example 4.1 **(T_Λ $(\Lambda \in P)$)**
Let $T_\Lambda = \{t_\Lambda\}$ *with* $\mathrm{wt}(t_\Lambda) = \Lambda$, $\varepsilon_i(t_\Lambda) = \varphi_i(t_\Lambda) = -\infty$, $\widetilde{e}_i t_\Lambda = \widetilde{f}_i t_\Lambda = 0$. *Tensoring a crystal* B *with the crystal* T_Λ *has the effect of shifting the weight* wt *by* Λ *and leaving the other data fixed.*

Example 4.2 **(B_i $(i \in I)$)** $B_i = \{b_i(n); n \in \mathbb{Z}\}$ *with* $\mathrm{wt}(b_i(n)) = n\alpha_i$, $\varepsilon_i(b_i(n)) = -n = -\varphi_i(b_i(n))$, $\varepsilon_j(b_i(n)) = -\infty = \varphi_j(b_i(n))$ *if* $j \neq i$; $\widetilde{e}_i b_i(n-1) = b_i(n) = \widetilde{f}_i b_i(n+1)$, $\widetilde{e}_j b_i(n) = \widetilde{f}_j b_i(n) = 0$ *if* $j \neq i$. *We write* b_i *for* $b_i(0)$.

4.1 Description of $B(\infty)$

$B(\infty)$ is the crystal associated with the crystal graph of $\mathbf{U}_q^-(\mathfrak{g})$ where $\mathfrak{g}$ is the Kac-Moody algebra defined from the Cartan data above. One can also define $B(\infty)$ as an abstract crystal. As such, it can be characterized by Kashiwara-Saito's Proposition 4.3 below.

Proposition 4.3 ([KS97] Proposition 3.2.3) *Let* B *be a crystal and* b_0 *an element of* B *with weight zero. Assume the following conditions.*

(B1) $\mathrm{wt}(B) \subset Q_-$.

(B2) b_0 *is the unique element of* B *with weight zero.*

(B3) $\varepsilon_i(b_0) = 0$ *for every* $i \in I$.

(B4) $\varepsilon_i(b) \in \mathbb{Z}$ *for any* $b \in B$ *and* $i \in I$.

(B5) For every $i \in I$, *there exists a strict embedding* $\Psi_i \colon B \to B \otimes B_i$.

(B6) $\Psi_i(B) \subset B \times \{\widetilde{f}_i^{\,n} b_i; n \geq 0\}$.

(B7) For any $b \in B$ *such that* $b \neq b_0$, *there exists* i *such that* $\Psi_i(b) = b' \otimes \widetilde{f}_i^{\,n} b_i$ *with* $n > 0$.

Then B *is isomorphic to* $B(\infty)$.

5 Module theoretic realizations of certain crystals

Let $\mathcal{B}$ denote the set of isomorphism classes of irreducible R-modules. Let $\mathbf{0}$ denote the zero module.

Let M be an irreducible $R(\nu)$-module, so that $[M] \in \mathcal{B}$. By abuse of notation, we identify M with $[M]$ in the following definitions. Hence, we are defining operators and functions on $\mathcal{B} \sqcup \{0\}$ below.

Recall from Section 3.2 the definitions of $\widetilde{e}_i, \widetilde{f}_i, \varepsilon_i$. For $\nu = \sum_{i \in I} \nu_i \alpha_i$, $i \in I$ and $M \in R(\nu)-\mathrm{mod}$ set

$$\mathrm{wt}(M) = -\nu, \qquad \mathrm{wt}_i(M) = \langle h_i, \mathrm{wt}(M) \rangle, \qquad \varphi_i(M) = \varepsilon_i(M) + \langle h_i, \mathrm{wt}(M) \rangle.$$

Proposition 5.1 *The tuple* $(\mathcal{B}, \varepsilon_i, \varphi_i, \widetilde{e}_i, \widetilde{f}_i, \mathrm{wt})$ *defines a crystal.*

We write $\mathbb{1} \in \mathcal{B}$ for the class of the trivial $R(\nu)$-module where $\nu = \emptyset$ and $|\nu| = 0$.

One of the main theorems of this paper is Theorem 7.4 that identifies the crystal $\mathcal{B}$ as $B(\infty)$. However we need the many auxiliary results that follow before we can prove this.

Let $\mathcal{B}^\Lambda$ denote the set of isomorphism classes of irreducible R^Λ-modules. As in the previous section, by abuse of notation we write $\mathcal{M}$ for $[\mathcal{M}]$ below. Recall from Section 3.3 the definitions of $\widetilde{e}_i^{\Lambda}$, $\widetilde{f}_i^{\Lambda}$, ε_i^Λ. Let $\mathrm{wt}^\Lambda(\mathcal{M}) = -\nu+\Lambda$ when $\mathcal{M}$ is an $R^\Lambda(\nu)$-module and $\varphi_i^\Lambda(\mathcal{M}) = \max\{k \in \mathbb{Z} \mid \mathrm{pr}_\Lambda \circ \widetilde{f}_i^{k} \circ \mathrm{infl}_\Lambda \mathcal{M} \neq \mathbf{0}\}$. Note $\varepsilon_i^\Lambda(\mathcal{M}) = \max\{k \in \mathbb{Z} \mid (\widetilde{e}_i^{\Lambda})^k \mathcal{M} \neq \mathbf{0}\}$, and $0 \leq \varphi_i^\Lambda(\mathcal{M}) < \infty$.

It is true, but not at all obvious, that with this definition $\varphi_i^\Lambda(\mathcal{M}) = \varepsilon_i^\Lambda(\mathcal{M}) + \langle h_i, \mathrm{wt}^\Lambda \mathcal{M}\rangle$; see Corollary 6.18. The proof that the data $(\mathcal{B}^\Lambda, \varepsilon_i^\Lambda, \varphi_i^\Lambda, \widetilde{e}_i^{\Lambda}, \widetilde{f}_i^{\Lambda}, \mathrm{wt}^\Lambda)$ defines a crystal is delayed until Section 7.

6 Understanding $R(\nu)$-modules and the crystal data of $\mathcal{B}$

This section contains a summary of how the quantities $\varepsilon_j^\vee$, ε_i, φ_i^Λ change with the application of $\widetilde{f}_j$.

Throughout this section we assume $j \neq i$ and set $a = a_{ij} = -\langle h_i, \alpha_j\rangle$.

6.1 Jump

Given an irreducible module M, $\mathrm{pr}_\Lambda \widetilde{f}_i M$ is either irreducible or zero. In the following subsection, we measure exactly when the latter occurs. More specifically, we compare $\varepsilon_i^\vee(M)$ to $\varepsilon_i^\vee(\widetilde{f}_i M)$ and compute when the latter "jumps" by $+1$. In this case, we show $\widetilde{f}_i M \cong \widetilde{f}_i^{\vee} M$. Understanding exactly when this jump occurs is a key ingredient in constructing the strict embedding of crystals in Section 7.1.

Proposition 6.1 *Let M be an irreducible $R(\nu)$-module.*

i) For any $i \in I$, either $\varepsilon_i^\vee(\widetilde{f}_i M) = \varepsilon_i^\vee(M)$ or $\varepsilon_i^\vee(M) + 1$.

ii) For any $i, j \in I$ with $i \neq j$, we have $\varepsilon_i^\vee(\widetilde{f}_j M) = \varepsilon_i^\vee(M)$ and $\varepsilon_i(\widetilde{f}_j^{\vee} M) = \varepsilon_i(M)$.

Definition 6.2 *Let M be an irreducible $R(\nu)$-module and let $\Lambda \in P^+$. Define $\varphi_i^\Lambda(M) = \max\{k \in \mathbb{Z} \mid \mathrm{pr}_\Lambda \widetilde{f}_i^{k} M \neq \mathbf{0}\}$, where we take the convention that $\widetilde{f}_i^{k} = \widetilde{e}_i^{-k}$ when $k < 0$.*

Definition 6.3 *Let M be a simple $R(\nu)$-module and let $i \in I$. Then* $\mathrm{jump}_i(M) := \max\{J \geq 0 \mid \varepsilon_i^\vee(M) = \varepsilon_i^\vee(\widetilde{f}_i^{J} M)\}$.

Lemma 6.4 (Jump Lemma) *Let M be irreducible. The following are equivalent:*

1) $\mathrm{jump}_i(M) = 0$

2) $\widetilde{f}_i M \cong \widetilde{f}_i^{\vee} M$

3) $\widetilde{f}_i^{m} M \cong (\widetilde{f}_i^{\vee})^m M$ *for all* $m \geq 1$

4) $\mathrm{Ind}\, M \boxtimes L(i) \cong \mathrm{Ind}\, L(i) \boxtimes M$

5) $\mathrm{Ind}\, M \boxtimes L(i^m) \cong \mathrm{Ind}\, L(i^m) \boxtimes M$ *for all* $m \geq 1$

6) $\widetilde{f}_i M \cong \mathrm{Ind}\, M \boxtimes L(i)$

6') $\widetilde{f}_i^{\vee} M \cong \mathrm{Ind}\, L(i) \boxtimes M$

7) $\mathrm{Ind}\, M \boxtimes L(i)$ *is irreducible*

7') $\mathrm{Ind}\, L(i) \boxtimes M$ *is irreducible*

8) $\mathrm{Ind}\, M \boxtimes L(i^m)$ *is irreducible for all* $m \geq 1$

8') $\mathrm{Ind}\, L(i^m) \boxtimes M$ *is irreducible for all* $m \geq 1$

9) $\varepsilon_i^\vee(\widetilde{f}_i M) = \varepsilon_i^\vee(M) + 1$

9') $\varepsilon_i(\widetilde{f}_i^{\vee} M) = \varepsilon_i(M) + 1$

10) $\mathrm{jump}_i(\widetilde{f}_i^{m} M) = 0$ *for all* $m \geq 0$

11) $\varepsilon_i^\vee(\widetilde{f}_i^{m} M) = \varepsilon_i^\vee(M) + m$ *for all* $m \geq 1$

Proposition 6.5 *Let M be a simple $R(\nu)$-module and let $i \in I$. Then the following hold.*

i) $\mathrm{jump}_i(M) = \max\{J \geq 0 \mid \varepsilon_i(M) = \varepsilon_i((\widetilde{f_i}^\vee)^J M)\}$

ii) $\mathrm{jump}_i(M) = \min\{J \geq 0 \mid \widetilde{f}_i(\widetilde{f_i}^J M) \cong \widetilde{f_i}^\vee(\widetilde{f_i}^J M)\}$

iii) $\mathrm{jump}_i(M) = \min\{J \geq 0 \mid \widetilde{f}_i((\widetilde{f_i}^\vee)^J M) \cong \widetilde{f_i}^\vee((\widetilde{f_i}^\vee)^J M)\}$

iv) If $\varphi_i^\Lambda(M) > -\infty$*, then* $\mathrm{jump}_i(M) = \varphi_i^\Lambda(M) + \varepsilon_i^\vee(M) - \lambda_i$.

v) $\mathrm{jump}_i(M) = \varepsilon_i(M) + \varepsilon_i^\vee(M) + \mathrm{wt}_i(M)$.

Remark 6.6 *Given $\Lambda, \Omega \in P^+$ and irreducible modules A and B with* $\mathrm{pr}_\Lambda A \neq \mathbf{0}$, $\mathrm{pr}_\Omega A \neq \mathbf{0}$, $\mathrm{pr}_\Lambda B \neq \mathbf{0}$, $\mathrm{pr}_\Omega B \neq \mathbf{0}$*, then* $\varphi_i^\Lambda(A) - \varphi_i^\Lambda(B) = \varphi_i^\Omega(A) - \varphi_i^\Omega(B)$ *since by Proposition 6.5* (iv) *we compute*

$$\begin{aligned}\varphi_i^\Lambda(A) - \varphi_i^\Lambda(B) &= (\mathrm{jump}_i(A) - \varepsilon_i^\vee(A) + \lambda_i) - (\mathrm{jump}_i(B) - \varepsilon_i^\vee(B) + \lambda_i)\\ &= \mathrm{jump}_i(A) - \mathrm{jump}_i(B) + \varepsilon_i^\vee(B) - \varepsilon_i^\vee(A) = \varphi_i^\Omega(A) - \varphi_i^\Omega(B).\end{aligned}$$

6.2 The Structure Theorems for simple $R(ci + j)$-modules

In this section we describe the structure of all simple $R(ci+j)$-modules. We will henceforth refer to Theorems 6.7, 6.8 as the Structure Theorems for simple $R(ci+j)$-modules. Throughout this section we assume $j \neq i$ and set $a = a_{ij} = -\langle h_i, \alpha_j \rangle$.

In the theorems below we introduce the notation $\mathcal{L}(i^{c-n}ji^n)$ and $\mathcal{L}(n) \stackrel{\text{def}}{=} \mathcal{L}(i^{a-n}ji^n)$ for the simple $R(ci+j)$-modules (up to grading shift) when $c \leq a$. They are characterized by $\varepsilon_i\left(\mathcal{L}(i^{c-n}ji^n)\right) = n$.

Theorem 6.7 *Let $c \leq a$ and let $\nu = ci + j$. Up to isomorphism and grading shift, there exists a unique irreducible $R(\nu)$-module denoted $\mathcal{L}(i^{c-n}ji^n)$ with $\varepsilon_i\left(\mathcal{L}(i^{c-n}ji^n)\right) = n$ for each n with $0 \leq n \leq c$. Furthermore, $\varepsilon_i^\vee\left(\mathcal{L}(i^{c-n}ji^n)\right) = c - n$ and*

$$\mathrm{ch}(\mathcal{L}(i^{c-n}ji^n)) = [c-n]_i![n]_i! i^{c-n}ji^n.$$

In particular, in the Grothendieck group $e_i^{(c-s)}e_je_i^{(s)}[\mathcal{L}(i^{c-n}ji^n)] = 0$ unless $s = n$.

In the previous theorem we introduced the notation $\mathcal{L}(i^{c-n}ji^n)$ for the unique (up to isomorphism and grading shift) simple $R(ci+j)$-module with $\varepsilon_i = n$ when $c \leq a$. Theorem 6.8 below extends this uniqueness to $c \geq a$. In the special case that $c = a$, we denote $\mathcal{L}(n) = \mathcal{L}(i^{a-n}ji^n)$. The following theorem motivates why we distinguish the special case $c = a$.

Theorem 6.8 *Let $0 \leq n \leq a$.*

i) The module

$$\mathrm{Ind}\, L(i^m) \boxtimes \mathcal{L}(n) \cong \mathrm{Ind}\, \mathcal{L}(n) \boxtimes L(i^m)$$

is irreducible for all $m \geq 0$.

ii) Let $c \geq a$. Let N be an irreducible $R(ci+j)$-module with $\varepsilon_i(N) = n$. Then $c - a \leq n \leq c$, and up to grading shift

$$N \cong \mathrm{Ind}\, \mathcal{L}(n - (c-a)) \boxtimes L(i^{c-a}).$$

The proofs of the Structure Theorems are given by careful calculations on special weight bases of the above modules. In the interest of space, they are ommitted here.

6.3 Understanding φ_i^Λ

The following theorems measure how the crystal data differs for M and $\widetilde{f}_j M$.

Theorem 6.9 *Let M be a simple $R(\nu)$-module and let $\Lambda \in P^+$ such that $\mathrm{pr}_\Lambda M \neq \mathbf{0}$ and $\mathrm{pr}_\Lambda \widetilde{f}_j M \neq \mathbf{0}$. Let $m = \varepsilon_i(M)$, $k = \varphi_i^\Lambda(M)$. Then there exists an n with $0 \leq n \leq a$ such that $\varepsilon_i(\widetilde{f}_j M) = m - (a-n)$ and $\varphi_i^\Lambda(\widetilde{f}_j M) = k + n$.*

Proof: This follows from Theorem 6.16 which proves the theorem in the case $\nu = ci + dj$ and from Proposition 6.17 which reduces it to this case. $\square$

One important rephrasing of the Theorem is

$$\varphi_i^\Lambda(\widetilde{f}_j M) - \varepsilon_i(\widetilde{f}_j M) = a + (\varphi_i^\Lambda(M) - \varepsilon_i(M)) = -\langle h_i, \alpha_j \rangle + (\varphi_i^\Lambda(M) - \varepsilon_i(M)).$$

First we introduce several lemmas that will be needed.

Lemma 6.10 *Suppose $c + d \leq a$.*

i) $\mathrm{Ind}\, \mathcal{L}(i^c j i^d) \boxtimes L(i^m)$ *has irreducible cosocle equal to*

$$\widetilde{f}_i^{\,m} \mathcal{L}(i^c j i^d) = \widetilde{f}_i^{\,m+d} \mathcal{L}(i^c j) = \begin{cases} \mathrm{Ind}\, \mathcal{L}(a-c) \boxtimes L(i^{m-a+c+d}) & m \geq a - (c+d) \\ \mathcal{L}(i^c j i^{d+m}) & m < a - (c+d). \end{cases}$$

ii) Suppose there is a nonzero map $\mathrm{Ind}\, \mathcal{L}(c_1) \boxtimes \mathcal{L}(c_2) \boxtimes \cdots \boxtimes \mathcal{L}(c_r) \boxtimes L(i^m) \longrightarrow Q$ where Q is irreducible. Then $\varepsilon_i(Q) = m + \sum_{t=1}^r c_t$ and $\varepsilon_i^\vee(Q) = m + \sum_{t=1}^r (a - c_t)$.

iii) Let B and Q be simple with a nonzero map $IndB \boxtimes \mathcal{L}(c) \to Q$. Then $\varepsilon_i(Q) = \varepsilon_i(B) + c$.

Lemma 6.11 *Let N be an irreducible $R(ci + dj)$-module with $\varepsilon_i(N) = 0$. Suppose $c + d > 0$.*

i) There exists irreducible $\overline{N}$ with $\varepsilon_i(\overline{N}) = 0$ and a surjection $\mathrm{Ind}\, \overline{N} \boxtimes \mathcal{L}(i^b j) \twoheadrightarrow N$ with $b \leq a$.

ii) There exists an $r \in \mathbb{N}$ and $b_t \leq a$ for $1 \leq t \leq r$ such that $\mathrm{Ind}\, \mathcal{L}(i^{b_1} j) \boxtimes \mathcal{L}(i^{b_2} j) \boxtimes \cdots \boxtimes \mathcal{L}(i^{b_r} j) \twoheadrightarrow N$.

Lemma 6.12 *Suppose Q is irreducible and we have a surjection*

$$\mathrm{Ind}\, \mathcal{L}(i^{b_1} j) \boxtimes \mathcal{L}(i^{b_2} j) \boxtimes \cdots \boxtimes \mathcal{L}(i^{b_r} j) \boxtimes L(i^h) \twoheadrightarrow Q.$$

i) Then for $h \gg 0$ we have a surjection $\mathrm{Ind}\, \mathcal{L}(a - b_1) \boxtimes \mathcal{L}(a - b_2) \boxtimes \cdots \boxtimes \mathcal{L}(a - b_r) \boxtimes L(i^g) \twoheadrightarrow Q$ where $g = h - \sum_{t=1}^r (a - b_t)$.

ii) In the case $h < ar - \sum_{t=1}^r b_t$, we have

$$\mathrm{Ind}\, \mathcal{L}(i^{b_1} j) \boxtimes \cdots \boxtimes \mathcal{L}(i^{b_{s-1}} j) \boxtimes \mathcal{L}(i^{b_s} j i^{g'}) \boxtimes \mathcal{L}(a - b_{s+1}) \boxtimes \cdots \boxtimes \mathcal{L}(a - b_r) \twoheadrightarrow Q$$

where $g' = h - \sum_{t=s+1}^r (a - b_t)$ and s is such that $\sum_{t=s+1}^r (a - b_t) \leq h < \sum_{t=s}^r (a - b_t)$.

Lemma 6.13 *Let M be an irreducible $R(\nu)$-module and suppose $\mathrm{Ind}\, A \boxtimes B \boxtimes L(i^h) \twoheadrightarrow M$ is a nonzero map where $\varepsilon_i(A) = 0$ and B is irreducible. Then there exists a surjective map $\mathrm{Ind}\, A \boxtimes \widetilde{f}_i^{\,h} B \twoheadrightarrow M$.*

The following two lemmas discuss ways of detecting when pr_Λ of an induced module is zero.

Lemma 6.14 *Let A be an irreducible $R(\nu)$-module with $\mathrm{pr}_\Lambda A \neq \mathbf{0}$ and $k = \varphi_i^\Lambda(A)$.*

i) Let U be an irreducible $R(\mu)$-module and let $t \geq 1$. Then $\mathrm{pr}_\Lambda \mathrm{Ind}\, A \boxtimes L(i^{k+t}) \boxtimes U = \mathbf{0}$.

ii) Let B be irreducible with $\varepsilon_i^\vee(B) > k$. Then $\mathrm{pr}_\Lambda \mathrm{Ind}\, A \boxtimes B = \mathbf{0}$. In particular, if Q is any irreducible quotient of $\mathrm{Ind}\, A \boxtimes B$, then $\mathrm{pr}_\Lambda Q = \mathbf{0}$.

Lemma 6.15 *Let A be an irreducible $R(\nu)$-module with $\mathrm{pr}_\Lambda A \neq \mathbf{0}$ and $k = \varphi_i^\Lambda(A)$. Further suppose $\varepsilon_i(A) = \varepsilon_j(A) = 0$ and that B is an irreducible $R(ci+dj)$-module with $\varepsilon_i^\vee(B) \leq k$. Let Q be irreducible such that $\mathrm{Ind}\, A \boxtimes B \twoheadrightarrow Q$ is nonzero. Then $\varepsilon_i^\vee(Q) \leq \lambda_i$. Further, if $\varepsilon_j^\vee(B) \leq \varphi_j^\Lambda(A)$ (or if $\lambda_j \gg 0$) then $\mathrm{pr}_\Lambda Q \neq \mathbf{0}$.*

Theorem 6.16 *Let M be an irreducible $R(ci + dj)$-module and let $\Lambda \in P^+$ be such that $\mathrm{pr}_\Lambda M \neq \mathbf{0}$ and $\mathrm{pr}_\Lambda \widetilde{f}_j M \neq \mathbf{0}$. Let $m = \varepsilon_i(M)$, $k = \varphi_i^\Lambda(M)$. Then there exists an n with $0 \leq n \leq a$ such that $\varepsilon_i(\widetilde{f}_j M) = m - (a - n)$ and $\varphi_i^\Lambda(\widetilde{f}_j M) = k + n$.*

We have just shown in Theorem 6.16 that Theorem 6.9 holds for all $R(ci + dj)$-modules. Next we show that to deduce the theorem for $R(\nu)$-modules for arbitrary ν it suffices to know the result for $\nu = ci + dj$.

Proposition 6.17 *Let $\Lambda \in P^+$ and let M be an irreducible $R(\nu)$-module such that $\mathrm{pr}_\Lambda M \neq \mathbf{0}$ and $\mathrm{pr}_\Lambda \widetilde{f}_j M \neq \mathbf{0}$. Suppose $\varepsilon_i(M) = m$ and $\varepsilon_i(\widetilde{f}_j M) = m - (a - n)$ for some $0 \leq n \leq a$. Then there exists c, d and an irreducible $R(ci + dj)$-module B such that $\varepsilon_i(B) = m$, $\varepsilon_i(\widetilde{f}_j B) = m - (a - n)$ and there exists $\Omega \in P^+$ with $\mathrm{pr}_\Omega(B) \neq \mathbf{0}$, $\mathrm{pr}_\Omega(\widetilde{f}_j B) \neq \mathbf{0}$, $\mathrm{pr}_\Omega(M) \neq \mathbf{0}$, $\mathrm{pr}_\Omega(\widetilde{f}_j M) \neq \mathbf{0}$, and furthermore*

$$\varphi_i^\Omega(\widetilde{f}_j M) - \varphi_i^\Omega(M) = \varphi_i^\Omega(\widetilde{f}_j B) - \varphi_i^\Omega(B).$$

Note that by Remark 6.6 $\varphi_i^\Lambda(\widetilde{f}_j M) - \varphi_i^\Lambda(M) = \varphi_i^\Omega(\widetilde{f}_j M) - \varphi_i^\Omega(M)$, so once we prove this proposition, it together with Theorem 6.16 proves Theorem 6.9.

Corollary 6.18 (Corollary of Theorem 6.9) *Let $\Lambda = \sum_{i \in I} \lambda_i \Lambda_i \in P^+$ and let M an irreducible $R(\nu)$-module such that $\mathrm{pr}_\Lambda M \neq \mathbf{0}$. Then*

$$\varphi_i^\Lambda(M) = \lambda_i + \varepsilon_i(M) + \mathrm{wt}_i(M).$$

7 Main Results

Now that we have built up the machinery of Section 6, we can prove the module theoretic crystal $\mathcal{B}$ is isomorphic to $B(\infty)$. Once we have completed this step, it is not much harder to show $\mathcal{B}^\Lambda \cong B(\Lambda)$.

7.1 Constructing the strict embedding Ψ

Proposition 7.1 *Let M be a simple $R(\nu)$-module, and write $c = \varepsilon_i^\vee(M)$.*

i) Suppose $\varepsilon_i^\vee(\widetilde{f}_i M) = \varepsilon_i^\vee(M) + 1$. Then $\widetilde{e}_i^{\vee} \widetilde{f}_i M \cong M$ up to grading shift.

ii) Suppose $\varepsilon_i^\vee(\widetilde{f}_j M) = \varepsilon_i^\vee(M)$ where i and j are not necessarily distinct. Then $(\widetilde{e}_i^{\vee})^c(\widetilde{f}_j M) \cong \widetilde{f}_j(\widetilde{e}_i^{\vee c} M)$ up to grading shift.

Proposition 7.2 *Let M be an irreducible $R(\nu)$-module, and write $c = \varepsilon_i^\vee(M)$, $\overline{M} = (\widetilde{e}_i^{\,\vee})^c(M)$.*

i) $\varepsilon_i(M) = \max\{\varepsilon_i(\overline{M}), c - \mathrm{wt}_i(\overline{M})\}$.

ii) Suppose $\varepsilon_i(M) > 0$. Then

$$\varepsilon_i^\vee(\widetilde{e}_i M) = \begin{cases} c & \text{if } \varepsilon_i(\overline{M}) \geq c - \mathrm{wt}_i(\overline{M}), \\ c-1 & \text{if } \varepsilon_i(\overline{M}) < c - \mathrm{wt}_i(\overline{M}). \end{cases}$$

iii) Suppose $\varepsilon_i(M) > 0$. Then

$$(\widetilde{e}_i^{\,\vee})^{\varepsilon_i^\vee(\widetilde{e}_i M)}(\widetilde{e}_i M) = \begin{cases} \widetilde{e}_i(\overline{M}) & \text{if } \varepsilon_i(\overline{M}) \geq c - \mathrm{wt}_i(\overline{M}), \\ \overline{M} & \text{if } \varepsilon_i(\overline{M}) < c - \mathrm{wt}_i(\overline{M}). \end{cases}$$

Proposition 7.3 *For each $i \in I$ define a map $\Psi_i \colon \mathcal{B} \to \mathcal{B} \otimes B_i$ by*

$$M \mapsto (\widetilde{e}_i^{\,\vee})^c(M) \otimes b_i(-c),$$

where $c = \varepsilon_i^\vee(M)$. Then Ψ_i is a strict embedding of crystals.

Theorem 7.4 (First Main Theorem) *The crystal $\mathcal{B}$ is isomorphic to $B(\infty)$.*

Now we will show the data $(\mathcal{B}^\Lambda, \varepsilon_i^\Lambda, \varphi_i^\Lambda, \widetilde{e}_i^{\,\Lambda}, \widetilde{e}_i^{\,\Lambda}, \mathrm{wt}^\Lambda)$ of Section 5 defines a crystal graph and identify it as the highest weight crystal $B(\Lambda)$.

Theorem 7.5 (Second Main Theorem) *$\mathcal{B}^\Lambda$ is a crystal; the crystal $\mathcal{B}^\Lambda$ is isomorphic to $B(\Lambda)$.*

7.2 $\mathbf{U}_q^+$-module structures

Set

$$G_0^*(R) = \bigoplus_\nu G_0(R(\nu))^* \qquad G_0^*(R^\Lambda) = \bigoplus_\nu G_0(R^\Lambda(\nu))^*$$

where, by V^* we mean the linear dual $\mathrm{Hom}_\mathcal{A}(V, \mathcal{A})$. Because $G_0(R)$ and $G_0(R^\Lambda)$ are ${}_\mathcal{A}\mathbf{U}_q^+$-modules, we can endow $G_0^*(R)$, $G_0^*(R^\Lambda)$ with a left ${}_\mathcal{A}\mathbf{U}_q^+$-module structure. $G_0(R(\nu))^*$ has basis given by $\{\delta_M \mid M \in \mathcal{B}, \mathrm{wt}(M) = -\nu\}$ defined by

$$\delta_M([N]) = \begin{cases} q^{-r} & M \cong N\{r\} \\ 0 & \text{otherwise,} \end{cases}$$

where N ranges over simple $R(\nu)$-modules. We set $\mathrm{wt}(\delta_M) = -\mathrm{wt}(M)$. Likewise $G_0(R^\Lambda(\nu))^*$ has basis $\{\mathfrak{d}_\mathcal{M} \mid \mathcal{M} \in \mathcal{B}^\Lambda, \mathrm{wt}(\mathcal{M}) = -\nu + \Lambda\}$ defined similarly. Note that if δ_M has degree d then $\delta_{M\{1\}} = q^{-1}\delta_M$ has degree $d-1$. Recall $\mathbb{1} \in \mathcal{B}$ denotes the trivial $R(0)$-module and we will also write $\mathbb{1} \in \mathcal{B}^\Lambda$ for the trivial $R^\Lambda(0)$-module.

Lemma 7.6 *The maps $F \colon {}_\mathcal{A}\mathbf{U}_q^+ \to G_0^*(R)$, $\mathcal{F} \colon {}_\mathcal{A}\mathbf{U}_q^+ \to G_0^*(R^\Lambda)$ defined by $F(y) = y \cdot \delta_\mathbb{1}$, $\mathcal{F}(y) = y \cdot \mathfrak{d}_\mathbb{1}$ respectively are surjective ${}_\mathcal{A}\mathbf{U}_q^+$-module homomorphisms, and $\ker \mathcal{F} \ni e_i^{(\lambda_i+1)}$ for all $i \in I$.*

Recall that as a left ${}_{\mathcal{A}}\mathbf{U}_q^+$-module ${}_{\mathcal{A}}V^*(\Lambda) \cong {}_{\mathcal{A}}\mathbf{U}_q^+ / \sum_{i\in I} {}_{\mathcal{A}}\mathbf{U}_q^+ \cdot e_i^{(\lambda_i+1)}$. So the map $\mathcal{F}$ factors through ${}_{\mathcal{A}}V^*(\Lambda)$, and the resulting induced map to $G_0^*(R^\Lambda)$ must be injective as the ranks of their weight spaces coincide by Theorem 7.5.

Theorem 7.7 *As ${}_{\mathcal{A}}\mathbf{U}_q^+$ modules*

$$1.\ {}_{\mathcal{A}}\mathbf{U}_q^+ \cong G_0^*(R), \qquad 2.\ {}_{\mathcal{A}}V^*(\Lambda) \cong G_0^*(R^\Lambda), \qquad 3.\ {}_{\mathcal{A}}V(\Lambda) \cong G_0(R^\Lambda).$$

References

[Ari02] S. Ariki. *Representations of quantum algebras and combinatorics of Young tableaux*, volume 26 of *University Lecture Series*. AMS, Providence, RI, 2002.

[AK94] S. Ariki and K. Koike. A Hecke algebra of $(\mathbf{Z}/r\mathbf{Z}) \wr \mathcal{S}_n$ and construction of its irreducible representations. *Adv. Math.*, 106(2):216–243, 1994.

[BK09] J. Brundan and A. Kleshchev. Graded decomposition numbers for cyclotomic Hecke algebras. *Adv. Math.*, 222(6):1883–1942, 2009.

[CR08] J. Chuang and R. Rouquier. Derived equivalences for symmetric groups and sl_2-categorification. *Ann. of Math.*, 167:245–298, 2008.

[KS97] M. Kashiwara and Y. Saito. Geometric construction of crystal bases. *Duke Math. J.*, 89(1):9–36, 1997.

[KL09] M. Khovanov and A. Lauda. A diagrammatic approach to categorification of quantum groups I. *Represent. Theory*, 13:309–347, 2009.

[KL08a] M. Khovanov and A. Lauda. A diagrammatic approach to categorification of quantum groups II. To appear in *Transactions of the AMS*, math.QA/0804.2080, 2008.

[KL08b] M. Khovanov and A. Lauda. A diagrammatic approach to categorification of quantum groups III. math.QA/0807.3250, 2008.

[Lau08] A. D. Lauda. A categorification of quantum sl(2). math.QA/0803.3652, 2008.

[Lau09] A. D. Lauda. Nilpotency in type A cyclotomic quotients. math.RT/0903.2992, 2009.

[LV09] A. D. Lauda and M. Vazirani. Crystals from categorified quantum groups math.RT/0909.1810, 2009.

[Rou08] R. Rouquier. 2-Kac-Moody algebras. arXiv:0812.5023, 2008.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 750–757

On the diagonal ideal of $(\mathbb{C}^2)^n$ and q,t-Catalan numbers

Kyungyong Lee[1]†and Li Li[2]

[1]*Department of Mathematics, Purdue University, West Lafayette, IN 47907*
[2]*Department of Mathematics, University of Illinois at Urbana Champaign, Urbana, IL 61801*

Abstract. Let I_n be the (big) diagonal ideal of $(\mathbb{C}^2)^n$. Haiman proved that the q,t-Catalan number is the Hilbert series of the graded vector space $M_n = \bigoplus_{d_1,d_2}(M_n)_{d_1,d_2}$ spanned by a minimal set of generators for I_n. We give simple upper bounds on dim $(M_n)_{d_1,d_2}$ in terms of partition numbers, and find all bi-degrees (d_1, d_2) such that dim $(M_n)_{d_1,d_2}$ achieve the upper bounds. For such bi-degrees, we also find explicit bases for $(M_n)_{d_1,d_2}$.

Résumé. Soit I_n l'idéal de la (grande) diagonale de $(\mathbb{C}^2)^n$. Haiman a démontré que le q,t-nombre de Catalan est la série de Hilbert de l'espace vectoriel gradué $M_n = \bigoplus_{d_1,d_2}(M_n)_{d_1,d_2}$ engendré par un ensemble minimal de générateurs de I_n. Nous obtenons des bornes supérieures simples pour dim $(M_n)_{d_1,d_2}$ en termes de nombres de partitions, ainsi que tous les bi-degrés (d_1, d_2) pour lesquels ces bornes supérieures sont atteintes. Pour ces bi–degrés, nous trouvons aussi des bases explicites de $(M_n)_{d_1,d_2}$.

Keywords: q,t-Catalan number, diagonal ideal

1 introduction

1.1 Background

The goal of this paper is to study the q,t-Catalan numbers and the (thick) diagonal ideal in $(\mathbb{C}^2)^n$, and discuss some technique that we have developed recently.

Let n be a positive integer. Consider the set of n-tuples $\{(x_i, y_i)\}_{1\le i\le n}$ in the plane $\mathbb{C}^2$. They form an affine space $(\mathbb{C}^2)^n$ with coordinate ring $\mathbb{C}[\mathbf{x}, \mathbf{y}] = \mathbb{C}[x_1, y_1, ..., x_n, y_n]$. There is a natural symmetric group S_n acting on $\mathbb{C}[\mathbf{x}, \mathbf{y}]$ by permuting the coordinates in $\mathbf{x}, \mathbf{y}$ simultaneously. With this group action, a polynomial $f \in \mathbb{C}[\mathbf{x}, \mathbf{y}]$ is called *alternating* if

$$\sigma(f) = \mathrm{sgn}(\sigma)f \quad \text{for all } \sigma \in S_n.$$

Define $\mathbb{C}[\mathbf{x}, \mathbf{y}]^\epsilon$ to be the vector space of alternating polynomials in $\mathbb{C}[\mathbf{x}, \mathbf{y}]$.

†Research of the first author partially supported by NSF grant DMS 0901367.

There is a more combinatorial way to describe $\mathbb{C}[\mathbf{x},\mathbf{y}]^{\epsilon}$. Denote by $\mathbb{N}$ the set of nonnegative integers. Let $\mathfrak{D}_n$ be the set of n-tuples $D = \{(\alpha_1,\beta_1),...,(\alpha_n,\beta_n)\} \subset \mathbb{N}\times\mathbb{N}$. For $D \in \mathfrak{D}_n$, define

$$\Delta(D) := \det \begin{bmatrix} x_1^{\alpha_1}y_1^{\beta_1} & x_1^{\alpha_2}y_1^{\beta_2} & ... & x_1^{\alpha_n}y_1^{\beta_n} \\ \vdots & \vdots & \ddots & \vdots \\ x_n^{\alpha_1}y_n^{\beta_1} & x_n^{\alpha_2}y_n^{\beta_2} & ... & x_n^{\alpha_n}y_n^{\beta_n} \end{bmatrix}$$

Then $\{\Delta(D)\}_{D\in\mathfrak{D}_n}$ forms a basis for the $\mathbb{C}$-vector space $\mathbb{C}[\mathbf{x},\mathbf{y}]^{\epsilon}$.

It is easy to see that any alternating polynomial vanishes on the thick diagonal of $(\mathbb{C}^2)^n$. (By thick diagonal we mean the set of n-tuples of points in $\mathbb{C}^2$ where at least two points coincide.) A theorem of Haiman asserts that the converse is also true: any polynomial that vanishes on the diagonal of $(\mathbb{C}^2)^n$ can be generated by alternating polynomials, i.e.

$$\bigcap_{1\le i<j\le n} (x_i - x_j, y_i - y_j) = \text{ideal generated by } \Delta(D)\text{'s}.$$

We call the above ideal the **diagonal ideal** and denote it by I_n. the number of minimal generators of I_n, which is the same as the dimension of the vector space $M_n = I_n/(\mathbf{x},\mathbf{y})I_n$, is equal to the n-th Catalan number. The space M_n is doubly graded as $\oplus_{d_1,d_2}(M_n)_{d_1,d_2}$. The q,t-Catalan number, originally introduced by A.M.Garsia and M.Haiman in [4], can be defined as

$$C_n(q,t) = \sum_{d_1,d_2} t^{d_1}q^{d_2}\dim(M_n)_{d_1,d_2}.$$

The q,t-Catalan number $C_n(q,t)$ also has a combinatorial interpretation using Dyck paths. To be more precise, take the $n\times n$ square whose southwest corner is $(0,0)$ and northeast corner is (n,n). Let $\mathcal{D}_n$ be the collection of Dyck paths, i.e. lattice paths from $(0,0)$ to (n,n) that proceed by NORTH or EAST steps and never go below the diagonal. For any Dyck path Π, let $a_i(\Pi)$ be the number of squares in the i-th row that lie in the region bounded by Π and the diagonal. A.M.Garsia and J.Haglund ([2], [3]) among others showed that

$$C_n(q,t) = \sum_{\Pi\in\mathcal{D}_n} q^{\text{area}(\Pi)}t^{\text{dinv}(\Pi)},$$

where

$$\text{dinv}(\Pi) := |\{(i,j)\,|\,i<j \text{ and } a_i(\Pi) = a_j(\Pi)\}| + |\{(i,j)\,|\,i<j \text{ and } a_i(\Pi)+1 = a_j(\Pi)\}|.$$

Haiman posed a question asking for a rule that associate to each Dyck path Π an element $D(\Pi)\in\mathfrak{D}_n$ such that $\deg_{\mathbf{x}}\Delta(D(\Pi)) = \text{area}(\Pi)$, $\deg_{\mathbf{y}}\Delta(D(\Pi)) = \text{dinv}(\Pi)$, and that the set $\{\Delta(D(\Pi))\}$ generates I_n. The last condition is equivalent to requiring the images $\{\overline{\Delta(D(\Pi))}\}$ form a basis of M_n). It is natural to ask the following more general question:

Question 1.1 *Given a bi-degree (d_1,d_2), is there a combinatorially significant construction of the basis for each $(M_n)_{d_1,d_2}$?*

1.2 Main results

This paper initiates the approach to the study of M_n by comparing it with $M_{n'}$ for a large integer n'. On the one hand, there is a natural map $M_n \to M_{n'}$ for any $n' > n$. On the other hand, for n' sufficiently large, the basis of $(M_{n'})_{d_1',d_2}$ becomes "stable" if we fix d_2 and fix

$$k = \binom{n}{2} - d_1 - d_2 = \binom{n'}{2} - d_1' - d_2.$$

Therefore we can take the "limit" of such basis for $n' \to \infty$. This basis is indexed by the partitions of k. As a consequence, $(M_{n'})_{d_1',d_2}$ can be imbedded as a subspace of the polynomial ring with infinite many variables $\mathbb{C}[\rho_1, \rho_2, \dots]$. The induced map

$$\bar{\varphi} : (M_n)_{d_1,d_2} \to \mathbb{C}[\rho_1, \rho_2, \dots],$$

which will be defined explicitly in subsection 1.2.3, provides a powerful tool to study M_n.

1.2.1 Asymptotic behavior when $k \ll n$

We shall show that if $k \ll n$, then $(M_n)_{d_1,d_2}$ has a basis $\{\overline{\Delta(D)}\}$ where D are so-called **minimal staircase forms** that will be defined later.

The essential step is to observe the following three linear relations that turn the questions into combinatorial games First we introduce some notations.

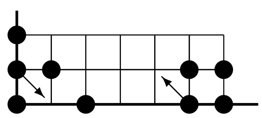
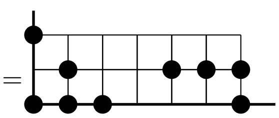
• For $D = \{P_1, \dots, P_n\} \in \mathfrak{D}_n$ where $P_i = (\alpha_i, \beta_i)$, define $|P_i| = \alpha_i + \beta_i$.

Relation 1. Given positive integers $1 \le i \ne j \le n$ such that $|P_i| = i - 1$, $|P_{i+1}| = i$, $|P_j| = j - 1$, $|P_{j+1}| = j$, $\beta_i > 0$, $\alpha_j > 0$ (we assume $|P_{n+1}| = n$). Let D' be obtained from D by moving P_i to southeast and P_j to northwest, i.e.

$$D' = \{P_1, \dots, P_{i-1}, P_i + (1, -1), P_{i+1}, \dots, P_{j-1}, P_j + (-1, 1), P_{j+1}, \dots, P_n\}.$$

Then $\overline{\Delta(D)} = \overline{\Delta(D')}$.

Example: $n = 9, i = 2, j = 6$.

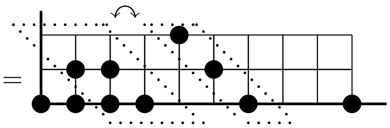
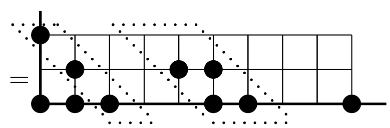
$D =$  $\longrightarrow$ $D' =$ 

Relation 2. Given positive integers h, ℓ and m such that $2 \le h < h+\ell+m \le n+1$, $|P_h| = h-1$, $|P_{h+\ell}| = h + \ell - 1$, $|P_{h+\ell+m}| = h + \ell + m - 1$ (by convention, the last equality holds if $h + \ell + m = n + 1$) and $\alpha_{h+\ell}, ..., \alpha_{h+\ell+m-1} \ge \ell$. Let D' be obtained from D by moving the m points $P_{h+\ell}, \dots, P_{h+\ell+m-1}$ to the left by ℓ units and moving the ℓ points $P_h, \dots, P_{h+\ell-1}$ to the right by m units, i.e.

$$\begin{aligned} D' = \{&P_1, P_2, \dots, P_{h-1}, P_{h+\ell} - (\ell, 0), P_{h+\ell+1} - (\ell, 0), \dots, P_{h+\ell+m-1} - (\ell, 0), \\ &P_h + (m, 0), P_{h+1} + (m, 0), \dots, P_{h+\ell-1} + (m, 0), P_{h+\ell+m}, \dots, P_n\}. \end{aligned}$$

Then $\overline{\Delta(D)} = \overline{\Delta(D')}$.

Example: $n = 10, h = 3, \ell = 4, m = 3$.

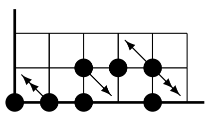
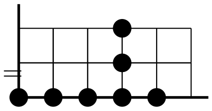
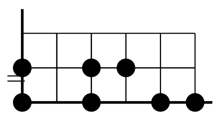
$D =$  $\longrightarrow$ $D' =$ 

Relation 3. Given positive integers j and s. Suppose P_{s_0} is the last point in D satisfying $|P_i| = i - 1$. Define $j = (s_0 - 1 - |P_{s_0}|) + (s_0 - |P_{s_0+1}|) + \cdots + (n - 1 - |P_n|)$. Suppose $|P_i| = i - 1$ for $1 \le i \le j+2$, $P_2 = (1, 0)$, $s_0 \le s \le n$, and $\alpha_s, \beta_s \ge 1$. Let

$$D^{\nwarrow} = \{P_1, \ldots, P_{j+1}, P_{j+2} + (1,-1), P_{j+3}, \ldots, P_{s-1}, P_s + (-1, 1), P_{s+1}, \ldots, P_n\},$$
$$D^{\searrow} = \{P_1, (0,1), P_3, \ldots, P_{s-1}, P_s + (1,-1), P_{s+2}, \ldots, P_n\}.$$

Then $2\overline{\Delta(D)} = \overline{\Delta(D^{\nwarrow})} + \overline{\Delta(D^{\searrow})}$.

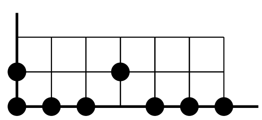
Example: $n = 9, i = 2, j = 6$.

$D =$  $\longrightarrow$ $D^{\nwarrow} =$  , $D^{\searrow} =$ 

We call $D = \{P_1, \ldots, P_n\}$ a *minimal staircase form* if $|P_i| = i - 1$ or $i - 2$ for every $1 \le i \le n$. For a minimal staircase form D, let $\{i_1 < i_2 < \cdots < i_\ell\}$ be the set of i's such that $|P_i| = i - 1$, we define the *partition type* of D to be the partition of $(\binom{n}{2} - \sum |P_i|)$ consisting of all the positive integers in the sequence

$$(i_1 - 1, i_2 - i_1 - 1, i_3 - i_2 - 1, \ldots, i_\ell - i_{\ell-1} - 1, n - i_\ell).$$

Example: Let $n = 8$ and $D = \{P_1, \ldots, P_8\}$ satisfying $(|P_1|, \ldots, |P_8|) = (0, 1, 1, 2, 4, 4, 5, 6)$. Then D is a minimal staircase form. The set $\{i \,\big|\, |P_i| = i - 1\}$ equals $\{1, 2, 5\}$. The positive integers in the sequence $(1 - 1, 2 - 1 - 1, 5 - 2 - 1, 8 - 5)$ are $(2, 3)$, so the partition type of D is $(2, 3)$.



Let $p(k)$ denote the number of partitions of an integer k and Π_k denote the set of partitions of k.

Theorem 1.2 *Let k be any positive integer. There are positive constants $c_1 = 8k + 5, c_2 = 2k + 1$ such that the following holds:*

For integers n, d_1, d_2 satisfying $n \ge c_1$, $d_1 \ge c_2 n$, $d_2 \ge c_2 n$ and $d_1 + d_2 = \binom{n}{2} - k$, the vector space $(M_n)_{d_1,d_2}$ has dimension $p(k)$, and the $p(k)$ elements

$$\big\{\text{a minimal staircase form of bi-degree } (d_1, d_2) \text{ and of partition type } \mu\big\}_{\mu \in \Pi_k}$$

form a basis of $(M_n)_{d_1,d_2}$.

Note that N.Bergeron and Z.Chen have found explicit bases for $(M_n)_{d_1,d_2}$ for certain bi-degrees using a different method [1].

1.2.2 For arbitrary k and n

Denote by $p(k)$ the partition number of k and by convention $p(0) = 1$ and $p(k) = 0$ for $k < 0$. Denote by $p(b, k)$ the partition number of k into no more than b parts, and by convention $p(0, k) = 0$ for $k > 0$, $p(b, 0) = 1$ for $b \geq 0$. One of our main results is as follows.

Theorem 1.3 *Let d_1, d_2 be non-negative integers d_1, d_2 with $d_1 + d_2 \leq \binom{n}{2}$. Define $k = \binom{n}{2} - d_1 - d_2$ and $\delta = \min(d_1, d_2)$. Then the coefficient of $q^{d_1} t^{d_2}$ in $C_n(q,t)$ is less than or equal to $p(\delta, k)$, and the equality holds if and only if one the following conditions holds:*

- $k \leq n - 3$, *or*
- $k = n - 2$ *and* $\delta = 1$, *or*
- $\delta = 0$.

This theorem is a consequence of Theorem C. It contains [8, Theorem 6] and a result of N.Bergeron and Z.Chen [1, Corollary 8.3.1] as special cases. In fact it proves [8, Conjecture 8]. We feel that the coefficient of $q^{d_1} t^{d_2}$ for general k can also be expressed in terms of partition numbers, only that the expression might be complicated. For example, we give the following conjecture which is verified for $6 \leq n \leq 10$.

Conjecture. Let n, d_1, d_2, δ, k be as in Theorem 1.3. If $n - 2 \leq k \leq 2n - 8$ and $\delta \geq k$, then the coefficient of $q^{d_1} t^{d_2}$ in $C_n(q,t)$ equals

$$p(k) - 2[p(0) + p(1) + \cdots + p(k - n + 1)] - p(k - n + 2).$$

As a corollary of Theorem 1.3 , we can compute some higher degree terms of the specialization at $t = q$.

Corollary 1.4

$$C_n(q,q) = \sum_{k=0}^{n-3} \left(p(k) \left(\binom{n}{2} - 3k + 1 \right) + 2 \sum_{i=1}^{k-1} p(i,k) \right) q^{\binom{n}{2} - k} + (\textit{lower degree terms}).$$

The following theorem immediately implies Theorem 1.3.

Theorem 1.5 *Let d_1, d_2 be non-negative integers d_1, d_2 with $d_1 + d_2 \leq \binom{n}{2}$. Define $k = \binom{n}{2} - d_1 - d_2$ and $\delta = \min(d_1, d_2)$. Then $\dim(M_n)_{d_1,d_2} \leq p(\delta, k)$, and the equality holds if and only if one the following conditions holds:*

- $k \leq n - 3$, *or*
- $k = n - 2$ *and* $\delta = 1$, *or*
- $\delta = 0$.

In case the equality holds, there is an explicit construction of a basis of $(M_n)_{d_1,d_2}$.

The idea of the construction of the basis in the above theorem consists of two parts:
(1) Prove that

$$\dim(M_n)_{d_1,d_2} \leq p(\delta, k)$$

using a new characterization of q, t-Catalan numbers. The characterization is as follows, and is discovered independently by A. Woo [10].

Let $\mathfrak{D}_n^{catalan}$ be the set consisting of $D \subset \mathbb{N} \times \mathbb{N}$, where D contains n points satisfying the following conditions.

(a) If $(p, 0) \in D$ then $(i, 0) \in D, \forall i \in [0, p]$.

(b) For any $p \in \mathbb{N}$,

$$\#\{j \mid (p+1, j) \in D\} + \#\{j \mid (p, j) \in D\} \geq \max\{j \mid (p, j) \in D\} + 1.$$

(If $\{j \mid (p, j) \in D\} = \emptyset$, then we require that no point $(i, j) \in D$ satisfies $i \geq p$.) Denote by $\deg_x D$ (resp. $\deg_y D$) the sum of the first (resp. second) components of the n points in D.

Proposition 1.6 *The coefficient of $q^{d_1}t^{d_2}$ in the q, t-Catalan number $C_n(q, t)$ is equal to*

$$\#\{D \in \mathfrak{D}_n^{catalan} \mid \deg_x D = d_1, \deg_y D = d_2\}.$$

(2) Construct a set of $p(\delta, k)$ linearly independent elements in $(M_n)_{d_1,d_2}$. It seems difficult (as least to the authors) to test directly whether a given set of elements in $(M_n)_{d_1,d_2}$ are linearly independent. We define a map φ sending an alternating polynomial $f \in \mathbb{C}[\mathbf{x}, \mathbf{y}]^\epsilon$ to a polynomial ring

$$\mathbb{C}[\rho] := \mathbb{C}[\rho_1, \rho_2, \rho_3, \ldots].$$

The map has two desirable properties: (i) for many f, $\varphi(f)$ can be easily computed, and (ii) for each bi-degree (d_1, d_2), φ induces a morphism $\bar{\varphi} : (M_n)_{d_1,d_2} \to \mathbb{C}[\rho]$ of $\mathbb{C}$-modules. Then we use the fact the linear dependency is easier to check in $\mathbb{C}[\rho]$ than in $(M_n)_{d_1,d_2}$. The map φ is defined as below.

1.2.3 Maps φ and $\bar{\varphi}$.

(a) Define the map $\varphi : \mathfrak{D}_n \to \mathbb{Z}[\rho]$ as follows. Let $D = \{(a_1, b_1), ..., (a_n, b_n)\} \in \mathfrak{D}_n$, $k = \binom{n}{2} - \sum_{i=1}^n (a_i + b_i)$, and define

$$\varphi(D) := (-1)^k \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n \left(\sum \rho_{w_1} \rho_{w_2} \cdots \rho_{w_{b_i}} \right),$$

where $(w_1, \ldots, w_{b_i})$ in the sum $\sum \rho_{w_1} \rho_{w_2} \cdots \rho_{w_{b_i}}$ runs through the set

$$\{(w_1, \ldots, w_{b_i}) \in \mathbb{N}^{b_i} \mid w_1 + ... + w_{b_i} = \sigma(i) - 1 - a_i - b_i\},$$

with the convention that

$$\sum \rho_{w_1} ... \rho_{w_{b_i}} = \begin{cases} 0 & \text{if } \sigma(i) - 1 - a_i - b_i < 0; \\ 0 & \text{if } b_i = 0 \text{ and } \sigma(i) - 1 - a_i - b_i > 0; \\ 1 & \text{if } b_i = 0 \text{ and } \sigma(i) - 1 - a_i - b_i = 0. \end{cases}$$

(b) Here is an equivalent definition of $\varphi(D)$. Define the weight of ρ_i to be i for $i \in \mathbb{N}^+$ and define the weight of $\rho_0 = 1$ to be 0. Naturally the weight of any monomial $c\rho_{i_1}...\rho_{i_n}$ ($c \in \mathbb{Z}$) is defined to be $i_1 + ... + i_n$. For $\mathrm{w} \in \mathbb{N}$ and a power series $f \in \mathbb{Z}[[\rho_1, \rho_2, \dots]]$, denote by $\{f\}_{\mathrm{w}}$ the sum of terms of weight-w in f, which is a polynomial. Define

$$h(b, \mathrm{w}) := \left\{(1 + \rho_1 + \rho_2 + \cdots)^b\right\}_{\mathrm{w}}, \quad b \in \mathbb{N}, \mathrm{w} \in \mathbb{Z}.$$

Naturally $h(b, \mathrm{w}) = 0$ if $\mathrm{w} < 0$. Also assume $(1 + \rho_1 + \rho_2 + \cdots)^0 = 1$. Then

$$\varphi(D) = (-1)^k \begin{vmatrix} h(b_1, -|P_1|) & h(b_1, 1 - |P_1|) & h(b_1, 2 - |P_1|) & \cdots & h(b_1, n - 1 - |P_1|) \\ h(b_2, -|P_2|) & h(b_2, 1 - |P_2|) & h(b_2, 2 - |P_2|) & \cdots & h(b_2, n - 1 - |P_2|) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h(b_n, -|P_n|) & h(b_n, 1 - |P_n|) & h(b_n, 2 - |P_n|) & \cdots & h(b_n, n - 1 - |P_n|) \end{vmatrix}.$$

(c) Let $D_1, \dots, D_\ell \in D'$ be of the same bi-degree and $\sum_{i=1}^{\ell} c_i D_i$ be the formal sum for any $c_i \in \mathbb{C}$ ($1 \le i \le \ell$). Define

$$\varphi(\sum_{i=1}^{\ell} c_i D_i) := \sum_{i=1}^{\ell} c_i \, \varphi(D_i).$$

For any bi-homogeneous alternating polynomials $f = \sum_{i=1}^{\ell} c_i \, \Delta(D_i) \in \mathbb{C}[\mathbf{x}, \mathbf{y}]^\epsilon$, we define

$$\varphi(f) := \varphi(\sum_{i=1}^{\ell} c_i D_i) = \sum_{i=1}^{\ell} c_i \, \varphi(D_i)$$

by abuse of notation. □

Proposition 1.7 *Fix any pair of nonnegative integers* (d_1, d_2)*, the map* φ *induces a well-defined linear map*

$$\bar{\varphi} : (M_n)_{d_1, d_2} \longrightarrow \mathbb{C}[\rho].$$

Moreover, this map $\bar{\varphi}$ is conjecturally injective. And our future work is to generalizing it to the case $I_n^m / (\mathbf{x}, \mathbf{y}) I_n^m$ for any positive integer m.

Acknowledgements

We are grateful to François Bergeron, Mahir Can, Jim Haglund, Nick Loehr, Alex Woo and Alex Yong for valuable discussions and correspondence.

References

[1] N. Bergeron and Z. Chen, Basis of Diagonally Alternating Harmonic Polynomials for low degree, arXiv: 0905.0377.

[2] A. M. Garsia and J. Haglund, A positivity result in the theory of Macdonald polynomials, Proc. Natl. Acad. Sci. USA **98** (2001), no. 8, 4313–4316 (electronic).

[3] A. M. Garsia and J. Haglund, A proof of the q, t-Catalan positivity conjecture. LaCIM 2000 Conference on Combinatorics, Computer Science and Applications (Montreal, QC). Discrete Math. 256 (2002), no. 3, 677–717.

[4] A. M. Garsia and M. Haiman, A Remarkable q; t-Catalan sequence and q-Lagrange inversion, J. Algebraic Combin. **5** (1996), 191–244.

[5] M. Haiman, Hilbert schemes, polygraphs and the Macdonald positivity conjecture, J. Amer. Math. Soc. **14** (2001), no. 4, 941–1006.

[6] M. Haiman, Vanishing theorems and character formulas for the Hilbert scheme of points in the plane, Invent. Math. **149** (2002), no. 2, 371–407.

[7] M. Haiman, Commutative algebra of n points in the plane, With an appendix by Ezra Miller. Math. Sci. Res. Inst. Publ., 51, Trends in commutative algebra, 153–180, Cambridge Univ. Press, Cambridge, 2004.

[8] K. Lee, L. Li, Notes on a minimal set of generators for the radical ideal defining the diagonal locus of $(\mathbb{C}^2)^n$, Arxiv math 0901.1176.

[9] K. Lee, L. Li, q, t-Catalan numbers and generators for the radical ideal defining the diagonal locus of $(\mathbb{C}^2)^n$, ArXiv math 0909.1612.

[10] A. Woo, private communication.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 758–765

Toric Ideals of Flow Polytopes

Matthias Lenz

Technische Universität Berlin, Sekretariat MA 4-5, Straße des 17. Juni 136, 10623 Berlin

Abstract. We show that toric ideals of flow polytopes are generated in degree 3. This was conjectured by Diaconis and Eriksson for the special case of the Birkhoff polytope. Our proof uses a hyperplane subdivision method developed by Haase and Paffenholz.

It is known that reduced revlex Gröbner bases of the toric ideal of the Birkhoff polytope B_n have at most degree n. We show that this bound is sharp for some revlex term orders. For $(m \times n)$-transportation polytopes, a similar result holds: they have Gröbner bases of at most degree $\lfloor mn/2 \rfloor$. We construct a family of examples, where this bound is sharp.

Résumé. Nous démontrons que les idéaux toriques des polytopes de flot sont engendrés par un ensemble de degré 3. Cela a été conjecturé par Diaconis et Eriksson pour le cas particulier du polytope de Birkhoff. Notre preuve utilise une méthode de subdivision par hyperplans, développée par Haase et Paffenholz.

Il est bien connu que les bases de Gröbner revlex réduite du polytope de Birkhoff B_n sont au plus de degré n. Nous démontrons que cette borne est optimale pour quelques ordres revlex. Pour les polytopes de transportation de dimension $(m \times n)$, il existe un résultat similaire : leurs bases de Gröbner sont au plus de degré $\lfloor mn/2 \rfloor$. Nous construisons une famille d'exemples pour lesquels cette borne est atteinte.

Resumen. Demostramos que los ideales tóricos de politopos de flujo se generan en grado 3. Esto fue conjeturado por Diaconis y Eriksson para el caso especial del politopo de Birkhoff. Nuestra demostración utiliza un método de subdivisión de hiperplanos desarrollado por Haase y Paffenholz.

Se sabe que las bases de Gröbner revlex reducidas de los ideales tóricos del politopo de Birkhoff B_n tienen como máximo grado n. Se demuestra que este límite es tight en algunos ordenes de termines revlex. Para politopos de transporte $(m \times n)$, existe un resultado similar: tienen bases de Gröbner de máximo grado $\lfloor mn/2 \rfloor$. Construimos una familia de ejemplos, mostrando que este límite es tight.

Keywords: Toric ideal, Flow polytope, Transportation polytope, Gröbner basis, Markov basis

1 Introduction

Let $G = (V, A)$ be a directed graph and $\boldsymbol{d} \in \mathbb{Z}^V$, $\boldsymbol{l}, \boldsymbol{u} \in \mathbb{N}^A$. A *flow* on G is a function $f : A \to \mathbb{R}_{\geq 0}$ that respects the lower and upper bounds $\boldsymbol{l}$ and $\boldsymbol{u}$ and satisfies the demand $\boldsymbol{d}$, *i. e.* for every vertex v, the flow entering v minus the flow leaving v equals d_v. A *flow polytope* is the set of all flows with fixed parameters $G, \boldsymbol{d}, \boldsymbol{u}, \boldsymbol{l}$.

An important special case are *transportation polytopes*. They can be written as sets of $(m \times n)$-matrices whose row and column sums equal some fixed positive integers.

Given a lattice polytope P, the relations among the lattice points in P define the *toric ideal* I_P. Generating sets of toric ideals correspond to Markov bases that are used in statistics *e. g.* for sampling from the set of all contingency tables with given marginals ([2]). In particular, small generating systems that can be handled by computers are of practical interest. *Diaconis and Eriksson* ([1]) conjectured that the toric ideal of the Birkhoff polytope B_n (the convex hull of all $(n \times n)$-permutation matrices) is generated in degree 3. They proved this by massive computations for $n \leq 6$. For arbitrary $n \geq 4$, they showed that I_{B_n} has a generating set of degree $n-1$.

Haase and Paffenholz ([4]) proved that the toric ideals of almost all (3×3)-transportation polytopes and particularly the smooth ones are generated in degree 2. The only exception is the Birkhoff polytope B_3, whose toric ideal is generated in degree 3.

Our Main Theorem proves the Diaconis-Eriksson conjecture and generalizes the result of Haase and Paffenholz:

Theorem 1.1 (Main Theorem) *Toric ideals of flow polytopes are generated in degree* 3*.*

Toric ideals define toric varieties, which are an important class of examples in algebraic geometry. A lattice polytope P is smooth if the normal fan at all vertices is unimodular. Equivalently, the corresponding projective variety X_P has to be smooth. It was conjectured that if P is smooth, then the defining ideal is generated by quadrics ([11, Conjecture 2.9]). The original motivation of our research was to check if this conjecture holds for flow polytopes.

In *Section 2*, we review some important definitions and theorems. In *Section 3*, we describe a method that can be used to prove degree bounds for generating sets and Gröbner bases of toric ideals: first, we choose a nice triangulation of the point set and then, we use a correspondence between Gröbner bases and triangulations established by *Sturmfels*. This hyperplane subdivision method was developed by Haase and Paffenholz in [4]. In *Section 4*, we briefly describe how the method described in Section 3 can be used to prove our Main Theorem.

Diaconis and Sturmfels showed that all revlex Gröbner bases of I_{B_n} are at most of degree n ([2, Theorem 6.1], [10, Theorem 14.8]). Computational experiments provide evidence that this bound might be optimal ([1, Remark 9]). In *Section 5* we show that for some revlex term orders, this bound is indeed optimal. Both the bound and the examples can be generalized to $(m \times n)$-transportation polytopes: reduced Gröbner bases with respect to a certain class of term orders are at most of degree $\lfloor mn/2 \rfloor$ and we construct a family of transportation polytopes and term orders, where this bound is almost sharp.

The proofs that are missing in this extended abstract are contained in the arXiv version ([6]).

2 Background

In this section, we review some important definitions and theorems.

Notation: $\mathbb{N} = \{0, 1, 2, \ldots\}$. Matrices are denoted by capital letters, vectors by bold faced small letters. Their entries are denoted by the corresponding small letters. Let $\boldsymbol{a} = (a_1, \ldots, a_n) \in \mathbb{N}^n$. We write $\boldsymbol{x}^{\boldsymbol{a}}$ to denote the monomial $\prod_{i=1}^{n} x_i^{a_i} \in k[\boldsymbol{x}] = k[x_1, \ldots, x_n]$.

The term *polytope* always refers to a convex lattice polytope *i. e.* all vertices of our polytopes are integral. For background information on polyhedral geometry and polytopes see *Schrijver*'s or *Ziegler*'s book [9, 12].

Flow polytopes: *Flow polytopes* (or transshipment polytopes) are the main geometric objects we are dealing with. Let $G = (V, A)$ be a directed graph and $\boldsymbol{d} \in \mathbb{Z}^V$, $\boldsymbol{l}, \boldsymbol{u} \in \mathbb{N}^A$. Let $M_G \in \{-1, 0, 1\}^{V \times A}$ denote the vertex-arc incidence matrix of G.

Note that the definition of *flow polytopes* given in the introduction is equivalent to

$$F = F_G = F_{G,\boldsymbol{d},\boldsymbol{u},\boldsymbol{l}} = \left\{ \boldsymbol{f} \in \mathbb{R}^A_{\geq 0} \,\middle|\, M_G \boldsymbol{f} = \boldsymbol{d},\, \boldsymbol{l} \leq \boldsymbol{f} \leq \boldsymbol{u} \right\} \tag{1}$$

It is a standard fact that M_G is totally-unimodular. This implies that the polytope F has integral vertices.

Throughout this paper, we suppose that all our flow polytopes F are *homogeneous*, *i. e.* F is contained in an affine hyperplane, that does not contain the origin. If $\boldsymbol{d} \neq \boldsymbol{0}$ this statement holds. Otherwise, we consider the homogenized polytope $\{1\} \times F$.

An important special case of flow polytopes are transportation polytopes. In statistics, they appear under the name 2-*way contingency tables*.

Let $m, n \in \mathbb{Z}_{\geq 1}, \boldsymbol{r} \in \mathbb{Z}^m_{\geq 1}, \boldsymbol{c} \in \mathbb{Z}^n_{\geq 1}$ be two vectors satisfying $\sum_{i=1}^n c_i = \sum_{i=1}^m r_i =: s$. The *transportation polytope* $T_{\boldsymbol{rc}}$ is defined as

$$T_{\boldsymbol{rc}} = \left\{ A \in \mathbb{R}^{m \times n}_{\geq 0} \,\middle|\, \sum_{i=1}^m a_{ij} = c_j, \sum_{j=1}^n a_{ij} = r_i \right\} \tag{2}$$

The upper $((m-1) \times (n-1))$-minor of a matrix $A \in T_{\boldsymbol{rc}}$ determines all other entries. Hence, the dimension of $T_{\boldsymbol{rc}}$ is at most $(m-1)(n-1)$. On the other hand, $a_{ij} = r_i c_j / s$ determines an interior point, so that the dimension is exactly $(m-1)(n-1)$.

If $\boldsymbol{r} = \boldsymbol{c} = (1, \ldots, 1)$, we obtain an important example: the Birkhoff polytope B_n.

Toric ideals and Gröbner bases: This paragraph defines toric ideals and Gröbner bases as in *Sturmfels*'s book ([10]). Let k be a field and let $P \subseteq \mathbb{R}^d$ be a homogeneous lattice polytope. The set of its lattice points $\mathcal{A} = \{\boldsymbol{a}_i \,|\, i \in I\}$ defines a semigroup homomorphism $\pi : \mathbb{N}^I \to \mathbb{Z}^d, \boldsymbol{u} \mapsto \sum_{i \in I} u_i \boldsymbol{a}_i$, which can be lifted to a ring homomorphism

$$\hat{\pi} : k[x_i \,|\, i \in I] \to k[t_1^{\pm 1}, \ldots t_d^{\pm 1}], \qquad x_i \mapsto \mathbf{t}^{\boldsymbol{a}_i} \tag{3}$$

Its kernel is the homogeneous ideal $I_{\mathcal{A}} = \left\langle \mathbf{x}^{\mathbf{u}} - \mathbf{x}^{\mathbf{v}} \;\middle|\; \sum_{i \in I} u_i \boldsymbol{a}_i = \sum_{i \in I} v_i \boldsymbol{a}_i \right\rangle$. This ideal is called the *toric ideal* associated to $\mathcal{A}$ (or P respectively).

A binomial $\boldsymbol{x}^{\boldsymbol{u}} - \boldsymbol{x}^{\boldsymbol{v}} \in I_{\mathcal{A}}$ corresponds to a relation between points in $\mathcal{A}$. For example, for $\mathcal{A} = B_3 \cap \mathbb{Z}^{3 \times 3}$, $I_{\mathcal{A}}$ is generated by the binomial that corresponds to the relation $\sum_{\det(M)=1} M = \sum_{\det(M)=-1} M$.

A total order $\prec$ on $\mathbb{N}^n$ is a *term order* if $\boldsymbol{a} \prec \boldsymbol{b}$ implies $\boldsymbol{a} + \boldsymbol{c} \prec \boldsymbol{b} + \boldsymbol{c}$ for all $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{N}^n$ and the zero vector is the unique minimal element. An important example is the graded reverse lexicographic (revlex) order: $\boldsymbol{a} \prec_{\text{revlex}} \boldsymbol{b}$ if $\sum_i a_i < \sum_i b_i$ or $\sum_i a_i = \sum_i b_i$ and the rightmost non-zero entry in $\boldsymbol{a} - \boldsymbol{b}$ is positive. Note that the revlex order depends on the order of the variables.

Let $I \subseteq k[x_1, \ldots, x_n]$ be an ideal and let $\prec$ denote a term order on $\mathbb{N}^n$. For $f \in k[x_1, \ldots, x_n]$, let $\text{in}_\prec(f)$ denote the initial (largest) term of f with respect to $\prec$. A finite set $\mathcal{G} \subseteq I$ is a Gröbner basis of I if for every $f \in I$, there exists a $g \in \mathcal{G}$, s. t. $\text{in}_\prec(g) | \text{in}_\prec(f)$. A Gröbner basis $\mathcal{G}$ is called *reduced* if for two distinct elements $g, g' \in \mathcal{G}$, no term of g' is divisible by $\text{in}_\prec(g)$. The reduced Gröbner basis is unique if we fix an ideal and a term order and require the coefficient of the initial term of every element to be 1.

Triangulations: We assume that the reader is familiar with *regular* and *pulling triangulations* (see *e. g.* [5] or [8]). A triangulation is called *unimodular* if the normalized volume of all simplices contained in it equals 1. Furthermore, we will use the fact that regular triangulations of a point set correspond to a certain class of term orders (see [10, Chapter 8]).

3 The Hyperplane Subdivision Method

In this section, we describe a general method for showing degree bounds of generating sets and Gröbner bases of toric ideals that was developed by Haase and Paffenholz ([4], see also [3, Proposition IV.2.5]).

A flow polytope has a canonical subdivision into polytopes contained in lattice translates of a unit cube: we slice F along hyperplanes of type $H_{ak} = \{\boldsymbol{x} \,|\, x_a = k\}$. Let $F = F_{G,\boldsymbol{d},\boldsymbol{u},\boldsymbol{l}}$ be a flow polytope. For $\boldsymbol{k} \in \mathbb{Z}^A$ we define a *cell of* F as $Z_F(\boldsymbol{k}) = \{\boldsymbol{f} \in F \,|\, k_a \leq f_a \leq k_a + 1 \text{ for all } a \in A\}$. Cells are flow polytopes using the same graph with tighter upper and lower bounds. Thus, they are lattice polytopes. For our purposes, it is acceptable to identify a cell with the translated cell $Z_F(\boldsymbol{k}) - \boldsymbol{k} \subseteq [0,1]^A$.

We will use a particular class of regular triangulations that we call *subdivide-and-pull triangulations*. They are obtained in the following way: start by subdividing the flow polytope along hyperplanes into cells as defined above. Then determine a pulling triangulation of each of the cells.

The following theorem is used in our proof of the Main Theorem:

Theorem 3.1 *Let F be a flow polytope and $k \geq 2$. I_F contains a generating set of at most degree k if the toric ideal of every cell of F contains a generating set of at most degree k.*

Let Δ be a subdivide-and-pull triangulation of F and $\mathcal{G}$ the reduced Gröbner basis with respect to the term order that corresponds to Δ. $\mathcal{G}$ has at most degree k if all cells have a Gröbner basis of at most degree k.

The main ingredient of the proof is the following theorem, which is a conglomerate of Corollaries 8.4 and 8.9 in Sturmfels's book ([10]):

Theorem 3.2 *Let P be a polytope and Δ be a regular, unimodular triangulation of P. Let $\prec_\Delta$ be the term order corresponding to Δ. Then, the initial ideal of I_P with respect to $\prec_\Delta$ is given by*

$$\mathrm{in}_{\prec_\Delta}(I_P) = \left\langle \boldsymbol{x}^F \,|\, F \textit{ is a minimal non-face of } \Delta \right\rangle \tag{4}$$

4 On the Proof of the Main Theorem

In this section, we briefly describe the idea of the proof of the Main Theorem. First, the result is proved for transportation polytopes. The general statement can be reduced to this special case.

Theorem 4.1 *Toric ideals of transportation polytopes are generated in degree* 3*.*

Proof (idea): Due to Theorem 3.1, it suffices to show that the bound holds for all cells. Let Z be a cell of an $(m \times n)$-transportation polytope and let $\mathcal{A} = Z \cap \mathbb{Z}^{m\times n}$. Consider the set $J_\mathcal{A}$ of binomials in $I_\mathcal{A}$ that cannot be expressed by binomials of smaller degree. Let $\boldsymbol{x}^{\boldsymbol{u}} - \boldsymbol{x}^{\boldsymbol{v}} \in J_\mathcal{A}$ be a binomial s. t. $\min\{d(M,N) \,|\, M \in \mathrm{supp}(\boldsymbol{u}), N \in \mathrm{supp}(\boldsymbol{v})\}$ is minimal over all binomials in $J_\mathcal{A}$, where d denotes the Hamming distance. Let $M \in \mathrm{supp}(\boldsymbol{u})$ and $N \in \mathrm{supp}(\boldsymbol{v})$ be two matrices realizing this minimal Hamming distance. Suppose $\deg(\boldsymbol{x}^{\boldsymbol{u}} - \boldsymbol{x}^{\boldsymbol{v}}) \geq 4$. By an analysis of the matrices M and N, we can reach a contradiction. □

Proof of the Main Theorem: The Main Theorem can be reduced to Theorem 4.1 by the transformation described in [9, 21.6a]. □

5 Gröbner Bases for Transportation Polytopes

In this section, we discuss degree bounds for Gröbner bases of transportation polytopes and we construct Gröbner bases in high degree.

For the toric ideal of the Birkhoff polytope, the following degree bound for their revlex Gröbner bases is known:

Theorem 5.1 ([2, Theorem 6.1],[10, Theorem 14.8]) *Let I_{B_n} be the toric ideal of the Birkhoff polytope B_n. Let $\mathcal{G}$ be a reduced Gröbner basis of I_{B_n} with respect to an arbitrary reverse lexicographic term order.*

Then, $\mathcal{G}$ has at most degree n.

The proof of this theorem can easily be generalized to prove degree bounds for Gröbner bases of transportation polytopes:

Theorem 5.2 *Let $T_{\boldsymbol{rc}}$ be an $(m \times n)$-transportation polytope and let $\prec$ denote a term order that corresponds to a subdivide-and-pull triangulation (as defined in Section 3).*

Then, the reduced Gröbner basis of $I_{T_{\boldsymbol{rc}}}$ has at most degree $\left\lfloor \frac{m \cdot n}{2} \right\rfloor$.

This improves a known degree bound for reduced Gröbner bases ([10, Proposition 13.15]) by a factor of approximately 2.

Both theorems are (almost) as good as they can get, in the following sense:

Theorem 5.3 ($\boldsymbol{B_n}$ has revlex Gröbner bases in degree $\boldsymbol{n}$) *Let n be even. Then there exists a revlex term order $\prec$, s. t. the reduced Gröbner basis $\mathcal{G}_\prec$ of I_{B_n} has exactly degree n.*

Theorem 5.4 (Gröbner bases in high degree for transportation polytopes) *Let m and n be even. Then there exists a smooth $(m \times n)$-transportation polytope $T_{\boldsymbol{rc}}$ and a term order $\prec$, s. t. the reduced Gröbner basis $\mathcal{G}_\prec$ of $I_{T_{\boldsymbol{rc}}}$ has degree at least $\frac{m(n-2)}{2}$.*

The term order can be chosen to be revlex or it can correspond to a subdivide-and-pull triangulation (as defined in Section 3).

Theorem 5.3 supports the experimental result of Diaconis and Eriksson, who suggested that Gröbner bases of I_{B_n} have exactly degree n ([1, Remark 9]).

We illustrate the constructions used to prove Theorems 5.3 and 5.4 in two examples:

Example 5.5 This example shows that Theorem 5.3 holds for $n = 6$.

$$\underbrace{\begin{bmatrix} 1&&&&& \\ &&&&1& \\ &&&&&1 \\ &1&&&& \\ &&&1&& \\ &&1&&& \end{bmatrix}}_{A_1} + \underbrace{\begin{bmatrix} &&&1&& \\ &1&&&& \\ &&&&&1 \\ 1&&&&& \\ &&1&&& \\ &&&&1& \end{bmatrix}}_{A_2} + \underbrace{\begin{bmatrix} &&&1&& \\ &&&&1& \\ &&1&&& \\ &&&&&1 \\ 1&&&&& \\ &1&&&& \end{bmatrix}}_{A_3} + \underbrace{\begin{bmatrix} &&&&1& \\ 1&&&&& \\ &&&&&1 \\ &&&1&& \\ &1&&&& \\ &&1&&& \end{bmatrix}}_{B_1} + \underbrace{\begin{bmatrix} &&&1&& \\ &&&&&1 \\ &1&&&& \\ 1&&&&& \\ &&&&1& \\ &&1&&& \end{bmatrix}}_{B_2} + \underbrace{\begin{bmatrix} &&1&&& \\ &&&1&& \\ &&&&1& \\ 1&&&&& \\ &1&&&& \\ &&&&&1 \end{bmatrix}}_{B_3} \tag{5}$$

$$= \underbrace{\begin{bmatrix} 1 &&&&& \\ & 1 &&&& \\ && 1 &&& \\ &&& 1 && \\ &&&& 1 & \\ &&&&& 1 \end{bmatrix}}_{J_6} + \underbrace{\begin{bmatrix} && 1 &&& \\ 1 &&&&& \\ & 1 &&&& \\ &&&&& 1 \\ &&& 1 && \\ &&&& 1 & \end{bmatrix}}_{B} + 2 \cdot \underbrace{\begin{bmatrix} &&& 1 && \\ &&&& 1 & \\ &&&&& 1 \\ 1 &&&&& \\ & 1 &&&& \\ && 1 &&& \end{bmatrix}}_{C} + \underbrace{\begin{bmatrix} &&&& 1 & \\ &&& 1 && \\ &&&&& 1 \\ 1 &&&&& \\ && 1 &&& \\ & 1 &&&& \end{bmatrix}}_{D_1} + \underbrace{\begin{bmatrix} &&& 1 && \\ &&&&& 1 \\ &&&& 1 & \\ & 1 &&&& \\ 1 &&&&& \\ && 1 &&& \end{bmatrix}}_{D_2} = \begin{bmatrix} 1 & & 1 & 3 & 1 & \\ 1 & 1 & & 1 & 2 & 1 \\ & 1 & 1 & & 1 & 3 \\ 3 & 1 & & 1 & & 1 \\ 1 & 2 & 1 & 1 & 1 & \\ & 1 & 3 & & 1 & 1 \end{bmatrix} \tag{6}$$

This equation corresponds to an element of I_{B_6}. We order the lattice points in B_6 s. t. J_6 is minimal and $A_1, A_2, A_3, B_1, B_2, B_3$ are smaller than all the remaining points. In the revlex order $\prec$ defined by this ordering, the left side of the equation corresponds to the initial term. One can show that it is a minimal generator of the initial ideal $\mathrm{in}_\prec(I_{B_6})$. Hence, the reduced Gröbner basis of I_{B_6} with respect to $\prec$ contains an element of degree 6.

Example 5.6 This example shows that Theorem 5.4 holds for $n = m = 6$. The following equation is a relation of lattice points in the polytope $T_{\boldsymbol{rc}}$ with $\boldsymbol{r} = \boldsymbol{c} = (3,3,3,3,3,3)$. A translated version of this relation is contained in the transportation polytope with marginals $\boldsymbol{r} = (39,39,39,39,39,39)$ and $\boldsymbol{c} = (3,3,3,3,3,219)$. By [4, Lemma 1], this polytope is smooth.

$$\underbrace{\begin{bmatrix} 1 &&& 1 && 1 \\ &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ & 1 & 1 && 1 & \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \end{bmatrix}}_{A_{12}} + \underbrace{\begin{bmatrix} 1 &&& 1 & 1 & \\ &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ & 1 & 1 &&& 1 \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \end{bmatrix}}_{A_{13}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ 1 &&& 1 && 1 \\ &&& 1 & 1 & 1 \\ 1 & 1 & 1 &&& \\ & 1 & 1 && 1 & \\ 1 & 1 & 1 &&& \end{bmatrix}}_{A_{22}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ 1 &&& 1 & 1 & \\ &&& 1 & 1 & 1 \\ 1 & 1 & 1 &&& \\ & 1 & 1 &&& 1 \\ 1 & 1 & 1 &&& \end{bmatrix}}_{A_{23}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ 1 &&& 1 && 1 \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \\ & 1 & 1 && 1 & \end{bmatrix}}_{A_{32}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ 1 &&& 1 & 1 & \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \\ & 1 & 1 &&& 1 \end{bmatrix}}_{A_{33}} \tag{7}$$

$$+ \underbrace{\begin{bmatrix} & 1 &&& 1 & 1 \\ &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ 1 && 1 & 1 && \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \end{bmatrix}}_{B_{12}} + \underbrace{\begin{bmatrix} && 1 && 1 & 1 \\ &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ 1 & 1 && 1 && \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \end{bmatrix}}_{B_{13}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ & 1 &&& 1 & 1 \\ &&& 1 & 1 & 1 \\ 1 & 1 & 1 &&& \\ 1 && 1 & 1 && \\ 1 & 1 & 1 &&& \end{bmatrix}}_{B_{22}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ && 1 && 1 & 1 \\ &&& 1 & 1 & 1 \\ 1 & 1 & 1 &&& \\ 1 & 1 && 1 && \\ 1 & 1 & 1 &&& \end{bmatrix}}_{B_{23}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ & 1 &&& 1 & 1 \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \\ 1 && 1 & 1 && \end{bmatrix}}_{B_{32}} + \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ && 1 && 1 & 1 \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \\ 1 & 1 && 1 && \end{bmatrix}}_{B_{33}} \tag{8}$$

$$= \begin{bmatrix} 2 & 1 & 1 & 10 & 11 & 11 \\ 2 & 1 & 1 & 10 & 11 & 11 \\ 2 & 1 & 1 & 10 & 11 & 11 \\ 10 & 11 & 11 & 2 & 1 & 1 \\ 10 & 11 & 11 & 2 & 1 & 1 \\ 10 & 11 & 11 & 2 & 1 & 1 \end{bmatrix} = 1 \cdot \underbrace{\begin{bmatrix} 1 &&&& 1 & 1 \\ 1 &&&& 1 & 1 \\ 1 &&&& 1 & 1 \\ & 1 & 1 & 1 && \\ & 1 & 1 & 1 && \\ & 1 & 1 & 1 && \end{bmatrix}}_{C} + 10 \cdot \underbrace{\begin{bmatrix} &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \end{bmatrix}}_{D} + \underbrace{\begin{bmatrix} 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \\ 1 & 1 & 1 &&& \\ &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \\ &&& 1 & 1 & 1 \end{bmatrix}}_{E} \tag{9}$$

We order the lattice points of $T_{\boldsymbol{rc}}$ s. t. E is minimal and the A_{ij}s and B_{ij}s are smaller than all the remaining points.

In the revlex order $\prec$ defined by this ordering, the left side of the equation corresponds to the initial term. One can show that it is a minimal generator of the initial ideal $\mathrm{in}_\prec(I_{T_{rc}})$. Hence, the reduced Gröbner basis of $I_{T_{rc}}$ with respect to $\prec$ contains an element of degree 12.

Acknowledgements

This article contains the main results of my *Diplomarbeit* (Master's thesis) [7], written in 2007 at Freie Universität Berlin. I would like to thank my advisor Christian Haase for many inspiring hints and discussions.

References

[1] Persi Diaconis and Nicholas Eriksson. Markov bases for noncommutative Fourier analysis of ranked data. *Journal of Symbolic Computation*, 41(2), February 2006. `arXiv:math/0405060v2`.

[2] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.*, 26(1):363–397, 1998. Available at `http://www-stat.stanford.edu/~cgates/PERSI/papers/sturm98.pdf`.

[3] Christian Haase. *Lattice Polytopes and Triangulations*. PhD thesis, Technische Universität Berlin, 2000. Available at `http://opus.kobv.de/tuberlin/frontdoor.php?source_opus=90`.

[4] Christian Haase and Andreas Paffenholz. Quadratic Gröbner bases for smooth 3×3 transportation polytopes. *Journal of Algebraic Combinatorics*, 30(4):477–489, 2009. `arXiv:math.CO/0607194`.

[5] Carl W. Lee. Regular triangulations of convex polytopes. In *Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift*, volume 4 of *DIMACS Series in Discrete Math. and Theor. Comput. Sci.*, pages 443–456. American Mathematical Society, Providence, 1991.

[6] Matthias Lenz. Toric ideals of flow polytopes. `arXiv:0801.0495`.

[7] Matthias Lenz. Torische Ideale von Flusspolytopen. Master's thesis, Freie Universität Berlin, July 2007. `arXiv:0709.3570v3`.

[8] Jörg Rambau, Francisco Santos, and Jesus A. De Loera. *Triangulations: Structures for Algorithms and Applications*, volume 25 of *Algorithms and Computation in Mathematics*. Springer, 2010. Not yet published. Available: October 2010.

[9] Alexander Schrijver. *Combinatorial Optimization – Polyhedra and Efficiency*. Number 24A in Algorithms and Combinatorics. Springer, Berlin, 2003.

[10] Bernd Sturmfels. *Gröbner Bases and Convex Polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, Providence, Rhode Island, 1996.

[11] Bernd Sturmfels. Equations defining toric varieties. Kollár, János (ed.) et al., Algebraic geometry. Proceedings of the Summer Research Institute, Santa Cruz, CA, USA, July 9–29, 1995. Providence, RI: American Mathematical Society. Proc. Symp. Pure Math. 62(pt.2), 437-449, 1997. `arXiv:alg-geom/9610018v1`.

[12] Günter M. Ziegler. *Lectures on Polytopes*, volume 152 of *GTM*. Springer-Verlag, New York, 1995.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 766–777

Nonzero coefficients in restrictions and tensor products of supercharacters of $U_n(q)$ (extended abstract)

Stephen Lewis[1] and Nathaniel Thiem[2]

[1] *Department of Mathematics, University of Washington*
[2] *Department of Mathematics, University of Colorado at Boulder*

Abstract. The standard supercharacter theory of the finite unipotent upper-triangular matrices $U_n(q)$ gives rise to a beautiful combinatorics based on set partitions. As with the representation theory of the symmetric group, embeddings of $U_m(q) \subseteq U_n(q)$ for $m \leq n$ lead to branching rules. Diaconis and Isaacs established that the restriction of a supercharacter of $U_n(q)$ is a nonnegative integer linear combination of supercharacters of $U_m(q)$ (in fact, it is polynomial in q). In a first step towards understanding the combinatorics of coefficients in the branching rules of the supercharacters of $U_n(q)$, this paper characterizes when a given coefficient is nonzero in the restriction of a supercharacter and the tensor product of two supercharacters. These conditions are given uniformly in terms of complete matchings in bipartite graphs.

Résumé. La théorie standard des supercaractères des matrices triangulaires supérieures unipotentes finies $U_n(q)$ donne lieu à une merveilleuse combinatoire basée sur les partitions d'ensembles. Comme avec la théorie des représentations du groupe symétrique, Les plongements $U_m(q) \subseteq U_n(q)$ pour $m \leq n$ mènent aux règles de branchement. Diaconis et Isaacs ont montré que la restriction d'un supercaractère de $U_n(q)$ est une combinaison linéaire des supercaractères de $U_m(q)$ avec des coefficients entiers non négatifs (en fait, elle est polynomiale en q). Dans une première étape vers la compréhension de la combinatoire des coefficients dans les règles de branchement des supercaractères de $U_n(q)$, ce texte caractérise les coefficients non nuls dans la restriction d'un supercaractère et dans le produit des tenseurs de deux supercaractères. Ces conditions sont données uniformément en termes des couplages complets dans des graphes bipartis.

Keywords: supercharacters, set-partitions, matching, bipartite graphs, unipotent upper-triangular matrices

1 Introduction

The representation theory of the finite groups of unipotent upper-triangular matrices $U_n(\mathbb{F}_q)$ has traditionally been a hard problem, where even enumerating the irreducible representations is a well-known wild problem. In fact, it is not even known if the number of irreducible representations is polynomial in q (the Higman conjecture [12] suggests the affirmative). However, André [1, 2, 3, 4] and later Yan [18] demonstrated if one decomposes the regular representation into "nearly irreducible" pieces (called super-representations) rather than the usual irreducible pieces, one obtains a theory that is far more tractable with beautiful combinatorial underpinnings. Later, work of Arias-Castro, Diaconis and Stanley [7] demonstrated that this theory could even be used in place of the usual representation theory in an application to random walks, and for more general supercharacter theories, Otto [15] has shown that they can be used to bound nilpotence classes of nilpotent algebras.

While it has been a guiding principle that the supercharacter theory of $U_n(q)$ is analogous to the representation theory of the symmetric group S_n, many U_n-analogues of S_n results remain to be worked out. Some of the known observations include

(a) The irreducible characters of S_n are indexed by partitions of n, and the supercharacters of $U_n(\mathbb{F}_q)$ are indexed by (a q-analogue of) set partitions of $\{1, 2, \ldots, n\}$ [1, 18, 7],

(b) For S_n, Young subgroups are a natural family of subgroups which give the corresponding character rings a Hopf structure through induction and restriction. Similarly, [16] defines an analogue to Young subgroups for $U_n(q)$, noting that while in the S_n-case the particular embedding of the subgroup typically does not matter, in the $U_n(q)$-case it is critical [14, 17]. These subgroups are indexed by set-partitions instead of integer partitions.

(c) As an algebra, the ring of symmetric functions model restriction/induction branching rules for the characters of S_n considered simultaneously for all $n \geq 0$. The corresponding ring for the supercharacters of $U_n(\mathbb{F}_q)$ seems to be the ring of symmetric functions in non-commuting variables [16].

This paper attempts to better understand the combinatorics of branching rule coefficients for the supercharacters of $U_n(\mathbb{F}_q)$.

In the symmetric group case, the irreducible character $\chi^\mu \times \chi^\nu$ appears in the decomposition of the restricted character

$$\mathrm{Res}^{S_{|\lambda|}}_{S_{|\mu|} \times S_{|\nu|}}(\chi^\lambda)$$

only if $\mu, \nu \subset \lambda$. For $U_n(q)$ this paper gives both necessary and sufficient conditions for analogous result, using the close relationship between tensor products and restrictions in this case. In particular, the main results of this abstract are

Theorem 3.1. Given a set partition λ and subgroup $U_K \subseteq U_n$, there is a bipartite graph $\Gamma_K(\lambda)$ such that the trivial character appears in the decomposition of $\mathrm{Res}^{U_n}_{U_K}(\chi^\lambda)$ if and only if the graph has a complete matching.

Theorem 4.1. Given set partitions λ, μ, and ν, there is a bipartite graph $\Gamma(\lambda, \mu, \nu)$ such that the χ^ν appears in the decomposition of $\chi^\lambda \otimes \chi^\mu$ if and only if the graph has a complete matching.

Theorem 4.3. Given set partitions λ, μ and subgroup $U_K \subseteq U_n$, there is a bipartite graph $\Gamma_K(\lambda, \mu)$ such that the χ^μ appears in the decomposition of $\mathrm{Res}^{U_n}_{U_K}(\chi^\lambda)$ if and only if the graph has a complete matching.

The bipartite graphs referenced in all three results have a uniform construction as described in Section 4.1, and are remarkably easy to construct given the initial data. However, the description of the bipartite graph in Theorem 3.1 is particularly nice, so we describe it separately in Section 3.1. A fundamental part of extending Theorem 3.1 to Theorem 4.1 and Theorem 4.3 is a result that rewrites tensor products as restriction, as follows.

Theorem 4.2 Given two supercharacters χ^λ and χ^μ of U_K, there exists a supercharacter χ^ν and groups $U_{K'} \subseteq U_L$ with $U_{K'} \cong U_K$ such that $\chi^\lambda \otimes \chi^\mu$ (under the internal point-wise product) is the same character of U_K as $q^{-r}\mathrm{Res}^{U_L}_{U'_K}(\chi^\nu)$ is of $U_{K'}$, where $r \in \mathbb{Z}_{\geq 0}$.

While these results do not give a complete description of the coefficients, [8] showed that these coefficients are always positive integers (in fact, polynomial in q). Thus, a more explicit understanding of the coefficients remains open.

The supercharacter theory studied in this paper is a particular example of a supercharacter theory that has a more general construction on algebra groups [8]. Generalizations of this approach have also been studied by André and Neto for maximal unipotents subgroups in other Lie types [5, 6]. This study of a particular supercharacter theory is somewhat different from recent work by [11], which attempts to find all the possible supercharacter theories for a given finite group.

Acknowledgements

Part of this work (in particular Theorem 3.1) was in Lewis' undergraduate honors thesis at the University of Colorado at Boulder. Thiem was supported in part by an NSF grant DMS-0854893.

2 Preliminaries

This section sets up the necessary combinatorics of set partitions, which differs from some of the more standard formulations. From this point of view, the parts of the set partition are less important than the relative sizes of the numbers in the same part. We then review the definition of a supercharacter theory, and recall the specific supercharacter theory of interest for the finite groups of unipotent upper-triangular matrices, as developed by André [1, 2, 3, 4] and Yan [18].

2.1 Set partition combinatorics

Fix a prime power q, and let $\mathbb{F}_q$ be the finite field with q elements with additive group $\mathbb{F}_q^+$ and multiplicative group $\mathbb{F}_q^\times$.

For a finite subset $K \subseteq \mathbb{Z}_{\geq 1}$, let

$$\mathcal{A}_K(q) = \{i \overset{a}{\frown} j \mid i, j \in K, i < j, a \in \mathbb{F}_q^\times\},$$

and

$$\mathcal{A}(q) = \bigcup_{\substack{K\subseteq \mathbb{Z}_{\geq 1}\\ |K|<\infty}} \mathcal{A}_K(q),$$

where $\mathcal{A}_\emptyset(q) = \{\emptyset\}$. We will refer to the non-emptyset elements of $\mathcal{A}(q)$ as *arcs*.

Let

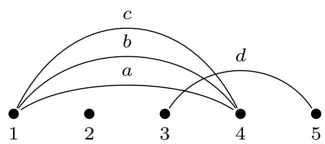
$$\mathcal{M}(q) = \bigcup_{\substack{K\subseteq \mathbb{Z}_{\geq 1}\\ |K|<\infty}} \mathcal{M}_K(q), \qquad \text{where} \qquad \mathcal{M}_K(q) = \{\text{finite multisets in } \mathcal{A}_K(q)\}.$$

For $\lambda \in \mathcal{M}_K(q)$ and $j, k \in K$, let

$$R_{jk} = \{j \overset{a}{\frown} k \in \lambda\}, \qquad m_{jk}(\lambda) = |R_{jk}|, \qquad \text{and} \qquad \mathrm{wt}_{jk}(\lambda) = \sum_{j \overset{a}{\frown} k \in R_{jk}} a \in \mathbb{F}_q, \tag{1}$$

where R_{jk} is a multiset. For example, if

$$\lambda = \{1 \overset{a}{\frown} 4, 1 \overset{b}{\frown} 4, 1 \overset{c}{\frown} 4, 3 \overset{d}{\frown} 5\} =$$



, for $a, b, c, d \in \mathbb{F}_q^{\times}$,

then

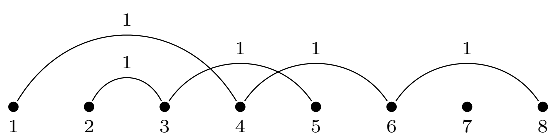
$$R_{14} = \{1 \overset{a}{\frown} 4, 1 \overset{b}{\frown} 4, 1 \overset{c}{\frown} 4\}, \quad m_{14}(\lambda) = 3, \quad \mathrm{wt}_{14}(\lambda) = a + b + c, \quad \text{and} \quad R_{23} = \emptyset.$$

Note we will use a diagrammatic representation of multisets $\lambda \in \mathcal{M}_K(q)$, by associating to each element of K a node (usually arranged along a horizontal line), and each arc $i \overset{a}{\frown} j$ in λ becomes a labeled edge connecting node i to node j.

A *q-set partition* of K is a multiset $\lambda \in \mathcal{M}_K(q)$ such that if $i \overset{a}{\frown} l, j \overset{b}{\frown} k \in \lambda$ are two distinct arcs, then $i \neq j$ and $k \neq l$. Let

$$\mathcal{S}(q) = \bigcup_{\substack{K \subseteq \mathbb{Z}_{\geq 1} \\ |K| < \infty}} \mathcal{S}_K(q), \qquad \text{where} \qquad \mathcal{S}_K(q) = \{q\text{-set partitions in } \mathcal{M}_K(q)\}.$$

Note that 2-set partitions of K are set partitions λ of K by the rule that i and j are in the same part if there is a sequence of arcs $i \overset{1}{\frown} j_1, j_1 \overset{1}{\frown} j_2, \ldots, j_{m-1} \overset{1}{\frown} j \in \lambda$. That is, the parts of the set partitions are the connected components of the diagrammatic representation of the 2-set partition. For example,



$\longleftrightarrow \{1, 4, 6, 8 \mid 2, 3, 5 \mid 7\}$.

In this sense, q-set partitions are a q-analogue of set partitions (although, strictly speaking, they are $(q-1)$-analogues of set partitions).

Let $\lambda \in \mathcal{M}(q)$. A *conflict in* λ *over* K is either

(CL) A pair of distinct arcs $i \overset{a}{\frown} l, j \overset{b}{\frown} k \in \lambda$ such that $i = j$ and $k < l$,

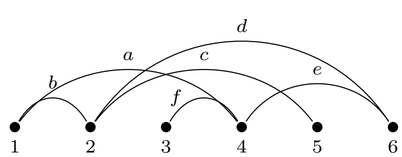
(CR) A pair of distinct arcs $i \overset{a}{\frown} l, j \overset{b}{\frown} k \in \lambda$ such that $i < j$ and $k = l$,

(CB) A pair of distinct arcs $i \overset{a}{\frown} l, j \overset{b}{\frown} k \in \lambda$ such that $i = j$ and $k = l$,

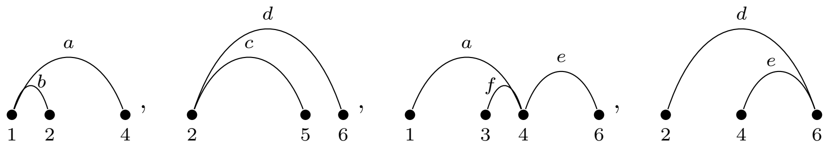
(CN) A nonempty multiset $\{i \overset{a}{\frown} k \in \lambda \mid i = j \text{ or } k = j\}$ for some $j \notin K$.

Example 1 *The multiset*

$\lambda =$



has conflicts



over $\{1, 2, 3, 5, 6\}$*, where the conflicts are of type (CL), (CL), (CN), and (CR).*

Conflicts are instances in a multiset which violate the conditions of membership in the set $\mathcal{S}_K(q)$.

Remark 1 *There are a variety of* q*-analogues of set partition or Stirling numbers in the literature. This particular* q*-analogue of set partitions is different from the one introduced [10] and only seems to appear in connection with supercharacters. There is also a standard construction for* q*-Stirling numbers (see for example [9]), where the Stirling number* $S(n, k; q)$ *is defined by "*q*-counting" the number of elements of* $\mathcal{S}_n(2)$ *with* $n - k$ *arcs. If we let* $S_q(n, k)$ *be the number of elements of* $\mathcal{S}_n(q)$ *with* $n - k$ *elements, we obtain a recursion*

$$\mathcal{S}_q(n, k) = \mathcal{S}_q(n-1, k-1) + k(q-1)\mathcal{S}_q(n-1, k),$$

which is different from the recursion for $S(n, k; q)$ *in [9].*

2.2 Supercharacters of $U_n(q)$

Supercharacters were first studied by André in relation to $U_n(q)$ as a way to find some more tractable way to understand the representation theory of $U_n(q)$. Diaconis and Isaacs [8] then generalized the concept to arbitrary finite groups, and we reproduce a version of this more general definition below.

A *supercharacter theory* of a finite group G is a pair $(\mathcal{K}, \mathcal{X})$ where $\mathcal{K}$ is a partition of G such that

(a) Each $K \in \mathcal{K}$ is a union of conjugacy classes,

(b) The identity element of G is in its own part in $\mathcal{K}$,

and $\mathcal{X}$ is a set of characters of G such that

(a) For each irreducible character ψ of G there is a unique $\chi \in \mathcal{X}$ such that $\langle \chi, \psi \rangle \neq 0$,

(b) The trivial character $\mathbb{1} \in \mathcal{X}$,

(c) The characters of $\mathcal{X}$ are constant on the parts of $\mathcal{K}$,

(d) $|\mathcal{K}| = |\mathcal{X}|$.

We will refer to the parts of $\mathcal{K}$ as *superclasses* and the characters of $\mathcal{X}$ as *supercharacters*. For more information on the implications of these axioms see [8] (including some redundancies in the definition).

For $n \in \mathbb{Z}_{\geq 1}$, let $M_n(\mathbb{F}_q)$ be the ring of $n \times n$ matrices with entries in the finite field $\mathbb{F}_q$ with q elements. Let

$$U_n(q) = \{u \in M_n(\mathbb{F}_q) \mid u_{ji} = 0, u_{ii} = 1, u_{ij} \in \mathbb{F}_q, i < j\}$$

be the group of unipotent upper-triangular matrices. For $K \subseteq \{1, 2, \ldots, n\}$, let

$$U_K(q) = \{u \in U_n(q) \mid u_{ij} = 0, i < j, \text{unless } i, j \in K\}.$$

Note that $U_K(q) \cong U_{|K|}(q)$.

For $U_K(q)$ there is a standard example of a supercharacter theory developed by André and Yan, where $\mathcal{K}$ and $\mathcal{X}$ are indexed by q-set partitions of K. For the purpose of this paper it suffices to recall the definition of the supercharacters. Fix a nontrivial group homomorphism,

$$\vartheta : \mathbb{F}_q^+ \longrightarrow \mathbb{C}^\times.$$

For $\lambda \in \mathcal{S}_K(q)$, there is a supercharacter χ^λ given by

$$\chi^\lambda = \bigotimes_{i \overset{a}{\frown} l \in \lambda} \chi^{i \overset{a}{\frown} l}, \tag{2}$$

where each $\chi^{i \overset{a}{\frown} l}$ is an irreducible character of $U_K(q)$ whose value on the superclass indexed by $\mu \in \mathcal{S}_K(q)$ is

$$\chi^{i \overset{a}{\frown} l}(\mu) = \begin{cases} 0, & \text{if } j \overset{b}{\frown} k \in \mu \text{ with } i = j < k < l \text{ or } i < j < k = l, \\ \dfrac{q^{|\{i<j<l \mid j \in K\}|}}{q^{|\{i<j<k<l \mid j \overset{b}{\frown} k \in \mu\}|}} \vartheta(a\mathrm{wt}_{il}(\mu)), & \text{otherwise.} \end{cases}$$

It can be quickly verified that the linear supercharacters of $U_K(q)$ correspond to $\lambda \in \mathcal{S}_K(q)$ with $i \overset{a}{\frown} l \in \lambda$ implies $\{i < j < l \mid j \in K\} = \emptyset$; the trivial character is $\chi^\emptyset$.

The superclass $\{1\}$ is indexed by $\emptyset \in \mathcal{S}_L(q)$. Thus, for $\lambda \in \mathcal{M}_L(q)$ with $L \subseteq \mathbb{Z}_{\geq 1}$, the degree of χ^λ is

$$\chi^\lambda(1) = \prod_{i \overset{a}{\frown} l \in \lambda} q^{|\{i<j<l \mid j \in L\}|}.$$

If $K \subseteq L$, then define

$$r_K^L(\lambda) = |\{(j, i \overset{a}{\frown} l) \in L \times \lambda \mid i < j < l, j \notin K\}|. \tag{3}$$

Note that if $\lambda \in \mathcal{M}_L(q) \cap \mathcal{M}_K(q)$ then $q^{r_K^L(\lambda)}$ is the ratio of the degrees of χ^λ as a character of $U_L(q)$ and χ^λ as a character of $U_K(q)$. It therefore is a constant that frequently comes up in the restriction of supercharacters. In fact, by inspection, if $\lambda \in \mathcal{M}_L(q) \cap \mathcal{M}_K(q)$, then

$$\mathrm{Res}_{U_K(q)}^{U_L(q)}(\chi^\lambda) = q^{r_K^L(q)} \chi^\lambda. \tag{4}$$

In general, supercharacters are orthogonal with respect to the usual inner product on class functions, and for $\lambda, \mu \in \mathcal{S}_K(q)$,

$$\langle \chi^\lambda, \chi^\mu \rangle = \delta_{\lambda\mu} q^{|\mathcal{C}(\lambda)|}, \qquad \text{where} \quad \mathcal{C}(\lambda) = \{(i \overset{a}{\frown} k, j \overset{b}{\frown} l) \in \lambda \times \lambda \mid i < j < k < l\}, \tag{5}$$

is the set of *crossings* in λ.

The papers [16, 17, 18] describe local rules for computing restrictions and tensor products. In principle, therefore, one can easily compute restrictions and tensor products in a recursive, algorithmic fashion (see [16] for a detailed description of this algorithm, and [13] for an implementation of this algorithm in Python). However, this algorithm does not give an obvious combinatorial interpretation of the resulting coefficients.

3 Coefficient of trivial character

This section investigates the coefficient of the trivial character $\mathbb{1} = \chi^{\emptyset}$ in the restriction from $U_L(q)$ to a subgroup $U_K(q)$. In particular, Theorem 3.1 characterizes when the coefficient of $\mathbb{1}$ is nonzero in the restriction of a supercharacter. Although the theorem seems somewhat specific, in later sections we will use it to analyze the coefficients of arbitrary supercharacters in both restrictions and tensor products.

3.1 Main result

Given a set partition $\lambda \in \mathcal{S}_L(q)$ and a subset $K \subseteq L$, let $\Gamma_K(\lambda)$ be the bipartite graph given by vertices

$$V_\bullet = \{i \frown j \in \lambda \mid i, j \in K\}$$
$$V_\circ = \{i \frown j \in \lambda \mid i, j \notin K\},$$

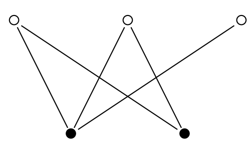
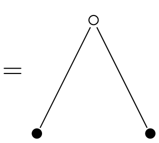
and an edge from $j \frown k \in V_\bullet$ to $i \frown l \in V_\circ$ if $i < j < k < l$. Note that this graph has in general far fewer vertices than λ has arcs. The following theorem is the main result of the paper, and is a model for the remaining results in this paper.

Theorem 3.1 *Let $K \subseteq L \subseteq \mathbb{Z}_{\geq 1}$ be finite sets, and let $\lambda \in \mathcal{S}_L(q)$ be a q-set partition. Then*

$$\langle \mathrm{Res}^{U_L}_{U_K}(\chi^\lambda), \mathbb{1} \rangle \neq 0,$$

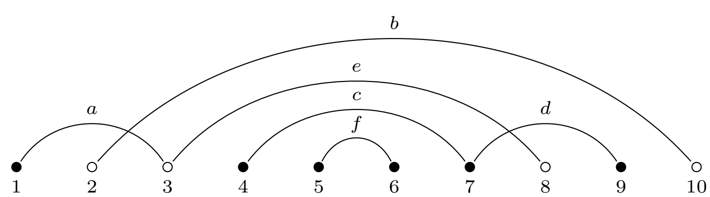
if and only if the graph $\Gamma_K(\lambda)$ has a complete matching from $V_\bullet$ to $V_\circ$.

Remark 2 *The complete matching referred to in Theorem 3.1 is a one-sided matching. That is, every element in $V_\bullet$ must be matched to a corresponding element in $V_\circ$, but there could potentially be elements of $V_\circ$ not matched to elements of $V_\bullet$. For example, if*

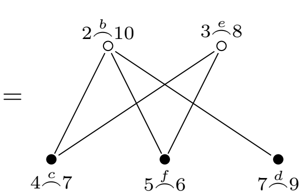
$\Gamma_1 =$  *and* $\Gamma_2 =$ 

then Γ_1 has a complete matching from $V_\bullet$ to $V_\circ$ and Γ_2 does not.

Example 2 *If $K = \{1, 4, 5, 6, 7, 9\}$ and*

$\lambda =$ 

then $V_\bullet = \{4 \overset{c}{\frown} 7, 5 \overset{f}{\frown} 6, 7 \overset{d}{\frown} 9\}$, $V_\circ = \{2 \overset{b}{\frown} 10, 3 \overset{e}{\frown} 8\}$, and

$\Gamma_K(\lambda) =$ 

Since this graph has no complete matchings from $V_\bullet$ *to* $V_\circ$*, by Theorem 3.1* $\langle \mathrm{Res}^{U_L}_{U_K}(\chi^\lambda), \mathbb{1}\rangle = 0$.

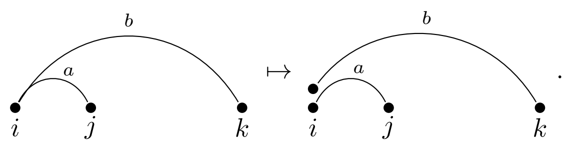
4 Tensor products and general restriction coefficients

Theorem 3.1 in fact is sufficiently strong that analogous statements can be made for the coefficients of arbitrary supercharacters in the decomposition of tensor products and restrictions. This section begins by developing the appropriate generalization to the graph $\Gamma_K(\lambda)$. We then state Theorem 4.1 for tensor products and Theorem 4.3 for restrictions. Along the way, Theorem 4.2 describes how characters corresponding to multisets are the same as restrictions from certain set partitions (up to a scalar multiple).

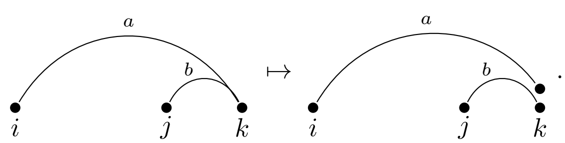
4.1 A generalized bipartite graph

Given $\lambda \in \mathcal{M}(q)$, perturb the arcs such that they stack on top of one-another in the following fashion.

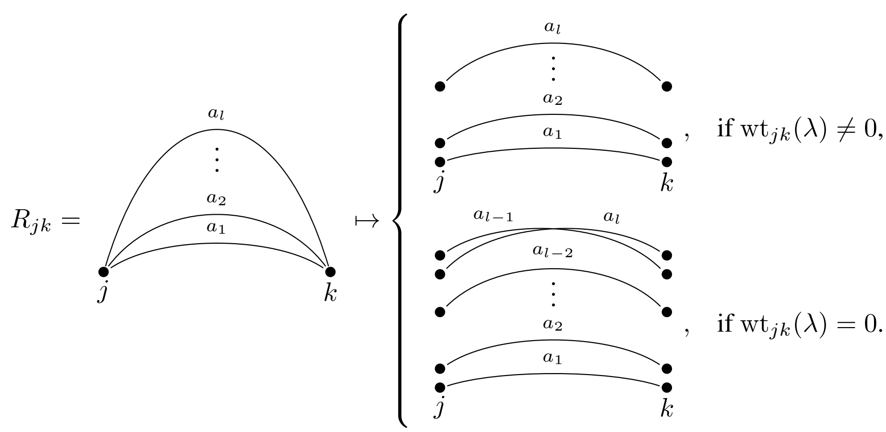
(TL) If $i \overset{a}{\frown} j, i \overset{b}{\frown} k \in \lambda$ with $j < k$, then the left endpoint of $i \overset{b}{\frown} k$ is above the left endpoint of $i \overset{a}{\frown} j$,



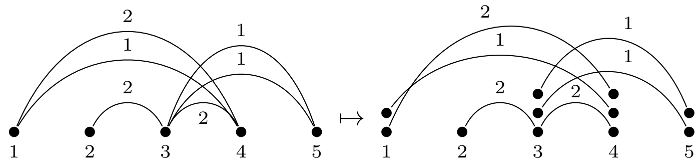
(TR) If $i \overset{a}{\frown} k, j \overset{b}{\frown} k \in \lambda$ with $i < j$, then the right endpoint of $i \overset{a}{\frown} k$ is above the right endpoint of $j \overset{b}{\frown} k$,



(TB) If $|R_{jk}| > 1$, then



Example 3 *For $q = 3$,*

$\lambda =$ 

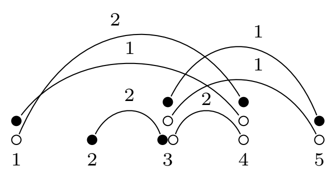
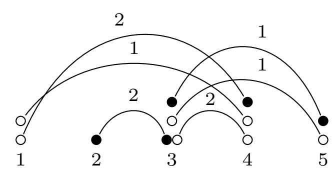
Define a labeling function $\Lambda_K : \lambda \to \{(\bullet,\bullet),(\circ,\bullet),(\bullet,\circ),(\circ,\circ)\}$, given by

$$\Lambda_K(j \overset{b}{\frown} k) = (\Lambda_K^L(j \overset{b}{\frown} k), \Lambda_K^R(j \overset{b}{\frown} k)),$$

where

$$\begin{aligned} \Lambda_K^L(j \overset{b}{\frown} k) &= \begin{cases} \circ & \text{if } j \notin K \text{ or } j \overset{b}{\frown} k \text{ starts below an arc starting at } j, \\ \bullet, & \text{otherwise.} \end{cases} \\ \Lambda_K^R(j \overset{b}{\frown} k) &= \begin{cases} \circ & \text{if } k \notin K \text{ or } j \overset{b}{\frown} k \text{ ends below an arc ending at } k, \\ \bullet, & \text{otherwise.} \end{cases} \end{aligned} \tag{6}$$

In the above example,

$\Lambda_{\{1,2,3,4,5\}}(\lambda) =$  and $\Lambda_{\{2,3,4,5\}}(\lambda) =$  ,

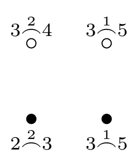
where we replace the endpoints of the arcs by their images under Λ_K.

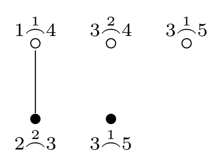
Construct a bipartite graph $\Gamma_K(\lambda)$ given by vertices

$$\begin{aligned} V_\bullet &= \{j \overset{a}{\frown} k \in \lambda \mid \Lambda_K(j \overset{a}{\frown} k) = (\bullet,\bullet)\} \\ V_\circ &= \{j \overset{a}{\frown} k \in \lambda \mid \Lambda_K(j \overset{a}{\frown} k) = (\circ,\circ)\}, \end{aligned}$$

and an edge from $i \overset{a}{\frown} l \in V_\circ$ to $j \overset{b}{\frown} k \in V_\bullet$ if $i < j < k < l$.

In our example,

$\Gamma_{\{1,2,3,4,5\}}(\lambda) =$ 

and $\Gamma_{\{2,3,4,5\}}(\lambda) =$ 

.

The main theorem of this section follows, and its proof can be found in Section 4.2.

Theorem 4.1 *Suppose $\lambda, \mu, \nu \in \mathcal{S}_K(q)$ with $K \subseteq \mathbb{Z}_{\geq 1}$ a finite subset. Then*

$$\langle \chi^\lambda \otimes \chi^\mu, \chi^\nu \rangle \neq 0$$

if and only if $\Gamma_K(\lambda \cup \mu \cup \bar{\nu})$ has a complete matching from $V_\bullet$ to $V_\circ$.

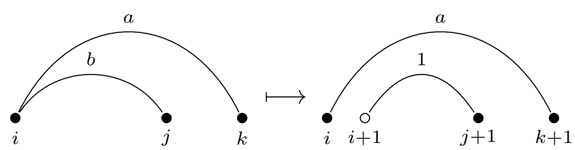
4.2 Straightening rules

Given $\lambda \in \mathcal{M}(q)$, this section describes "straightening" rules that allow us to create a sequence

$$\lambda = \lambda^{(0)}, \lambda^{(1)}, \cdots, \lambda^{(\ell)},$$

where at each stage we remove a conflict of type (CL), (CR), or (CB) until we arrive at $\lambda^{(\ell)} \in \mathcal{S}(q)$. Furthermore, there is an underlying sequence of pairs $(K^{(0)}, L^{(0)}), (K^{(1)}, L^{(1)}), \ldots, (K^{(\ell)}, L^{(\ell)})$ of finite subsets such that $|K^{(i)}| = |K|$ and $\lambda^{(i)} \in \mathcal{M}_{K^{(i)} \cup L^{(i)}}(q)$. While the order in which one applies the straightening rules does matter in terms of which set partition one obtains, for our purposes in this paper (Theorem 4.2, below) the differences are irrelevant. The rules are as follows.

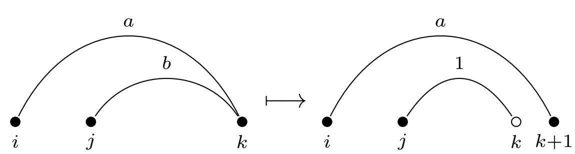
For $a, b \in \mathbb{F}_q^\times$, in moving from $\lambda^{(m-1)}$ to $\lambda^{(m)}$ we can



(SL)

with

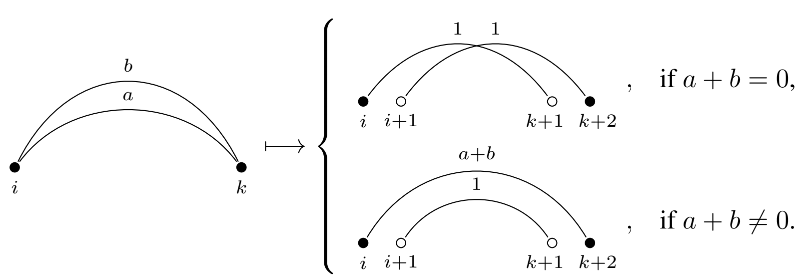
$$\begin{aligned} K^{(m)} &= ([1, i] \cap K^{(m-1)}) \cup ((\mathbb{Z}_{\geq i+1} \cap K^{(m-1)}) + 1) \\ L^{(m)} &= ([1, i] \cap L^{(m-1)}) \cup \{i+1\} \cup ((\mathbb{Z}_{\geq i+1} \cap L^{(m-1)}) + 1); \end{aligned}$$



(SR)

with

$$\begin{aligned} K^{(m)} &= ([1, k-1] \cap K^{(m-1)}) \cup ((\mathbb{Z}_{\geq k} \cap K^{(m-1)}) + 1), \\ L^{(m)} &= ([1, k-1] \cap L^{(m-1)}) \cup \{k\} \cup ((\mathbb{Z}_{\geq k} \cap L^{(m-1)}) + 1); \end{aligned}$$



(SB)

with

$$K^{(m)} = ([1,i] \cap K^{(m-1)}) \cup (([i+1,k-1] \cap K^{(m-1)}) + 1) \cup ((\mathbb{Z}_{\geq k} \cap K^{(m-1)}) + 2)$$
$$L^{(m)} = ([1,i] \cap L^{(m-1)}) \cup \{i+1\}(([i+1,k-1] \cap L^{(m-1)}) + 1) \cup \{k+1\} \cup ((\mathbb{Z}_{\geq k} \cap L^{(m-1)}) + 2).$$

In each case there are "new nodes" indicated by $\circ$ that push all the other node values to the right up (note, we view the $\bullet$-nodes as being the same though their number labels may change). In fact, $K^{(m)}$ is the set of original nodes (up to being pushed around) and $L^{(m)}$ is the set of nodes that were at some point $\circ$-nodes (see example after Theorem 4.2).

The following lemma states that these rules (SL), (SR) and (SB) do not fundamentally change the underlying character.

Lemma 4.1 *Let $\lambda \in \mathcal{M}_K(q)$ and apply (SL), (SR) or (SB) to obtain $\tilde{\lambda} = \lambda^{(1)} \in \mathcal{M}_{K^{(1)} \cup L^{(1)}}(q)$. Then as a character of $U_K \cong U_{K^{(1)}}$,*

$$\chi^\lambda = q^{-r_{K^{(1)}}^{K^{(1)} \cup L^{(1)}}(\tilde{\lambda})} \mathrm{Res}^{U_{K^{(1)} \cup L^{(1)}}}_{U_{K^{(1)}}}(\chi^{\tilde{\lambda}}).$$

By iterating Lemma 4.1 to remove all the conflicts of a multiset, we see that up to shifting of indices every tensor product is the same (up to a scalar multiple) as restriction from some supercharacter.

Theorem 4.2 *Let $\lambda \in \mathcal{M}_K(q)$. Then there exists $\tilde{\lambda} \in \mathcal{S}_{K' \cup L'}(q)$ with $|K| = |K'|$, such that*

$$\chi^\lambda = q^{-r_{K'}^{K' \cup L'}(\tilde{\lambda})} \mathrm{Res}^{U_{K' \cup L'}}_{U_{K'}}(\chi^{\tilde{\lambda}}).$$

Remark 3 *Note that if we assume there are no (CN) conflicts, or $\lambda \in \mathcal{M}_K(q)$, then we may simplify the definition of the labeling function Λ_K as follows. If $\lambda \in \mathcal{M}_K(q)$ and $\tilde{\lambda} \in \mathcal{S}_{K^{(\ell)} \cup L^{(\ell)}}(q)$ is obtained by applying (SL), (SR) and (SB), then*

$$\Lambda_K(j \overset{b}{\frown} k) = (\Lambda_K^L(j \overset{b}{\frown} k), \Lambda_K^R(j \overset{b}{\frown} k)),$$

where

$$\Lambda_K^L(j \overset{b}{\frown} k) = \begin{cases} \circ & \text{if } j \in L^{(\ell)}, \\ \bullet, & \text{otherwise.} \end{cases}$$
$$\Lambda_K^R(j \overset{b}{\frown} k) = \begin{cases} \circ & \text{if } k \in L^{(\ell)}, \\ \bullet, & \text{otherwise.} \end{cases}$$

Theorem 4.2 also allows us to extend Theorem 3.1 to the coefficient of arbitrary supercharacters.

Theorem 4.3 *Suppose $\lambda \in \mathcal{S}_L(q)$ and $\mu \in \mathcal{S}_K(q)$ with $K \subseteq L \subseteq \mathbb{Z}_{\geq 1}$ finite sets. Then*

$$\langle \mathrm{Res}^{U_L}_{U_K}(\chi^\lambda), \chi^\mu \rangle \neq 0$$

if and only if $\Gamma_K(\lambda \cup \bar{\mu})$ has a complete matching from $V_\bullet$ to $V_\circ$.

References

[1] André, C. Basic characters of the unitriangular group, *J. Algebra* **175** (1995), 287–319.

[2] André, C. Irreducible characters of finite algebra groups, *Matrices and group representations Coimbra, 1998* Textos Mat. Sér B **19** (1999), 65–80.

[3] André, C. The basic character table of the unitriangular group, *J. Algebra* **241** (2001), 437–471.

[4] André, C. Basic characters of the unitriangular group (for arbitrary primes), *Proc. Amer. Math. Soc.* **130** (2002), 1934–1954.

[5] André, C; Neto, A. Super-characters of finite unipotent groups of types B_n, C_n and D_n, *J. Algebra* **305** (2006), 394–429.

[6] André, C; Neto, A. Supercharacters of the Sylow p-subgroups of the finite symplectic and orthogonal groups, *Pacific J. Math.* **239** (2009), 201–230.

[7] Arias-Castro, E; Diaconis, P; Stanley, R. A super-class walk on upper-triangular matrices, *J. Algebra* **278** (2004), 739–765.

[8] Diaconis, P; Isaacs, M. Supercharacters and superclasses for algebra groups, *Trans. Amer. Math. Soc.* **360** (2008), 2359–2392.

[9] Garsia, A; Remmel, J. Q-counting rook configurations and a formula of Frobenius, *J. Combin. Theory Ser. A* **41** (1986), 246–275.

[10] Halverson, T; Thiem, N. q-Partition algebra combinatorics. To appear in *J. Combin. Theory Ser. A*.

[11] Hendrickson, A. *Supercharacter theories of finite cyclic groups*. Unpublished Ph.D. Thesis, Department of Mathematics, University of Wisconsin, 2008.

[12] Higman, G. Enumerating p-groups I: Inequalities, *Proc. London Math. Soc.* **10** (1960), 24–30.

[13] Lewis, S. Restriction of Supercharacters in $U_n(\mathbb{F}_2)$. Honors thesis (2009), University of Colorado at Boulder.

[14] Marberg, E; Thiem, N. Superinduction for pattern groups, *J. Algebra* **321** (2009), 3681–3703.

[15] Otto, B. Two bounds for the nilpotence class of an algebra, 2009 preprint.

[16] Thiem, N. Branching rules in the ring of superclass functions of unipotent upper-triangular matrices. To appear in *J. Algebraic Combin.*

[17] Thiem, N; Venkateswaran, V. Restricting supercharacters of the finite group of unipotent uppertriangular matrices, *Electron. J. Combin.* **16**(1) Research Paper 23 (2009), 32 pages.

[18] Yan, N. *Representation theory of the finite unipotent linear groups*, Unpublished Ph.D. Thesis, Department of Mathematics, University of Pennsylvania, 2001.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 778–789

Equivalence Relations of Permutations Generated by Constrained Transpositions

Stephen Linton[1†] and James Propp[2‡] and Tom Roby[3] and Julian West[4§]

[1] *Centre for Interdisciplinary Research in Computational Algebra, University of St Andrews, St Andrews, Fife KY16 9SX, Scotland*
[2] *University of Massachusetts, Lowell, MA 01854, USA*
[3] *University of Connecticut, Storrs, CT 06269, USA*
[4] *University of Bristol, Bristol BS8 1TW, England*

Abstract. We consider a large family of equivalence relations on permutations in S_n that generalise those discovered by Knuth in his study of the Robinson-Schensted correspondence. In our most general setting, two permutations are equivalent if one can be obtained from the other by a sequence of pattern-replacing moves of prescribed form; however, we limit our focus to patterns where two elements are transposed, conditional upon the presence of a third element of suitable value and location. For some relations of this type, we compute the number of equivalence classes, determine how many n-permutations are equivalent to the identity permutation, or characterise this equivalence class. Although our results include familiar integer sequences (e.g., Catalan, Fibonacci, and Tribonacci numbers) and special classes of permutations (layered, connected, and 123-avoiding), some of the sequences that arise appear to be new.

Résumé. Nous considérons une famille de relations d'equivalence sur l'ensemble S_n des permutations, qui généralisent les relations de Knuth liées à la correspondance Robinson-Schensted. Dans notre contexte général, deux permutations sont considérées comme équivalentes si l'une peut être obtenue de l'autre auprès d'une séquence de remplacements d'un motif par un autre selon des règles précisées. Désormais, nous ne considérons dans l'oeuvre actuelle que les motifs qui correspondent à la transposition de deux éléments, conditioné sur la presence d'un élément de valeur et de position approprié. Pour plusieurs exemples de ce problème, nous énumérons les classes d'équivalence, nous déterminons combien de permutations sur n éléments sont équivalentes à l'identité, ou nous précisons la forme des éléments dans cette dernière classe. Bien que nos résultats retrouvent des séquences des entiers très bien connues (nombres de Catalan, de Fibonacci, de Tribonacci . . .) ainsi que des classes de permutations déjà étudiées (en couches, connexes, sans motif 123), nous trouvons également des séquences qui paraissent être nouvelles.

Keywords: permutation patterns, equivalence classes, Knuth relations, integer sequences, Catalan numbers, layered permutations, connected permutations, pattern-avoiding permutations

---

[†]Supported by EPSRC grant EP/C523229/01.
[‡]Supported by NSF
[§]Supported by NSERC operating grant OGP0105492.

1 Introduction and motivation

We consider a family of equivalence relations on permutations in S_n, in which two permutations are considered to be equivalent if one can be converted into the other by replacing a short subsequence of elements by the same elements permuted in a specific fashion, or (extending by transitivity) by a sequence of such moves. These generalise the relations discovered by Knuth in his study of the Robinson-Schensted correspondence, though our original motivations were unrelated. We begin the systematic study of such equivalence relations, connecting them with integer sequences both familiar and (apparently) new.

As a simple first example, the permutation 12**345**6 can be converted to 12**543**6 by replacing the subsequence **345** by **543**. The same permutation 1**2**34**56** can be also be converted into 1**6**34**52**. In each of these examples, the subsequence removed is of *pattern* 123 (reducing the elements to form a permutation whose elements are in the same relative order) and the subsequence replacing it is of pattern 321.

We could therefore say that 123456 and 125436 are equivalent under the replacement $123 \to 321$. Since we want all our replacement rules to be bi-directional, we will actually say that these permutations are equivalent under $123 \leftrightarrow 321$, or, using set notation, under $\{123, 321\}$. Since 163452 and 123456 are equivalent under the same replacement, by transitivity we also have that 163452 and 125436 are equivalent under $\{123, 321\}$.

Interesting enumerative questions arise when the elements being replaced are allowed to be in general position (Section 2), but also when the replacements are further constrained to affect only adjacent elements as in the very first example above (Section 3), or even when constrained to affect only blocks of consecutive elements representing a run of consecutive values, again as in the first example (Section 4).

We may also wish to allow more than one type of (bi-directional) replacement, such as both $123 \leftrightarrow 321$ and $123 \leftrightarrow 132$. If the intersection of these sets is nonempty, the new relation corresponds simply to a union of the two sets: $\{123, 132, 321\} = \{123, 321\} \cup \{123, 132\}$. However, the Knuth relations (described below) require two disjoint types of replacements.

Let $\pi \in S_n$, and let $P = \{B_1, B_2, \ldots, B_t\}$ be a (set) partition of S_k, where $k \leq n$. Each block B_l of P represents a list of k-length patterns which can replace one another. We call two permutations P°-equivalent if one can be obtained from the other by a sequence of replacements, each replacing a σ_i-pattern subsequence with the same elements in the pattern σ_j, where σ_i and σ_j lie in the same block B_l of P. Then $\mathrm{Eq}^{\circ}(\pi, P)$ will denote the set of permutations equivalent to π under P°-equivalence. Thus $163452 \in \mathrm{Eq}^{\circ}\left(123456, \left\{\{123, 321\}\right\}\right)$.

Similarly, we will use $P^{\shortparallel}$ to denote the equivalence relation and $\mathrm{Eq}^{\shortparallel}(\pi, P)$ for the equivalence class of π under replacement within P only of adjacent elements, e.g. $125436 \in \mathrm{Eq}^{\shortparallel}\left(123456, \left\{\{123, 321\}\right\}\right)$. And we will use $\mathrm{Eq}^{\square}(\pi, P)$ for the case where both positions and values are constrained, e.g. $125436 \in \mathrm{Eq}^{\square}\left(123456, \left\{\{123, 321\}\right\}\right)$. To refer to such classes generally we use the notation $\mathrm{Eq}^{\star}(\pi, P)$. The set of distinct equivalence classes into which S_n splits under an equivalence $P^{\star}$ will be denoted $\mathrm{Classes}^{\star}(n, P)$.

The present paper begins the study of these equivalence relations by considering three types of question:

(A) Compute the number of equivalence classes, $\#\mathrm{Classes}^{\star}(n, P)$, into which S_n is partitioned.

(B) Compute the size, $\#\mathrm{Eq}^{\star}(\iota_n, P)$, of the equivalence class containing the identity $\iota_n = 123 \cdots n$.

(C) (More generally) characterise the set $\mathrm{Eq}^{\star}(\iota_n, P)$ of permutations equivalent to the identity.

Although the framework above allows for much greater generality, in this paper we will restrict our attention to replacements by patterns of length $k = 3$, and usually to replacement patterns built up from pairs in which one permutation is the identity, and the other is a transposition (i.e., fixes one of the elements). Omitting some cases by symmetry, we have the following possible partitions of S_3, where (as usual) we omit singleton blocks:

$$P_1 = \big\{\{123, 132\}\big\},$$
$$P_2 = \big\{\{123, 213\}\big\},$$
$$P_4 = \big\{\{123, 321\}\big\}.$$

We will also consider applying two of these replacement operations simultaneously, and we will number the appropriate partitions as

$$P_3 = \big\{\{123, 132, 213\}\big\},$$
$$P_5 = \big\{\{123, 132, 321\}\big\},$$
$$P_6 = \big\{\{123, 213, 321\}\big\},$$

following the convention $P_{i+j} := P_i \vee P_j$, the *join* of these two partitions [EC1, ch.3]. Indeed we can allow all three replacements: $P_7 = \big\{\{123, 132, 213, 321\}\big\}$. (In fact, the cases P_1 and P_2 are equivalent by symmetry, as are P_5 and P_6. We list P_1 and P_2 separately so as to be able to consider their join.)

Our motivation for focussing attention on pairs of this form is that we can then think of an operation, not in terms of replacing one pattern by another, but simply in terms of *swapping* two elements within the pattern, with the third serving as a witness enabling the swap.

By far the best-known example of constrained swapping in permutations is certainly the *Knuth Relations* [Knu70], which allow the swap of adjacent entries provided an intermediate value lies immediately to the right or left. In the notation of this paper, they correspond to $P_K^{\shortparallel} = \big\{\{213, 231\}, \{132, 312\}\big\}$. Permutations equivalent under this relation map to the same first coordinate (P-tableau) under the Robinson-Schensted correspondence.

Mark Haiman introduced the notion of *dual equivalence* of permutations: π and τ are dual equivalent if one can be obtained from the other by swaps of adjacent *values* from the above P_K, i.e., if their inverses are Knuth-equivalent, or if they map to the same second coordinate (Q-tableau) under the Robinson-Schensted correspondence [Hai92]. For the enumerative problems in this paper, we get the same answers for Knuth and dual equivalence.

In her dissertation [SA07] Sami Assaf constructed graphs (with some extra structure) whose vertices are tableaux of a fixed shape (which may be viewed as permutations via their "reading words"), and whose edges represent (elementary) dual equivalences between vertices. For this particular relation (equivalently for the Knuth relations), she was able to characterise the local structure of these graphs, which she later used to give a combinatorial formula for the Schur expansion of LLT polynomials and MacDonald Polynomials. She also used these, along with crystal graphs, to give a combinatorial realization of Schur-Weyl duality [SA08].

S. Fomin has a very clear elementary exposition of how Knuth and dual equivalence are related to the Robinson-Schensted correspondence, Schützenberger's *jeu de taquin*, and the Littlewood-Richardson rule in [EC2, Ch. 7, App. 1]. For the problems considered above, the answers for $P_K^{\shortparallel}$ are well known to be: (A) the number of involutions in S_n; (B) 1; and (C) $\{id\}$. In fact one can compute $\#\mathrm{Eq}^{\shortparallel}(\pi, P_K)$ for any permutation π by using the Frame-Robinson-Thrall hook-length formula to compute the number of standard Young tableaux of the *shape* output by the R-S correspondence applied to π.

Given that the Knuth relations act on adjacent elements, and lead to some deep combinatorial results, it is perhaps not surprising that the most interesting problems and proofs in this paper are to be found in Section 3. A summary of our numbers and results is given in Figure 1.

Fig. 1: Summary of Results

These tables give numerical values and names (when available) of the sequences associated with our two main enumerative questions. All sequences begin with the value for $n = 3$. Results proven in this paper have a grey background; for other cases we lack even conjectural formulae. Six-digit codes preceded by "A" cite specific sequences in Sloane [OEIS].

Number of classes		§ 1 neither ("classical")	§ 2 only indices adjacent	§ 3 indices and values adjacent
(1)	$123 \leftrightarrow 132$	[5, 14, 42, 132, 429] Catalan	[5, 16, 62, 284, 1507, 9104]	[5, 20, 102, 626, 4458, 36144]
(2)	$123 \leftrightarrow 213$			
(4)	$123 \leftrightarrow 321$	[5, 10, 3, 1, 1, 1] trivial	[5, 16, 60, 260, 1260, 67442]	[5, 20, 102, 626, 4458, 36144]
(3)	$123 \leftrightarrow 132 \leftrightarrow 213$	[4, 8, 16, 32, 64, 128] powers of 2	[4, 10, 26, 76, 232, 764] involutions	[4, 17, 89, 556, 4011, 32843]
(5)	$123 \leftrightarrow 132 \leftrightarrow 321$	[4, 2, 1, 1, 1, 1] trivial	[4, 8, 14, 27, 68, 159, 496]	[4, 16, 84, 536, 3912, 32256]
(6)	$123 \leftrightarrow 213 \leftrightarrow 321$			
(7)	$123 \leftrightarrow 132$ $\leftrightarrow 213 \leftrightarrow 321$	[3, 2, 1, 1, 1, 1] trivial	[3, 4, 5, 8, 11, 20, 29, 57]	[3, 13, 71, 470, 3497]

Size of class with ι_n		§ 1 neither ("classical")	§ 2 only indices adjacent	§ 3 indices and values adjacent
(1)	$123 \leftrightarrow 132$	[2, 6, 24, 120, 720] (n-1)!	[2, 4, 12, 36, 144, 576, 2880] product of two factorials	[2, 3, 5, 8, 13, 21, 34, 55] Fibonacci numbers
(2)	$123 \leftrightarrow 213$			
(4)	$123 \leftrightarrow 321$	[2, 4, 24, 720] trivial	[2, 3, 6, 10, 20, 35, 70, 126] central binomial coefficients	[2, 3, 4, 6, 9, 13, 19, 28] A000930
(3)	$123 \leftrightarrow 132 \leftrightarrow 213$	[3, 13, 71, 461] connected A003319	[3, 7, 35, 135, 945, 5193] terms are always odd	[3, 4, 8, 12, 21, 33, 55, 88] A052952
(5)	$123 \leftrightarrow 132 \leftrightarrow 321$	[3, 23, 120, 720] trivial	[3, 9, 54, 285, 2160, 15825] separate formulae for odd/even	[3, 5, 9, 17, 31, 57, 105, 193] tribonacci numbers A000213
(6)	$123 \leftrightarrow 213 \leftrightarrow 321$			
(7)	$123 \leftrightarrow 132$ $\leftrightarrow 213 \leftrightarrow 321$	[3, 23, 120, 720] trivial	[4, 21, 116, 713, 5030]	[4, 6, 13, 23, 44, 80, 149, 273] tribonacci A000073 $-[n$ even$]$

If $\tau \in \mathrm{Eq}^\star(\pi, P)$ we will say that τ is *reachable* from π (under P). If $\mathrm{Eq}^\star(\iota_n, P) = S_n$, then every permutation in S_n is reachable from every other, and we will say that S_n is *connected* by P. If $\mathrm{Eq}^\star(\pi, P) = \{\pi\}$ we will say that π is *isolated* (under P).

It is obvious that if P_i refines P_j as partitions of S_k (i.e., $P_i \le P_j$ in the lattice of partitions of S_k), then the partition of S_n induced by P_i refines the one induced by P_j, because a permutation reachable from π under P_i is also reachable under P_j. This enables the following simple observations:

Proposition 1 *If P_i refines P_j (as partitions of S_k), then for all $\pi \in S_n$ with $n \ge k$*

$$\begin{aligned} \mathrm{Eq}^\star(\pi, P_i) &\subseteq \mathrm{Eq}^\star(\pi, P_j) \\ \#\mathrm{Eq}^\star(\pi, P_i) &\le \#\mathrm{Eq}^\star(\pi, P_j) \\ \#\mathrm{Classes}^\star(n, P_i) &\ge \#\mathrm{Classes}(n, P_j) \end{aligned}$$

2 General pattern equivalence

Many of the equivalence relations in this section are trivial, and follow immediately from the following observation. The others lead to familar combinatorial numbers and objects.

Proposition 2 *Let P be any partition of S_k for $2 \leq k \leq n-1$. If $\#\mathrm{Classes}^\circ(n-1,P) = 1$, then $\#\mathrm{Classes}^\circ(n,P) = 1$.*

Proof: We will show that any $\pi \in S_n$ can be reached from the identity, under the supposition that any two permutations in S_{n-1} are equivalent. If $\pi(1) \neq n$, simply apply the supposition to the elements $1 \dots n-1$ to obtain any permutation beginning with $\pi(1)$. Then apply the supposition to the elements now occupying positions $2 \dots n$ to complete the construction of π.

If $\pi(1) = n$, it is necessary to add an additional step at the beginning in order to detach 1 from the tail of the permutation and move it into (say) position $n-1$. But we know we can do this by applying the supposition to the elements in positions $2 \dots n$. And as long as $n-1 \geq 2$, position $n-1$ is among the positions $2 \dots n$. □

The following results follow.

Proposition 3 $\#\mathrm{Classes}^\circ\left(n, \left\{\{123,132,321\}\right\}\right) = 1$ *and* $\#\mathrm{Classes}^\circ\left(n, \left\{\{123,132,213,321\}\right\}\right) = 1$ *for $n \geq 5$; and* $\#\mathrm{Classes}^\circ\left(n, \left\{\{123,321\}\right\}\right) = 1$ *for $n \geq 6$.*

Proof: It is easy to verify by hand, or by computer, that all permutations in S_5 are reachable from 12345 by moves in $P_5 = \left\{\{123,132,321\}\right\}$. (Indeed, all permutations in S_4 are reachable from 1234 except for 3412, which is isolated.) As S_5 is connected, it follows (by induction) from the preceding proposition that S_n is connected for all $n \geq 5$. Since $P_7 \geq P_5$, Proposition 1 tells us that S_n is connected under $P_7 = \left\{\{123,132,213,321\}\right\}$ whenever it is connected under P_5. (In S_4, the permutation 3412 remains isolated.) Finally, we can check by computer that under $P_4 = \left\{\{123,321\}\right\}$ S_6 is connected; whence, S_n is connected for $n \geq 6$. □

We remark that under P_4, S_4 splits into 10 equivalence classes, and S_5 into three classes. The class containing 12345 contains 24 elements. This suggests a possible bar bet. Hand your mark six cards numbered 1 through 6 and invite him or her to lay them out in any sequence. By applying moves of the form $123 \leftrightarrow 321$ ("interchange two cards if and only if an intermediate (value) card intervenes") you will always be able to put the cards in order (although it may take some practice to become efficient at this!). Now "go easy" on your mark by reducing the number of cards to 5. Even from a random sequence, the mark has only one chance in five of being able to reach the identity.

Of course from Proposition 3 it immediately follows that:

Corollary 4 $\#\mathrm{Eq}^\circ\left(\iota_n, \left\{\{123,132,321\}\right\}\right) = n!$ *for $n \geq 5$,* $\#\mathrm{Eq}^\circ\left(\iota_n, \left\{\{123,132,213,321\}\right\}\right) = n!$ *for $n \geq 5$, and* $\#\mathrm{Eq}^\circ\left(\iota_n, \left\{\{123,321\}\right\}\right) = n!$ *for $n \geq 6$.*

Proposition 5 $\#\mathrm{Eq}^\circ\left(\iota_n, \left\{\{123,213\}\right\}\right) = (n-1)!$ *for $n \geq 2$.*

Proof: Obviously the largest element n cannot be moved away from the end of the permutation. Equally obviously the n, remaining at the far right, facilitates the free permutation of all other elements. □

Proposition 6 *For $n \geq 1$,* $\#\mathrm{Classes}^\circ\left(n, \left\{\{123,213\}\right\}\right) = c_n = \frac{2n!}{n!(n+1)!}$, *the nth Catalan number.*

Proof: (Sketch) If $i < j < k$, and $p(i) < p(j) < p(k)$, then $p(k)$ facilitates the swapping of $p(i)$ and $p(j)$ to arrive at a permutation with a strictly larger number of inversions.

Thus the "largest" (by number of inversions) elements in each equivalence class are exactly the 123-avoiding permutations, of which there are c_n [Bon06, ch. 14] or [Bon04, Sec. 4.2]. Likewise the "smallest" elements are the 213-avoiding permutations. □

The next two propositions study a class equivalent under symmetry (complementation) to $\mathrm{Eq}^\circ(\iota_n, P_3)$. The first references the indecomposable permutations [OEIS, A003319] or [Bon04, p. 145], and the second the layered permutations introduced by W. Stromquist [Stro93], and studied carefully by A. Price in his thesis [Pri97].

Proposition 7 *Let ρ_n denote the permutation $n, n-1, \ldots, 1$. Then $\mathrm{Eq}^\circ\left(\rho_n, \{\{321, 312, 231\}\}\right)$ is the set of indecomposable permutations.*

Proof: When viewed as a matrix, any permutation decomposes into irreducible blocks along the main diagonal. The identity ι_n decomposes into n singleton blocks, while ρ_n is indecomposable and is one large block.

First note that if a transformation $(a_1, a_2, a_3) \to (b_1, b_2, b_3)$ applied within a block causes it to split into more than one block, then b_1 must be in the leftmost/lowest of the new blocks, and b_3 in the rightmost/highest. Therefore b_1 must be less than b_3, which is exactly what doesn't happen with any of our possible transformations, because the first element is larger than the third in each of 321, 312 and 231. Thus, in particular, if we start with an indecomposable permutation such as ρ_n, successive applications of the permitted operations will always produce indecomposable permutations.

Next, we have to show that all indecomposable permutations are in fact reachable from ρ_n. Remembering that our replacement operations are all reversible, we will instead show that we can always return to ρ_n from an arbitrary indecomposable permutation.

Take $n \geq 3$, and let $\tau = t_1, t_2, \ldots, t_n$ be an arbitrary indecomposable permutation other than ρ_n. We will show that τ always contains at least one of 312 or 231.

If we are not in ρ_n then somewhere there is $i < j$ such that $t_i < t_j$. It can't be the case that $t_i > t_{i+1} > \ldots > t_{j-1} > t_j$, so somewhere there is a consecutive rise, say t_k, t_{k+1}. Now if any element to the right of t_{k+1} is less than t_k we have a 231, so assume there are none such. Similarly, assume there is no element to the left of t_k and greater than t_{k+1}.

But there must be some x to the left of t_k which is greater than some y to the right of t_{k+1}, or otherwise the permutation decomposes between t_k and t_{k+1}. These four elements x, t_k, t_{k+1}, y form a 3142, which contains both a 312 (x, t_k, y) and a 231 (x, t_{k+1}, y).

Having now located a 312 or 231, we can then apply either $312 \to 321$ or $231 \to 321$, as appropriate. Each of these operations simply switches a pair of elements, and (as we have seen in the proof of Proposition 6) strictly increases the number of inversions, progressing us toward ρ_n. This completes the proof that all indecomposable permutations are reachable, and therefore the proof that the reachable permutations are exactly the indecomposable permutations. □

Proposition 8 $\#\mathrm{Classes}^\circ\left(n, \{\{321, 312, 231\}\}\right) = 2^{n-1}$ *for* $n \geq 1$.

Proof: This follows fairly easily from comments made in the proof of the previous proposition. We found that the equivalence class of the anti-identity consisted of the indecomposable permutations.

Any permutation decomposes as a direct sum of irreducible blocks. We found that in our case, our operations cannot cause a block to split. Therefore they also cannot cause blocks to join up, because then the operation could be reversed, splitting the blocks. So the block composition is preserved.

By the arguments already given, we can work within any indecomposable block to restore it to an anti-identity. Therefore each equivalence class consists of all the permutations with a given block structure under direct sum of indecomposables.

In particular, each equivalence class contains exactly one permutation which is a direct sum of anti-identities. These are exactly the layered permutations, and there are clearly 2^{n-1} of them, with a factor of 2 according to whether each consecutive pair of elements is or is not in the same layer. □

Finally we apply the reversal involution on S_n to the above result to get our result for the partition P_3.

Theorem 9 $\#\mathrm{Classes}^{\circ}\left(n,\left\{\{123,132,213\}\right\}\right)=2^{n-1}$ *for* $n\geq 1$.

3 Adjacent transformations

As mentioned in the introduction, this section contains our most interesting results and proofs. The first rediscovers sequence A010551 from Sloane [OEIS].

Theorem 10 $\#\mathrm{Eq}^{\shortparallel}\left(\iota_n,\left\{\{123,213\}\right\}\right)=\lfloor n/2\rfloor!\lceil n/2\rceil!$ *for* $n\geq 1$.

Proof: The largest element, n, never comes unglued from the end, because there is nothing to enable it. And therefore $n-1$ must stay somewhere in the last three positions (as only n can enable its movement), and $n-2$ somewhere in the last five, and so on; such restrictions apply to $\lfloor n/2\rfloor$ of the elements. This limits the number of potentially reachable elements to $\lfloor n/2\rfloor!\lceil n/2\rceil!$: placing the elements from largest to smallest, one has a choice of $1,2,3,\ldots,\lfloor n/2\rfloor,\lceil n/2\rceil,\ldots,3,2,1$ places to put each element.

Next we will show that all permutations conforming to these restrictions are indeed reachable. We will do this in two stages. In Stage 1 we advance each of the large, constrained elements as far left as it can go In Stage 2 we construct the target permutation from left to right, two elements at a time.

Stage 1: The elements $\lfloor n/2\rfloor,\ldots,n-1,n$ are to be positioned. First move $\lfloor n/2\rfloor$ one step left, using a move of type $123\rightarrow 213$, in which $\lfloor n/2\rfloor+1$ plays the role of the facilitating "3". In just the same way, more the element $\lfloor n/2\rfloor+1$ to the left, continuing until the entire block $\lfloor n/2\rfloor,\ldots,n-1$ has been shifted one to the left. The element $n-1$ has now reached its leftmost permitted position, and will remain in place as we now move the block $\lfloor n/2\rfloor,\ldots,n-2$. This moves $n-2$ as far left as it will go, and we now move the next smaller block, etc. As promised, this places each constrained element as far left as possible. These elements will now serve as a "skeleton" enabling the construction of the target permutation.

Stage 2: The key observation making this stage possible is that the small, unconstrained elements can be freely moved about, leaving the large elements in the skeleton fixed. This is because, if $\{a,b\}<X<Y$, we can always execute the following sequence of moves: $aXbY\rightarrow abXY\rightarrow baXY\rightarrow bXaY$. Also in the case where n is odd, the leftmost element in the skeleton is in position 3, and the two small elements in positions 1 and 2 can be interchanged if desired.

Now we examine the target permutation and move the required element(s) into the first position (if n is even), or the first two positions (if n is odd). At this point, the elements occupying the next two positions are reclassified as small, so that the skeleton terminates two positions further to the right, and we continue by placing and ordering the next pair of elements. By continuing two elements at a time, we can build the entire target permutation. □

Theorem 11 *(a)* $\#\mathrm{Eq}''\left(\iota_n, \left\{\{123, 132, 321\}\right\}\right) = \frac{3}{2}(k)(k+1)(2k-1)!$*, for* $n = 2k+1$ *odd and* $n \geq 3$*.*
(b)$\#\mathrm{Eq}''\left(\iota_n, \left\{\{123, 132, 321\}\right\}\right) = \frac{3}{2}(k)(k-\frac{1}{3})(2k-2)! - (2k-3)!!$*, for* $n = 2k$ *even and* $n \geq 2$*.*
(Here $(2k-3)!! = (2k-3)\dots(3)(1)$*, the product of odd natural numbers less than or equal to* $(2k-3)$*.)*

Proof: (Sketch) As in the previous proof, we begin by giving a necessary set of conditions for a permutation to be reachable. We then show how to reach all such permutations, except for a small number of permutations in the even case, which we explain separately.

Observe that the element 1 cannot occupy a position of even index, and the element 2 cannot occupy a position of odd index to the left of 1. Call this $\mathcal{A}$, the class of *admissible* permutations. A simple counting argument shows that this characterization produces the given formula for odd n, and also gives formula for even n upon suppression of the double-factorial correction term. This discrepancy is because when $n = 2k$ is even there are a small number of exceptional permutations which must be excluded; we will turn to these at the end of the proof.

Continuing on the assumption that $n = 2k+1$ is odd, it remains to show that all admissible permutations are in fact reachable. We do this in two stages.

Stage 1: First we will show that all permutations beginning with a 1 are reachable from the identity. We proceed in steps; after each, we will have a monotonically increasing initial segment, followed by a completed target segment. Each step increases the length of the completed segment by 1 by selecting and moving one element. If the selected element $(b > 1)$ is an even number of positions $(2k)$ away from the place to which it is to be moved, perform k times the following pair of moves: $bxy \to yxb \to xyb$; if an odd number of positions away, prepend $abcxy \to acbxy$, then continue as before. This means that we can travel between the identity and any permutation that begins with a 1.

Stage 2: It remains to show that the element 1 can always be moved to the front of any admissible permutation. Actually, we only need to show that the element 1 can always be moved *toward* the front.

If the 1 is at the very end of the permutation, the 2 must be to its left and in a position of opposite parity. Move the 2 rightward using moves 123 or 132 $\to$ 321 until it is adjacent to the 1; the 1 can then be moved.

If the 1 is not at the very end of the permutation, then consider the run of five elements centered on the 1. (Note that the existence of this 5-factor depends on the assumption that n is odd.) There are 24 cases. Most can be handled locally, but those of the form 2–1– – require more care. Checking these cases completes the proof for odd n. Now, for even n we have to consider which permutations have been included in the given characterization $\mathcal{A}$, but which are not in fact reachable.

Let $n = 2k$. Here is a description of a small class of exceptional permutations, $\mathcal{X}$, which are not reachable. Fill the positions in order $n-1, n, n-3, n-2, n-5, n-4 \dots 3, 4, 1, 2$. When filling positions of odd index, the smallest available element must be chosen; the subsequent selection of an element to place to its right is then unconstrained. None of these $(n-1)!!$ permutations is reachable. However, most of them are also not in $\mathcal{A}$, because most have the 2 in position $n-3$; the only ones we have counted are the ones where the 2 is in position n, of which there are $(n-3)!!$.

To see that none of the permutations in $\mathcal{X}$ is reachable, consider their 3-factors. These are all 213, 312, or 231; therefore these permutations are isolated points, and not in the equivalence class of the identity.

Now we have to consider which permutations in $\mathcal{A}$ are not in fact reachable. The proof for odd n only fails when the element 1 lies in the penultimate position $n-1$. We have already seen that the permutations belonging to $\mathcal{X} \cap \mathcal{A}$ are not reachable; we will show that all others are. Take any permutation $\pi \notin \mathcal{X}$, but with the minimal element 1 placed in position $n-1$. Checking the conditions from right to left, suppose all odd positions from j to $n-1$ are occupied by left-to-right minima, but suppose that the smallest element situtated in positions 1 through $j-1$ is not in position $j-2$, as expected, and its value is x.

As before, all we need to do is show that we can move the element 1 to the left. This exploits two facts: that x is the minimal element in a lefthand region, and the righthand region is alternating. First, take the element x and shift it rightward, two positions at a time, until it arrives in position $j-2$ or $j-1$. In either case, x now lies in a 321 which begins in an odd position. We check that we can propagate either of these odd 321s rightward until they capture the smallest element, which can then be moved. This completes the missing step in the proof for even n. $\square$

Theorem 12 $\#\mathrm{Eq}''\left(\iota_n, \{\{123, 321\}\}\right) = \binom{n-1}{\lfloor (n-1)/2 \rfloor}$.

Proof: (Sketch) The permutations in this class are direct sums of singletons and of blocks of odd size greater than one, where within each block the even elements are on the diagonal, and the odd elements form a *plus-indecomposable* [AAK03] 321-avoiding permutation. Let us call the set that we have just described $\mathcal{A}_n$. First we will show that $\mathcal{A}_n$ is closed under $123 \leftrightarrow 321$; since the identity is in $\mathcal{A}_n$ this will establish that the equivalence class of the identity is a subset of $\mathcal{A}_n$. Then we will show that we can return to the identity from any permutation in $\mathcal{A}_n$, which will establish that the two sets are identical.

Let π be an arbitrary permutation belonging to $\mathcal{A}_n$. Call the non-singleton blocks of π *large*; unless π is the identity, it contains at least one large block. Note that large blocks begin and end with descents.

First show that any application of $123 \to 321$ to π produces an element of $\mathcal{A}_n$. Consider the different ways that a $\pi_i, \pi_{i+1}, \pi_{i+2}$ of form 123 might occur within π. The cases to consider are (a) all three elements are in singleton blocks, (b) exactly two of the elements are in singleton blocks, (c) only the middle element is in a singleton block, (d) all three elements are in a single large block, (e) two consecutive elements are in a large block, but the third is not (which cannot arise). In each of these cases the replacement $123 \to 321$ winds up gluing together all the blocks which it straddles.

Now consider applications of $321 \to 123$ in a permutation $\rho \in \mathcal{A}_n$. Clearly, any 321 must lie within a single block, as in any two blocks, all the elements in the block to the right are larger than all the elements in the block to the left. Because the even elements within a block increase monotonically, the 321 is composed of odd, even, odd elements. It might, indeed, be the pattern resulting from a replacement $123 \to 321$ in any one of the cases (a) through (d) above; therefore by undoing it we might return to any one of these four configurations. But no matter, as each one corresponds to a permutation in $\mathcal{A}_n$.

Now we need to show that we can return to the identity from any permutation σ in $\mathcal{A}_n$. Observe that every large block of σ contains a 321 as a factor, because the first element of the block must lie below the diagonal and the last element must lie above it; therefore two consecutive odd elements exist with the first below and the second above the diagonal. Together with the even element (*on* the diagonal) which separates them, this forms a 321. Unless π is itself the identity, it contains a large block, and therefore a 321. By replacing this with a 123, we move to a permutation ρ having strictly fewer inversions than σ. But as $\mathcal{A}_n$ is closed under such replacements, we may iterate this process, until we arrive at the identity.

This establishes that the reachable permutations are as described; now to enumerate them. The number of plus-indecomposable 321-avoiding permutations on $m+1$ elements is the Catalan number $\frac{1}{m+1}\binom{2m}{m}$, so this is the number of possible blocks of size $2m+1$. We will need the following generating functions:

$$\begin{aligned} A &= \frac{1}{\sqrt{1-4x}} &&= 1 + 2x + 6x^2 + 20x^3 + 70x^4 + \ldots \\ B &= \frac{\frac{1}{\sqrt{1-4x}} - 1}{2x} &&= 1 + 3x + 10x^2 + 35x^3 + 126x^4 + \ldots \end{aligned}$$

$$C \quad = \tfrac{1-\sqrt{1-4x}}{2x} \quad = 1 + x + 2x^2 + 5x^3 + 14x^4 + \dots$$

Now, a reachable permutation of even size $2k+2$ is the direct sum of an indecomposable block of size $2i+1$ ($i \geq 0$) and a reachable permutation of odd size $2(k-i)+1$. Likewise a reachable permutation of odd size $2k+1$ is the direct sum of a block of size $2i+1$ and a reachable permutation of even size $2(k-i)$. The theorem then follows by checking that $B = AC$ and $A = (1+xB)C$. □

Theorem 13 *(a)* $\#\mathrm{Classes}''\left(n, \left\{\{123, 132, 213\}\right\}\right) = \mathrm{inv}_n$, *the number of involutions of order* n.
(b) $\#\mathrm{Eq}''\left(\pi, \left\{\{123, 132, 213\}\right\}\right)$ *is odd for all* n *and for each* $\pi \in S_n$.

Proof: (Sketch) An involution is a partition of $[n]$ into 1-cycles and 2-cycles. Write each involution as a product of cycles, with the elements increasing within each 2-cycle, and with the cycles in decreasing order of largest element. Then drop the parentheses.

The resulting set C of permutations covers all the classes, because each permutation π can be reduced to an element of C as follows: if n is at the front of π, it must stay there. (This corresponds to having n as a fixed point in the involution.) Otherwise, use $123 \to 132$ and $213 \to 132$ to push n leftward into position 2, which is as far as it will go. The element which is thus pushed into position 1 is the minimal element which was to the left of n to begin with. This is because this minimal element can never trade places with n under the given operations, as 1 is left of 3 in all of 123, 132 and 213.

This shows that the number of classes is at most the same as the number of involutions. To show that they are the same, it remains to show that each π can be reduced to a *unique* member of C. An equivalent statement (a') is that it is not possible to move from one member of C to another.

We will prove this by induction on n. At the same time we will prove statement (b) of the theorem. Assume as an induction hypothesis that both statements have been demonstrated for $n-1$ and $n-2$.

If the largest element, n, is at the front of a permutation, then it cannot move from there, so the equivalence classes split into two kinds: *special* equivalence classes, in which n is always at the front, and *ordinary* equivalence classes, in which n is never at the front. The special equivalence classes for S_n correspond upon deletion of the first elements to *all* the equivalence classes for S_{n-1}; therefore we can assert by induction the truth of both (a') and (b) as they apply to the *special* equivalence classes.

We turn to the *ordinary* equivalence classes. Consider a (directed) graph in which the vertices correspond to the permutations in S_n, and there is a blue (directed) edge from π to ρ if ρ can be obtained from π by applying $123 \to 132$, a red edge for each $213 \to 132$, and a green edge for each $123 \to 213$.

Now consider the forest of rooted trees which one obtains by taking only those red and blue edges in which the element n plays the role of the "3". The roots (i.e., sinks) of these trees are exactly the permutations in which the n has advanced as far as possible, to position 2. Each node in this forest has either zero or two children, because if it has a blue child (obtained by travelling backwards along a blue edge) then it also has a red child, and vice versa. Because each node has either zero or two children, each rooted tree has an odd number of nodes; indeed all of its level-sums are even except the zeroth level sum, which corresponds to the root vertex, which we call the *ground state*. Now we will selectively glue trees together into larger components. Namely, two trees with ground states g and h will be combined if $g(1) = h(1)$ and if, upon deleting the first two elements the shortened permutations, g'' and h'' are equivalent, regarded (in the obvious way) as members of S_{n-2}.

We claim that these larger components are exactly the connected components of our directed graph. Therefore, to complete the proof we show, by examining various cases, that there are no directed edges in the graph which escape from one component to another, in other words that all allowable moves carry us between two permutations which have equivalent ground states. □

This result is particularly interesting because the equivalence relation has the same number of classes as Knuth equivalence, yet the two relations appear to be materially different. For example, for $n = 3$, the equivalence classes for P_K have sizes 1,1,2,2, whereas for $P_3 = \{\{123, 132, 213\}\}$ the sizes are 1,1,1,3.

4 Doubly adjacent transformations

For completeness, we include a brief treatment of the situation where both indices and values are simultaneously constrained to be adjacent. As the situation is highly constrained, it is perhaps not surprising that the permutations reachable from the identity are in each case easy to classify and enumerate. Since all the treatments are similar, we can wrap them up in one proposition.

As in the previous section, we have as yet no results related to the enumeration of equivalence classes.

This proposition uses the Iverson bracket; [P] is equal to 1 if the statement P is true, and 0 otherwise.

Proposition 14 $\#\mathrm{Eq}^{\square}(\iota_n, P_1)$ *obeys the recurrence* $a(n) = a(n-1) + a(n-2)$ *with* $a_1 = a_2 = 1$. *(Fibonacci numbers* $F(n)$*, [OEIS, A000045]).*

$\#\mathrm{Eq}^{\square}(\iota_n, P_4)$ *obeys the recurrence* $a(n) = a(n-1) + a(n-3)$ *with* $a_0 = 0$*,* $a_1 = a_2 = 1$ *([OEIS, A000930]).*

$\#\mathrm{Eq}^{\square}(\iota_n, P_3) = F(n+1) - [n \text{ is even}]$.

$\#\mathrm{Eq}^{\square}(\iota_n, P_5)$ *obeys the recurrence* $a(n) = a(n-1) + a(n-2) + a(n-3)$ *with* $a(0) = a(1) = a(2) = 1$ *(Tribonacci numbers, [OEIS, A000213]).*

$\#\mathrm{Eq}^{\square}(\iota_n, P_7) = T(n+2) - [n \text{ is even}]$*, where* $T(n)$ *is sequence A000073 from Sloane[OEIS], obeying the recurrence* $a(n) = a(n-1) + a(n-2) + a(n-3)$ *with* $a(0) = a(1) = 0, a(2) = 1$.

Proof: (Sketch) We begin by characterizing the various equivalence classes. In each case, these are subsets of the layered permutations, and indeed consist of direct sums of anti-identities of dimensions either 1, 2 or 3, as follows:

P_1 ($123 \leftrightarrow 132$): direct sums of ρ_1 and ρ_2, not beginning with ρ_2

P_4 ($123 \leftrightarrow 321$): direct sums of ρ_1 and ρ_3

P_3 ($123 \leftrightarrow 132 \leftrightarrow 213$): direct sums of ρ_1 and ρ_2, including at least one ρ_1

P_5 ($123 \leftrightarrow 132 \leftrightarrow 321$): direct sums of ρ_1, ρ_2 and ρ_3, not beginning with ρ_2

P_7 ($123 \leftrightarrow 132 \leftrightarrow 213 \leftrightarrow 321$): direct sums of ρ_1, ρ_2, ρ_3, at least one of odd dimension

In each case it is easy to see that the given class remains closed under application of the appropriate operations. It is also easy in general to see how to reach a given target, especially if we cast the block sizes in the language of regular expressions. In the following, the notation $\{xy\}$ means a single block of size either x or y. An asterisk following a number means zero or more copies of that number. An asterisk following a string in [ ] means zero or more copies of that string.

P_1: The block sizes are [12*]*. Build each 12* from right to left.

P_4: The block sizes are $\{13\}$*. Build each block freely.

P_3: Split each run of 2s freely to get [12*][2*12*]*[2*1]. Build each block from the edges to the 1.

P_5: The block sizes are [$\{13\}$2*]*. First use $123 \to 132$ to build all the runs of 2 from right to left. Then use $123 \to 321$ to place 3s.

P_7: Build the 2s first, as in the case of P_3, and then place the 3s.

One now verifies all the necessary base cases, as trivially $a_1 = 1$, $a_2 = 1$, and $a_3 =$ the size of the non-singleton block of P_j.

As for the recurrences, for $n > 3$:

P_1: $a_n = a_{n-1} + a_{n-2}$, by appending respectively a ρ_1 or a ρ_2

P_2: $a_n = a_{n-1} + a_{n-3}$, by appending respectively a ρ_1 or a ρ_3

P_5: $a_n = a_{n-1} + a_{n-2} + a_{n-3}$, by appending ρ_1, ρ_2 or ρ_3

P_3: Count all direct sums of ρ_1 and ρ_2 (obviously Fibonacci) and then subtract 1 from the even terms to remove the special case 2*.

P_7: Count all direct sums of ρ_1, ρ_2, ρ_3 to get A000073, and subtract 1 from the even terms because 2* is disallowed. Alternatively, verify the recurrence $a_n = a_{n-2} + U_n$, where U_n is A000213, by noting that a permutation in $\mathrm{Eq}^{\square}(\iota_n, P_7)$ is either a ρ_2 prepended to a permutation in $\mathrm{Eq}^{\square}(\iota_{n-2}, P_7)$, or else belongs to $\mathrm{Eq}^{\square}(\iota_{n-2}, P_5)$. □

References

[AAK03] M. H. ALBERT, M. D. ATKINSON, AND M. KLAZAR, *The enumeration of simple permutations*, J. Integer Sequences, **6**, Art. 03.4.4, (2003).

[SA08] S. ASSAF, *A combinatorial realization of Schur-Weyl duality via crystal graphs and dual equivalence graphs*, FPSAC 2008, pp. 141–152, Discrete Math. Theor. Comput. Sci. Proc., Nancy, France, 2008.

[SA07] S. ASSAF, *Dual equivalence graphs, ribbon tableaux and Macdonald polynomials*, Ph.D. Thesis, UC Berkeley, 2007

[Bon04] M. BONA, *Combinatorics of Permutations*, Chapman & Hall/CRC, 2004.

[Bon06] M. BONA, *A Walk Through Combinatorics, 2nd Ed.*, World Scientific Publishing Co., 2006.

[Hai92] M. HAIMAN, *Dual equivalence with applications, including a conjecture of Proctor*, Discrete Math., **99** (1992), pp. 79–113.

[Knu70] D. KNUTH, *Permutations, matrices and generalized Young tableaux*, Pacific J. Math., **34** (1970), pp. 709–727.

[OEIS] N.J.A. Sloane, *The On-line Encyclopedia of Integer Sequences*, available at `http://www.research.att.com/~njas/sequences/`.

[Pri97] A. L. PRICE, *Packing Densities of Layered Patterns*, Ph.D. Thesis, University of Pennsylvania, 1997.

[EC1] R. STANLEY, *Enumerative Combinatorics Volume 1*, no. 49 in Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1999.

[EC2] R. STANLEY, *Enumerative Combinatorics Volume 2*, no. 62 in Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1999. With appendix 1 by Sergey Fomin.

[Stro93] W. STROMQUIST, *Packing Layered Posets Into Posets*, preprint, 1993, available at `http://walterstromquist.com/`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 790–801

Products of Geck-Rouquier conjugacy classes and the Hecke algebra of composed permutations

Pierre-Loïc Méliot[1]

[1]*Institut Gaspard Monge — Université Paris-Est Marne-La-Vallée — 77454 Marne-La-Vallée cedex 2*
meliot@phare.normalesup.org

Abstract. We show the q-analog of a well-known result of Farahat and Higman: in the center of the Iwahori-Hecke algebra $\mathscr{H}_{n,q}$, if $(a_{\lambda\mu}^{\nu}(n,q))_{\nu}$ is the set of structure constants involved in the product of two Geck-Rouquier conjugacy classes $\Gamma_{\lambda,n}$ and $\Gamma_{\mu,n}$, then each coefficient $a_{\lambda\mu}^{\nu}(n,q)$ depend on n and q in a polynomial way. Our proof relies on the construction of a projective limit of the Hecke algebras; this projective limit is inspired by the Ivanov-Kerov algebra of partial permutations.

Résumé. Nous démontrons le q-analogue d'un résultat bien connu de Farahat et Higman : dans le centre de l'algèbre d'Iwahori-Hecke $\mathscr{H}_{n,q}$, si $(a_{\lambda\mu}^{\nu}(n,q))_{\nu}$ est l'ensemble des constantes de structure mises en jeu dans le produit de deux classes de conjugaison de Geck-Rouquier $\Gamma_{\lambda,n}$ et $\Gamma_{\mu,n}$, alors chaque coefficient $a_{\lambda\mu}^{\nu}(n,q)$ dépend de façon polynomiale de n et de q. Notre preuve repose sur la construction d'une limite projective des algèbres d'Hecke ; cette limite projective est inspirée de l'algèbre d'Ivanov-Kerov des permutations partielles.

Keywords: Iwahori-Hecke algebras, Geck-Rouquier conjugacy classes, symmetric functions.

In this paper, we answer a question asked in [FW09] that concerns the products of Geck-Rouquier conjugacy classes in the Hecke algebras $\mathscr{H}_{n,q}$. If $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_r)$ is a partition with $|\lambda| + \ell(\lambda) \leq n$, we consider the completed partition

$$\lambda \to n = (\lambda_1 + 1, \lambda_2 + 1, \dots, \lambda_r + 1, 1^{n-|\lambda|-\ell(\lambda)}),$$

and we denote by $C_{\lambda,n} = C_{\lambda\to n}$ the corresponding conjugacy class, that is to say, the sum of all permutations with cycle type $\lambda \to n$ in the center of the symmetric group algebra $\mathbb{C}\mathfrak{S}_n$. Notice that in particular, $C_{\lambda,n} = 0$ if $|\lambda| + \ell(\lambda) > n$. It is known since [FH59] that the products of completed conjugacy classes write as

$$C_{\lambda,n} * C_{\mu,n} = \sum_{|\nu|\leq|\lambda|+|\mu|} a_{\lambda\mu}^{\nu}(n)\, C_{\nu,n},$$

where the structure constants $a_{\lambda\mu}^{\nu}(n)$ depend on n in a polynomial way. In [GR97], some deformations Γ_λ of the conjugacy classes C_λ are constructed. These central elements form a basis of the center $\mathscr{Z}_{n,q}$ of the Iwahori-Hecke algebra $\mathscr{H}_{n,q}$, and they are characterized by the two following properties, see [Fra99]:

1. The element Γ_λ is central and specializes to C_λ for $q = 1$.
2. The difference $\Gamma_\lambda - C_\lambda$ involves no permutation of minimal length in its conjugacy class.

As before, $\Gamma_{\lambda,n} = \Gamma_{\lambda\to n}$ if $|\lambda| + \ell(\lambda) \leq n$, and 0 otherwise. Our main result is the following:

Theorem 1 *In the center of the Hecke algebra $\mathscr{H}_{n,q}$, the products of completed Geck-Rouquier conjugacy classes write as*

$$\Gamma_{\lambda,n} * \Gamma_{\mu,n} = \sum_{|\nu|\leq|\lambda|+|\mu|} a^{\nu}_{\lambda\mu}(n,q)\ \Gamma_{\nu,n},$$

and the structure constants $a^{\nu}_{\lambda\mu}(n,q)$ are in $\mathbb{Q}[n,q,q^{-1}]$.

The first part of Theorem 1 — that is to say, that elements $\Gamma_{\nu,n}$ involved in the product satisfy the inequality $|\nu| \leq |\lambda| + |\mu|$ — was already in [FW09, Theorem 1.1], and the polynomial dependance of the coefficients $a^{\nu}_{\lambda\mu}(n,q)$ was Conjecture 3.1; our paper is devoted to a proof of this conjecture. We shall combine two main arguments:

- We construct a projective limit $\mathscr{D}_{\infty,q}$ of the Hecke algebras, which is essentially a q-version of the algebra of Ivanov and Kerov, see [IK99]. We perform *generic computations* inside various subalgebras of $\mathscr{D}_{\infty,q}$, and we project then these calculations on the algebras $\mathscr{H}_{n,q}$ and their centers.

- The centers of the Hecke algebras admit numerous bases, and these bases are related one to another in the same way as the bases of the symmetric function algebra Λ. This allows to separate the dependance on q and the dependance on n of the coefficients $a^{\nu}_{\lambda\mu}(n,q)$.

Before we start, let us fix some notations. If n is a non-negative integer, $\mathfrak{P}_n$ is the set of partitions of n, $\mathfrak{C}_n$ is the set of compositions of n, and $\mathfrak{S}_n$ is the set of permutations of the interval $[\![1,n]\!]$. The **type** of a permutation $\sigma \in \mathfrak{S}_n$ is the partition $\lambda = t(\sigma)$ obtained by ordering the sizes of the orbits of σ; for instance, $t(24513) = (3,2)$. The **code** of a composition $c \in \mathfrak{C}_n$ is the complementary in $[\![1,n]\!]$ of the set of descents of c; for instance, the code of $(3,2,3)$ is $\{1,2,4,6,7\}$. Finally, we denote by $\mathscr{Z}_n = Z(\mathbb{C}\mathfrak{S}_n)$ the center of the algebra $\mathbb{C}\mathfrak{S}_n$; the conjugacy classes C_λ form a linear basis of $\mathscr{Z}_n$ when λ runs over $\mathfrak{P}_n$.

1 Partial permutations and the Ivanov-Kerov algebra

Since our argument is essentially inspired by the construction of [IK99], let us recall it briefly. A **partial permutation** of order n is a pair (σ, S) where S is a subset of $[\![1,n]\!]$, and σ is a permutation in $\mathfrak{S}(S)$. Alternatively, one may see a partial permutation as a permutation σ in $\mathfrak{S}_n$ together with a subset containing the non-trivial orbits of σ. The product of two partial permutations is

$$(\sigma,S)\,(\tau,T) = (\sigma\tau, S\cup T),$$

and this operation yield a semigroup whose complex algebra is denoted by $\mathscr{B}_n$. There is a natural projection $\mathrm{pr}_n : \mathscr{B}_n \to \mathbb{C}\mathfrak{S}_n$ that consists in forgetting the support of a partial permutation, and also natural compatible maps

$$\phi_{N,n} : (\sigma,S) \in \mathscr{B}_N \mapsto \begin{cases} (\sigma,S) \in \mathscr{B}_n & \text{if } S \subset [\![1,n]\!]\,, \\ 0 & \text{otherwise,} \end{cases}$$

whence a projective limit $\mathscr{B}_\infty = \varprojlim \mathscr{B}_n$ with respect to this system $(\phi_{N,n})_{N\geq n}$ and in the category of filtered algebras. Now, one can lift the conjugacy classes C_λ to the algebras of partial permutations. Indeed, the symmetric group $\mathfrak{S}_n$ acts on $\mathscr{B}_n$ by

$$\sigma \cdot (\tau, S) = (\sigma\tau\sigma^{-1}, \sigma(S)),$$

and a linear basis of the invariant subalgebra $\mathscr{A}_n = (\mathscr{B}_n)^{\mathfrak{S}_n}$ is labelled by the partitions λ of size less than or equal to n:

$$\mathscr{A}_n = \bigoplus_{|\lambda|\leq n} \mathbb{C}A_{\lambda,n}, \qquad \text{where } A_{\lambda,n} = \sum_{\substack{|S|=|\lambda| \\ \sigma\in\mathfrak{S}(S),\ t(\sigma)=\lambda}} (\sigma, S).$$

Since the actions $\mathfrak{S}_n \curvearrowright \mathscr{B}_n$ are compatible with the morphisms $\phi_{N,n}$, the inverse limit $\mathscr{A}_\infty = (\mathscr{B}_\infty)^{\mathfrak{S}_\infty}$ of the invariant subalgebras has a basis $(A_\lambda)_\lambda$ indexed by all partitions $\lambda \in \mathfrak{P} = \bigsqcup_{n\in\mathbb{N}} \mathfrak{P}_n$, and such that $\phi_{\infty,n}(A_\lambda) = A_{\lambda,n}$ (with by convention $A_{\lambda,n} = 0$ if $|\lambda| > n$). As a consequence, if $(a_{\lambda\mu}^{\nu})_{\lambda,\mu,\nu}$ is the family of structure constants of the **Ivanov-Kerov algebra**[(i)] $\mathscr{A}_\infty$ in the basis $(A_\lambda)_{\lambda\in\mathfrak{P}}$, then

$$\forall n,\ \ A_{\lambda,n} * A_{\mu,n} = \sum_{\nu} a_{\lambda\mu}^{\nu} A_{\nu,n},$$

with $A_{\lambda,n} = 0$ if $|\lambda| \geq n$. Moreover, it is not difficult to see that $a_{\lambda\mu}^{\nu} \neq 0$ implies $|\nu| \leq |\lambda| + |\mu|$, and also $|\nu| - \ell(\nu) \leq |\lambda| - \ell(\lambda) + |\mu| - \ell(\mu)$, *cf.* [IK99, §10], for the study of the filtrations of $\mathscr{A}_\infty$. Now, $\mathrm{pr}_n(\mathscr{A}_n) = \mathscr{Z}_n$, and more precisely,

$$\mathrm{pr}_n(A_{\lambda,n}) = \binom{n - |\lambda| + m_1(\lambda)}{m_1(\lambda)} C_{\lambda-1,n}.$$

where $\lambda - 1 = (\lambda_1 - 1, \ldots, \lambda_s - 1)$ if $\lambda = (\lambda_1, \ldots, \lambda_s \geq 2, 1, \ldots, 1)$. The result of Farahat and Higman follows immediately, and we shall try to mimic this construction in the context of Iwahori-Hecke algebras.

2 Composed permutations and their Hecke algebra

We recall that the **Iwahori-Hecke algebra** of type A and order n is the quantized version of the symmetric group algebra defined over $\mathbb{C}(q)$ by

$$\mathscr{H}_{n,q} = \left\langle S_1, \ldots, S_{n-1} \;\middle|\; \begin{array}{l} \text{braid relations: } \forall i,\ S_iS_{i+1}S_i = S_{i+1}S_iS_{i+1} \\ \text{commutation relations: } \forall |j-i|>1,\ S_iS_j = S_jS_i \\ \text{quadratic relations: } \forall i,\ (S_i)^2 = (q-1)\,S_i + q \end{array} \right\rangle.$$

When $q = 1$, we recover the symmetric group algebra $\mathbb{C}\mathfrak{S}_n$. If $\omega \in \mathfrak{S}_n$, let us denote by T_ω the product $S_{i_1}S_{i_2}\cdots S_{i_r}$, where $\omega = s_{i_1}s_{i_2}\cdots s_{i_r}$ is any reduced expression of ω in elementary transpositions $s_i = (i, i+1)$. Then, it is well known that the elements T_ω do not depend on the choice of reduced expressions, and that they form a $\mathbb{C}(q)$-linear basis of $\mathscr{H}_{n,q}$, see [Mat99].

(i) It can be shown that $\mathscr{A}_\infty$ is isomorphic to the algebra of shifted symmetric polynomials, see Theorem 9.1 in [IK99].

In order to construct a *projective* limit of the algebras $\mathscr{H}_{n,q}$, it is very tempting to mimic the construction of Ivanov and Kerov, and therefore to build an Hecke algebra of partial permutations. Unfortunately, this is not possible; let us explain why by considering for instance the transposition $\sigma = 1432$ in $\mathfrak{S}_4$. The possible supports for σ are $\{2,4\}$, $\{1,2,4\}$, $\{2,3,4\}$ and $\{1,2,3,4\}$. However,

$$\sigma = s_2 s_3 s_2,$$

and the support of s_2 (respectively, of s_3) contains at least $\{2,3\}$ (resp., $\{3,4\}$). So, if we take account of the Coxeter structure of $\mathfrak{S}_4$ — and it should obviously be the case in the context of Hecke algebras — then the only valid supports for σ are the connected ones, namely, $\{2,3,4\}$ and $\{1,2,3,4\}$. This problem leads to consider *composed permutations* instead of *partial permutations*. If c is a composition of n, let us denote by $\pi(c)$ the corresponding set partition of $[\![1,n]\!]$, *i.e.*, the set partition whose parts are the intervals $[\![1,c_1]\!]$, $[\![c_1+1,c_1+c_2]\!]$, etc. A **composed permutation** of order n is a pair (σ,c) with $\sigma \in \mathfrak{S}_n$ and c composition in $\mathfrak{C}_n$ such that $\pi(c)$ is coarser than the set partition of orbits of σ. For instance, $(32154867,(5,3))$ is a composed permutation of order 8; we shall also write this $32154|867$. The product of two composed permutations is then defined by

$$(\sigma,c)\,(\tau,d) = (\sigma\tau, c \vee d),$$

where $c \vee d$ is the finest composition of n such that $\pi(c \vee d) \geq \pi(c) \vee \pi(d)$ in the lattice of set partitions. For instance,

$$321|54|867 \,\times\, 12|435|687 = 42153|768.$$

One obtains so a semigroup of composed permutations; its complex semigroup algebra will be denoted by[(ii)] $\mathscr{D}_n$, and the dimension of $\mathscr{D}_n$ is the number of composed permutations of order n.

Now, let us describe an Hecke version $\mathscr{D}_{n,q}$ of the algebra $\mathscr{D}_n$. As for $\mathscr{H}_{n,q}$, one introduces generators $(S_i)_{1\leq i\leq n-1}$ corresponding to the elementary transpositions s_i, but one has also to introduce generators $(I_i)_{1\leq i\leq n-1}$ that allow to join the parts of the composition of a composed permutation. Hence, the **Iwahori-Hecke algebra of composed permutations** is defined (over the ground field $\mathbb{C}(q)$) by $\mathscr{D}_{n,q} = \langle S_1,\ldots,S_{n-1},I_1,\ldots,I_{n-1}\rangle$ and the following relations:

$$\begin{aligned}
&\forall i,\;\; S_iS_{i+1}S_i = S_{i+1}S_iS_{i+1}\\
&\forall |j-i|>1,\;\; S_iS_j = S_jS_i\\
&\forall i,\;\; (S_i)^2 = (q-1)\,S_i + q\,I_i\\
&\forall i,j,\;\; S_iI_j = I_jS_i\\
&\forall i,j,\;\; I_iI_j = I_jI_i\\
&\forall i,\;\; S_iI_i = S_i\\
&\forall i,\;\; (I_i)^2 = I_i
\end{aligned}$$

The generators S_i correspond to the composed permutations $1|2|\ldots|i-1|i+1,i|i+2|\ldots|n$, and the generators I_i correspond to the composed permutations $1|2|\ldots|i-1|i,i+1|i+2|\ldots|n$.

(ii) If one considers pairs (σ,π) where π is any set partition of $[\![1,n]\!]$ coarser than $\mathrm{orb}(\sigma)$ (and not necessarily a set partition associated to a composition), then one obtains an algebra of *split permutations* whose subalgebra of invariants is related to the connected Hurwitz numbers $H_{n,g}(\lambda)$.

Proposition 2 *The algebra $\mathscr{D}_{n,q}$ specializes to the algebra of composed permutations $\mathscr{D}_n$ when $q = 1$; to the Iwahori-Hecke algebra $\mathscr{H}_{n,q}$ when $I_1 = I_2 = \cdots = I_{n-1} = 1$; and to the algebra $\mathscr{D}_{m,q}$ of lower order $m < n$ when $I_m = I_{m+1} = \cdots = I_{n-1} = 0$ and $S_m = S_{m+1} = \cdots = S_{n-1} = 0$.*

In the following, we shall denote by pr_n the specialization $\mathscr{D}_{n,q} \to \mathscr{H}_{n,q}$; it generalizes the map $\mathscr{D}_n \to \mathbb{C}\mathfrak{S}_n$ of the first section. The first part of Proposition 2 is actually the only one that is non trivial, and it will be a consequence of Theorem 3. If ω is a permutation with reduced expression $\omega = s_{i_1} s_{i_2} \cdots s_{i_r}$, we denote as before by T_ω the product $S_{i_1} S_{i_2} \ldots S_{i_r}$ in $\mathscr{D}_{n,q}$. On the other hand, if c is a composition of $[\![1, n]\!]$, we denote by I_c the product of the generators I_j with j in the code of c (so for instance, $I_{(3,2,3)} = I_1 I_2 I_4 I_6 I_7$ in $\mathscr{D}_{8,q}$). These elements are central idempotents, and I_c correspond to the composed permutation (id, c). Finally, if (σ, c) is a composed permutation, $T_{\sigma,c}$ is the product $T_\sigma I_c$.

Theorem 3 *In $\mathscr{D}_{n,q}$, the products T_σ do not depend on the choice of reduced expressions, and the products $T_{\sigma,c}$ form a linear basis of $\mathscr{D}_{n,q}$ when (σ, c) runs over composed permutations of order n. There is an isomorphism of $\mathbb{C}(q)$-algebras between*

$$\mathscr{D}_{n,q} \quad \textit{and} \quad \bigoplus_{c \in \mathfrak{C}_n} \mathscr{H}_{c,q},$$

where $\mathscr{H}_{c,q}$ is the Young subalgebra $\mathscr{H}_{c_1,q} \otimes \mathscr{H}_{c_2,q} \otimes \cdots \otimes \mathscr{H}_{c_r,q}$ of $\mathscr{H}_{n,q}$.

Proof: If $\sigma \in \mathfrak{S}_n$, the Matsumoto theorem ensures that it is always possible to go from a reduced expression $s_{i_1} s_{i_2} \cdots s_{i_r}$ to another reduced expression $s_{j_1} s_{j_2} \cdots s_{j_r}$ by braid moves $s_i s_{i+1} s_i \leftrightarrow s_{i+1} s_i s_{i+1}$ and commutations $s_i s_j \leftrightarrow s_j s_i$ when $|j - i| > 1$. Since the corresponding products of S_i in $\mathscr{D}_{n,q}$ are preserved by these substitutions, a product T_σ in $\mathscr{D}_{n,q}$ does not depend on the choice of a reduced expression. Now, let us consider an arbitrary product Π of generators S_i and I_j (in any order). As the elements I_j are central idempotents, it is always possible to reduce the product to

$$\Pi = S_{i_1} S_{i_2} \cdots S_{i_p}\, I_c$$

with c composition of n — here, $s_{i_1} s_{i_2} \cdots s_{i_p}$ is *a priori* not a reduced expression. Moreover, since $S_i\, I_i = S_i$, we can suppose that the code of c contains $\{i_1, \ldots, i_p\}$. Now, suppose that $\sigma = s_{i_1} s_{i_2} \cdots s_{i_p}$ is not a reduced expression. Then, by using braid moves and commutations, we can transform the expression in one with two consecutive letters that are identical, that is to say that if $j_k = j_{k+1}$,

$$\sigma = s_{j_1} \cdots s_{j_k} s_{j_{k+1}} \cdots s_{j_p} = s_{j_1} \cdots s_{j_{k-1}} s_{j_{k+2}} \cdots s_{j_p}.$$

We apply the same moves to the S_i in $\mathscr{D}_{n,q}$ and we obtain $\Pi = S_{j_1} \cdots S_{j_k} S_{j_{k+1}} \cdots S_{j_p}\, I_c$; notice that the code of c still contains $\{j_1, \ldots, j_p\} = \{i_1, \ldots, i_p\}$. By using the quadratic relation in $\mathscr{D}_{n,q}$, we conclude that if $j_k = j_{k+1}$,

$$\begin{aligned}\Pi &= (q-1)\, S_{j_1} \cdots S_{j_{k-1}} S_{j_k} S_{j_{k+2}} \cdots S_{j_p}\, I_c + q\, S_{j_1} \cdots S_{j_{k-1}} I_{j_k} S_{j_{k+2}} \cdots S_{j_p}\, I_c \\ &= (q-1)\, S_{j_1} \cdots S_{j_{k-1}} S_{j_k} S_{j_{k+2}} \cdots S_{j_p}\, I_c + q\, S_{j_1} \cdots S_{j_{k-1}} S_{j_{k+2}} \cdots S_{j_p}\, I_c\end{aligned}$$

because $I_{j_k} I_c = I_c$. Consequently, by induction on p, any product Π is a $\mathbb{Z}[q]$-linear combination of products $T_{\tau,c}$ (and with the same composition c for all the terms of the linear combination). So, the reduced products $T_{\sigma,c}$ span linearly $\mathscr{D}_{n,q}$ when (σ, c) runs over composed permutations of order n. If c is in $\mathfrak{C}_n$, we define a morphism of $\mathbb{C}(q)$-algebras from $\mathscr{D}_{n,q}$ to $\mathscr{H}_{c,q}$ by

$$\psi_c(S_i) = \begin{cases} S_i & \text{if } i \text{ is in the code of } c, \\ 0 & \text{otherwise,} \end{cases} \quad ; \quad \psi_c(I_i) = \begin{cases} 1 & \text{if } i \text{ is in the code of } c, \\ 0 & \text{otherwise.} \end{cases}$$

The elements $\psi_c(S_i)$ and $\psi_c(I_i)$ sastify in $\mathscr{H}_{c,q}$ the relations of the generators S_i and I_i in $\mathscr{D}_{n,q}$. So, there is indeed such a morphism of algebras $\psi_c : \mathscr{D}_{n,q} \to \mathscr{H}_{c,q}$, and one has in fact $\psi_c(T_{\sigma,b}) = T_\sigma$ if $\pi(b) \leq \pi(c)$, and 0 otherwise. Let us consider the direct sum of algebras $\mathscr{H}_{\mathfrak{C}_n,q} = \bigoplus_{c \in \mathfrak{C}_n} \mathscr{H}_{c,q}$, and the direct sum of morphisms $\psi = \bigoplus_{c \in \mathfrak{C}_n} \psi_c$. We denote the basis vectors $[0, 0, \ldots, (T_\sigma \in \mathscr{H}_{c,q}), \ldots, 0]$ of $\mathscr{H}_{\mathfrak{C}_n,q}$ by $T_{\sigma \in \mathscr{H}_{c,q}}$; in particular,

$$\psi(T_{\sigma,c}) = \sum_{d \geq c} T_{\sigma \in \mathscr{H}_{d,q}}$$

for any composed permutation (σ, c). As a consequence, the map ψ is surjective, because

$$\psi\left(\sum_{d \geq c} \mu(c,d)\, T_{\sigma,c}\right) = T_{\sigma \in \mathscr{H}_{c,q}}$$

where $\mu(c,d) = \mu(\pi(c), \pi(d)) = (-1)^{\ell(c)-\ell(d)}$ is the Möbius function of the hypercube lattice of compositions. If σ is a permutation, we denote by $\mathrm{orb}(\sigma)$ the set partition whose parts are the orbits of σ. Since the families $(T_{\sigma,c})_{\mathrm{orb}(\sigma) \leq \pi(c)}$ and $(T_{\sigma \in \mathscr{H}_{c,q}})_{\mathrm{orb}(\sigma) \leq \pi(c)}$ have the same cardinality $\dim \mathscr{D}_n$, we conclude that $(T_{\sigma,c})_{\mathrm{orb}(\sigma) \leq \pi(c)}$ is a $\mathbb{C}(q)$-linear basis of $\mathscr{D}_{n,q}$ and that ψ is an isomorphism of $\mathbb{C}(q)$-algebras. $\square$

Notice that the second part of Theorem 3 is the q-analog of Corollary 3.2 in [IK99]. To conclude this part, we have to build the inverse limit $\mathscr{D}_{\infty,q} = \varprojlim \mathscr{D}_{n,q}$, but this is easy thanks to the specializations evoked in the third part of Proposition 2. Hence, if $\phi_{N,n} : \mathscr{D}_{N,q} \to \mathscr{D}_{n,q}$ is the map that sends the generators $I_{i \geq n}$ and $S_{i \geq n}$ to zero and that preserves the other generators, then $(\phi_{N,n})_{N \geq n}$ is a system of compatible maps, and these maps behave well with respect to the filtration $\deg T_{\sigma,c} = |\mathrm{code}(c)|$. Consequently, there is a projective limit $\mathscr{D}_{\infty,q}$ whose elements are the infinite linear combinations of $T_{\sigma,c}$, with σ finite permutation in $\mathfrak{S}_\infty$ and c infinite composition compatible with σ and with almost all its parts of size 1.

It is not true that two elements x and y in $\mathscr{D}_{\infty,q}$ are equal if and only if their projections $\mathrm{pr}_n(\phi_{\infty,n}(x))$ and $\mathrm{pr}_n(\phi_{\infty,n}(y))$ are equal for all n: for instance,

$$T[21|34|5|6|\cdots] = S_1 I_1 I_3 \quad \text{and} \quad T[2134|5|6|\cdots] = S_1 I_1 I_2 I_3$$

have the same projections in all the Hecke algebras (namely, S_1 if $n \geq 4$ and 0 otherwise), but they are not equal. However, the result is true if we consider only the subalgebras $\mathscr{D}'_{n,q} \subset \mathscr{D}_{n,q}$ spanned by the $T_{\sigma,c}$ with $c = (k, 1, \ldots, 1)$ — then, σ may be considered as a partial permutation of $[\![1, k]\!]$.

Proposition 4 *For any n, the vector space $\mathscr{D}'_{n,q}$ spanned by the $T_{\sigma,c}$ with $c = (k, 1^{n-k})$ is a subalgebra of $\mathscr{D}_{n,q}$. In the inverse limit $\mathscr{D}'_{\infty,q} \subset \mathscr{D}_{\infty,q}$, the projections $\mathrm{pr}_{\infty,n} = \mathrm{pr}_n \circ \phi_{\infty,n}$ separate the vectors:*

$$\forall x, y \in \mathscr{D}'_{\infty,q}, \qquad \big(\forall n,\ \mathrm{pr}_{\infty,n}(x) = \mathrm{pr}_{\infty,n}(y)\big) \iff \big(x = y\big).$$

Proof: The supremum of two compositions $(k, 1^{n-k})$ and $(l, 1^{n-l})$ is $(m, 1^{n-m})$ with $m = \max(k, l)$; consequently, $\mathscr{D}'_{n,q}$ is indeed a subalgebra of $\mathscr{D}_{n,q}$. Any element x of the projective limit $\mathscr{D}'_{\infty,q}$ writes uniquely as

$$x = \sum_{k=0}^{\infty} \sum_{\sigma \in \mathfrak{S}_k} a_{\sigma,k}(x)\, T_{\sigma,(k,1^\infty)}.$$

Suppose that x and y have the same projections, and let us fix a permutation σ. There is a minimal integer k such that $\sigma \in \mathfrak{S}_k$, and $a_{\sigma,k}(x)$ is the coefficient of T_σ in $\mathrm{pr}_{\infty,k}(x)$; consequently, $a_{\sigma,k}(x) = a_{\sigma,k}(y)$. Now, $a_{\sigma,k}(x) + a_{\sigma,k+1}(x)$ is the coefficient of T_σ in $\mathrm{pr}_{\infty,k+1}(x)$, so one has also $a_{\sigma,k}(x) + a_{\sigma,k+1}(x) = a_{\sigma,k}(y) + a_{\sigma,k+1}(y)$, and $a_{\sigma,k+1}(x) = a_{\sigma,k+1}(y)$. By using the same argument and by induction on l, we conclude that $a_{\sigma,k+l}(x) = a_{\sigma,k+l}(y)$ for every l, and therefore $x = y$. We have then proved that the projections separate the vectors in $\mathscr{D}'_{\infty,q}$. $\square$

3 Bases of the center of the Hecke algebra

In the following, $\mathscr{Z}_{n,q}$ is the center of $\mathscr{H}_{n,q}$. We have already given a characterization of the **Geck-Rouquier central elements** Γ_λ, and they form a linear basis of $\mathscr{Z}_{n,q}$ when λ runs over $\mathfrak{P}_n$. Let us write down explicitly this basis when $n = 3$:

$$\Gamma_3 = T_{231} + T_{312} + (q-1)q^{-1}\, T_{321} \qquad ; \qquad \Gamma_{2,1} = T_{213} + T_{132} + q^{-1}\, T_{321} \qquad ; \qquad \Gamma_{1,1,1} = T_{123}$$

The first significative example of Geck-Rouquier element is actually when $n = 4$. Thus, if one considers

$$\begin{aligned}\Gamma_{3,1} &= T_{1342} + T_{1423} + T_{2314} + T_{3124} + q^{-1}\,(T_{2431} + T_{4132} + T_{3214} + T_{4213}) \\ &\quad + (q-1)q^{-1}\,(T_{1432} + T_{3214}) + (q-1)q^{-2}\,(T_{3421} + T_{4312} + 2\,T_{4231}) + (q-1)^2 q^{-3}\, T_{4321},\end{aligned}$$

the terms with coefficient 1 are the four minimal 3-cycles in $\mathfrak{S}_4$; the terms whose coefficients specialize to 1 when $q = 1$ are the eight 3-cycles in $\mathfrak{S}_4$; and the other terms are not minimal in their conjugacy classes, and their coefficients vanish when $q = 1$.

It is really unclear how one can lift these elements to the Hecke algebras of composed permutations; fortunately, the center $\mathscr{Z}_{n,q}$ admits other linear bases that are easier to pull back from $\mathscr{H}_{n,q}$ to $\mathscr{D}_{n,q}$. In [Las06], seven different bases for $\mathscr{Z}_{n,q}$ are studied(iii), and it is shown that up to diagonal matrices that depend on q in a polynomial way, the transition matrices between these bases are the same as the transition matrices between the usual bases of the algebra of symmetric functions. We shall only need the **norm basis** N_λ, whose properties are recalled in Proposition 5. If c is a composition of n and $\mathfrak{S}_c$ is the corresponding Young subgroup of $\mathfrak{S}_n$, it is well-known that each coset in $\mathfrak{S}_n/\mathfrak{S}_c$ or $\mathfrak{S}_c\backslash\mathfrak{S}_n$ has a unique representative ω of minimal length which is called the **distinguished representative** — this fact is even true for parabolic double cosets. In what follows, we rather work with right cosets, and the distinguished representatives of $\mathfrak{S}_c\backslash\mathfrak{S}_n$ are precisely the permutation words whose recoils are contained in the set of descents of c. So for instance, if $c = (2, 3)$, then

$$\mathfrak{S}_{(2,3)}\backslash\mathfrak{S}_5 = \{12345, 13245, 13425, 13452, 31245, 31425, 31452, 34125, 34152, 34512\} = 12 \sqcup\!\sqcup\, 345.$$

(iii) One can also consult [Jon90] and [Fra99].

Proposition 5 *[Las06, Theorem 7] If c is a composition of n, let us denote by N_c the element*

$$\sum_{\omega\in\mathfrak{S}_c\backslash\mathfrak{S}_n} q^{-\ell(\omega)}\,T_{\omega^{-1}}\,T_\omega$$

in the Hecke algebra $\mathscr{H}_{n,q}$. Then, N_c does not depend on the order of the parts of c, and the N_λ form a linear basis of $\mathscr{Z}_{n,q}$ when λ runs over $\mathfrak{P}_n$ — in particular, the norms N_c are central elements. Moreover,

$$(\Gamma_\lambda)_{\lambda\in\mathfrak{P}_n} = D\cdot M2E\cdot(N_\mu)_{\mu\in\mathfrak{P}_n},$$

where $M2E$ is the transition matrice between monomial functions m_λ and elementary functions e_μ, and D is the diagonal matrix with coefficients $(q/(q-1))^{n-\ell(\lambda)}$.

So for instance, $\Gamma_3 = q^2\,(q-1)^{-2}\,(3\,N_3 - 3\,N_{2,1} + N_{1,1,1})$, because $m_3 = 3\,e_3 - 3\,e_{2,1} + e_{1,1,1}$. Let us write down explicitly the norm basis when $n = 3$:

$$N_3 = T_{123} \qquad ; \qquad N_{2,1} = 3\,T_{123} + (q-1)q^{-1}\,(T_{213}+T_{132}) + (q-1)q^{-2}\,T_{321}$$
$$N_{1,1,1} = 6\,T_{123} + 3(q-1)q^{-1}\,(T_{213}+T_{132}) + (q-1)^2q^{-2}\,(T_{231}+T_{312}) + (q^3-1)q^{-3}\,T_{321}$$

We shall see hereafter that these norms have natural preimages by the projections pr_n and $\mathrm{pr}_{\infty,n}$.

4 Generic norms and the Hecke-Ivanov-Kerov algebra

Let us fix some notations. If c is a composition of size $|c|$ less than n, then $c\uparrow n$ is the composition $(c_1,\dots,c_r,n-|c|)$, $J_c = I_1I_2\cdots I_{|c|-1}$, and

$$M_{c,n} = \sum_{\omega\in\mathfrak{S}_{c\uparrow n}\backslash\mathfrak{S}_n} q^{-\ell(\omega)}\,T_{\omega^{-1}}\,T_\omega\,J_c,$$

the products T_ω being considered as elements of $\mathscr{D}_{n,q}$. So, $M_{c,n}$ is an element of $\mathscr{D}_{n,q}$, and we set $M_{c,n} = 0$ if $|c| > n$.

Proposition 6 *For any N,n and any composition c, $\phi_{N,n}(M_{c,N}) = M_{c,n}$, and $\mathrm{pr}_n(M_{c,n}) = N_{c\uparrow n}$ if $|c|\leq n$, and 0 otherwise. On the other hand, $M_{c,n}$ is always in $\mathscr{D}'_{n,q}$.*

Proof: Because of the description of distinguished representatives of right cosets by positions of recoils, if $|c|\leq n$, then the sum $M_{c,n}$ is over permutation words ω with recoils in the set of descents of c (notice that we include $|c|$ in the set of descents of c). Let us denote by $R_{c,n}$ this set of words, and suppose that $|c|\leq n-1$. If $\omega\in R_{c,n}$ is such that $\omega(n)\neq n$, then T_ω involves S_{n-1}, so the image by $\phi_{n,n-1}$ of the corresponding term in $M_{c,n}$ is zero. On the other hand, if $\omega(n) = n$, then any reduced decomposition of T_ω does not involve S_{n-1}, so the corresponding term in $M_{c,n}$ is preserved by $\phi_{n,n-1}$. Consequently, $\phi_{n,n-1}(M_{c,n})$ is a sum with the same terms as $M_{c,n}$, but with ω running over $R_{c,n-1}$; so, we have proved that $\phi_{n,n-1}(M_{c,n}) = M_{c,n-1}$ when $|c|\leq n-1$. The other cases are much easier: thus, if $|c| = n$, then

$M_{c,n-1} = 0$, and $\phi_{n,n-1}(M_{c,n})$ is also zero because $\phi_{n,n-1}(J_c) = 0$. And if $|c| > n$, then $M_{c,n}$ and $M_{c,n-1}$ are both equal to zero, and again $\phi_{n,n-1}(M_{c,n}) = M_{c,n-1}$. Since

$$\phi_{N,n} = \phi_{n+1,n} \circ \phi_{n+2,n+1} \circ \cdots \circ \phi_{N,N-1},$$

we have proved the first part of the proposition, and the second part is really obvious.

Now, let us show that $M_{c,n}$ is in $\mathscr{D}'_{n,q}$. Notice that the result is trivial if $|c| > n$, and also if $|c| = n$, because we have then $J_c = I_{(n)}$, and therefore $d = (n)$ for any composed permutation (σ, d) involved in $M_{c,|c|}$. Suppose then that $|c| \leq n-1$. Because of the description of $\mathfrak{S}_d \backslash \mathfrak{S}_{|d|}$ as a shuffle product, any distinguished representative ω of $\mathfrak{S}_{c\uparrow n} \backslash \mathfrak{S}_n$ is the shuffle of a distinguished representative ω_c of $\mathfrak{S}_c \backslash \mathfrak{S}_{|c|}$ with the word $|c|+1, |c|+2, \ldots, n$. For instance, 5613724 is the distinguished representative of a right $\mathfrak{S}_{(2,2,3)}$-coset, and it is a shuffle of 567 with the distinguished representative 1324 of a right $\mathfrak{S}_{(2,2)}$-coset. Let us denote by $s_{i_1} \cdots s_{i_r}$ a reduced expression of ω_c, and by $j_{|c|+1}, \ldots, j_n$ the positions of $|c|+1, \ldots, n$ in ω. Then, it is not difficult to see that

$$s_{i_1} \cdots s_{i_r} \times (s_{|c|} s_{|c|-1} \cdots s_{j_{|c|+1}})\,(s_{|c|+1} s_{|c|} \cdots s_{j_{|c|+2}}) \cdots (s_{n-1} s_{n-2} \cdots s_{j_n})$$

is a reduced expression for ω; for instance, s_2 is the reduced expression of 1324, and

$$s_2 \times (s_4 s_3 s_2 s_1)\,(s_5 s_4 s_3 s_2)\,(s_6 s_5)$$

is a reduced expression of 5613724. From this, we deduce that $T_\omega\, J_c = T_{\omega,(k,1^{n-k})}$, where k is the highest integer in $[\![|c|+1, n]\!]$ such that $j_k < k$ — we take $k = |c|$ if $\omega = \omega_c$. Then, the multiplication by $T_{\omega^{-1}}$ cannot fatten the composition anymore, so $T_{\omega^{-1}} T_\omega\, J_c$ is a linear combination of $T_{\tau,(k,1^{n-k})}$, and we have proved that $M_{n,c}$ is indeed in $\mathscr{D}'_{n,q}$. $\square$

From the previous proof, it is now clear that if we consider the infinite sum $M_c = \sum q^{-\ell(\omega)}\, T_{\omega^{-1}}\, T_\omega\, J_c$ over permutation words $\omega \in \mathfrak{S}_\infty$ with their recoils in the set of descents of c, then M_c is the unique element of $\mathscr{D}_{\infty,q}$ such that $\phi_{\infty,n}(M_c) = M_{c,n}$ for any positive integer n, and also the unique element of $\mathscr{D}'_{\infty,q}$ such that $\mathrm{pr}_{\infty,n}(M_c) = N_{c\uparrow n}$ for any positive integer n (with by convention $N_{c\uparrow n} = 0$ if $|c| > n$). In particular, M_c does not depend on the order of the parts of c, because this is true for the $N_{c\uparrow n}$ and the projections separate the vectors in $\mathscr{D}'_{\infty,q}$. Consequently, we shall consider only elements M_λ labelled by partitions λ of arbitrary size, and call them **generic norms**. For instance:

$$M_{(2),3} = T_{12|3} + 2\,T_{123} + (1-q^{-1})\,(T_{132} + T_{213}) + (q^{-1} - q^{-2})\,T_{321}$$

In what follows, if $i < n$, we denote by $(S_i)^{-1}$ the element of $\mathscr{D}_{n,q}$ equal to:

$$(S_i)^{-1} = q^{-1}\, S_i + \left(q^{-1} - 1\right) I_i$$

The product $S_i\,(S_i)^{-1} = (S_i)^{-1} S_i$ equals I_i in $\mathscr{D}_{n,q}$, and by the specialization $\mathrm{pr}_n : \mathscr{D}_{n,q} \to \mathscr{H}_{n,q}$, one recovers $S_i\,(S_i)^{-1} = 1$ in the Hecke algebra $\mathscr{H}_{n,q}$.

Theorem 7 *The M_λ span linearly the subalgebra $\mathscr{C}_{\infty,q} \subset \mathscr{D}'_{\infty,q}$ that consists in elements $x \in \mathscr{D}'_{\infty,q}$ such that $I_i\, x = S_i\, x\, (S_i)^{-1}$ for every i. In particular, any product $M_\lambda * M_\mu$ is a linear combination of M_ν, and moreover, the terms M_ν involved in the product satisfy the inequality $|\nu| \leq |\lambda| + |\mu|$.*

Proof: If $I_i\,x = S_i\,x\,(S_i)^{-1}$ and $I_i\,y = S_i\,y\,(S_i)^{-1}$, then

$$I_i\,xy = I_i\,x\,I_i\,y = S_i\,x\,(S_i)^{-1} S_i\,y\,(S_i)^{-1} = S_i\,x\,I_i\,y\,(S_i)^{-1} = S_i\,xy\,(S_i)^{-1},$$

so the elements that "commute" with S_i in $\mathscr{D}_{\infty,q}$ form a subalgebra. As an intersection, $\mathscr{C}_{\infty,q}$ is also a subalgebra of $\mathscr{D}_{\infty,q}$; let us see why it is spanned by the generic norms. If $\mathscr{D}'_{\infty,q,i}$ is the subspace of $\mathscr{D}_{\infty,q}$ spanned by the $T_{\sigma,c}$ with $c = (k, 1^{\infty}) \vee (1^{i-1}, 2, 1^{\infty})$, then the projections separate the vectors in this subspace — this is the same proof as in Proposition 4. For $\lambda \in \mathfrak{P}$, $I_i\,M_\lambda$ and $S_i\,M_\lambda\,(S_i)^{-1}$ belong to $\mathscr{D}'_{\infty,q,i}$, and they have the same projections in $\mathscr{H}_{n,q}$, because $\mathrm{pr}_{\infty,n}(M_\lambda)$ is a norm and in particular a central element. Consequently, $I_i\,M_\lambda = S_i\,M_\lambda\,(S_i)^{-1}$, and the M_λ are indeed in $\mathscr{C}_{\infty,q}$. Now, if we consider an element $x \in \mathscr{C}_{\infty,q}$, then for $i < n$, $\mathrm{pr}_n(x) = S_i\,\mathrm{pr}_n(x)\,(S_i)^{-1}$, so $\mathrm{pr}_n(x)$ is in $\mathscr{Z}_{n,q}$ and is a linear combination of norms:

$$\forall n \in \mathbb{N},\ \ \mathrm{pr}_n(x) = \sum_{\lambda \in \mathfrak{P}_n} a_\lambda(x)\, N_\lambda$$

Since the same holds for any difference $x - \sum b_\lambda\, M_\lambda$, we can construct by induction on n an infinite linear combination S_∞ of M_λ that has the same projections as x:

$$\mathrm{pr}_1(x) = \sum_{|\lambda|=1} b_\lambda\, N_\lambda \quad \Rightarrow \quad \mathrm{pr}_1\!\left(x - \sum_{|\lambda|=1} b_\lambda\, M_\lambda\right) = 0,\ \ S_1 = \sum_{|\lambda|=1} b_\lambda\, M_\lambda$$

$$\mathrm{pr}_2\,(x - S_1) = \sum_{|\lambda|=2} b_\lambda\, N_\lambda \quad \Rightarrow \quad \mathrm{pr}_{1,2}\!\left(x - \sum_{|\lambda|\leq 2} b_\lambda\, M_\lambda\right) = 0,\ \ S_2 = \sum_{|\lambda|\leq 2} b_\lambda\, M_\lambda$$

$$\vdots$$

$$\mathrm{pr}_{n+1}\,(x - S_n) = \sum_{|\lambda|=n+1} b_\lambda\, N_\lambda \quad \Rightarrow \quad S_{n+1} = S_n + \sum_{|\lambda|=n+1} b_\lambda\, M_\lambda = \sum_{|\lambda|\leq n+1} b_\lambda\, M_\lambda$$

Then, $S_\infty = \sum_{\lambda \in \mathfrak{P}} b_\lambda\, M_\lambda$ is in $\mathscr{D}'_{\infty,q}$ and has the same projections as x, so $S_\infty = x$. In particular, since $\mathscr{C}_{\infty,q}$ is a subalgebra, a product $M_\lambda * M_\mu$ is in $\mathscr{C}_{\infty,q}$ and is an *a priori* infinite linear combination of M_ν:

$$\forall \lambda, \mu,\ \ M_\lambda * M_\mu = \sum g^{\nu}_{\lambda\mu}\, M_\nu$$

Since the norms N_λ are defined over $\mathbb{Z}[q, q^{-1}]$, by projection on the Hecke algebras $\mathscr{H}_{n,q}$, one sees that the $g^{\nu}_{\lambda\mu}$ are also in $\mathbb{Z}[q, q^{-1}]$ — in fact, they are *symmetric* polynomials in q and q^{-1}. It remains to be shown that the previous sum is in fact over partitions $|\nu|$ with $|\nu| \leq |\lambda| + |\mu|$; we shall see why this is true in the last paragraph[iv]. $\Box$

For example, $M_1 * M_1 = M_1 + (q + 1 + q^{-1})\, M_{1,1} - (q + 2 + q^{-1})\, M_2$, and from this generic identity one deduces the expression of any product $(N_{(1)\uparrow n})^2$, *e.g.*,

$$N_{1,1}^2 = (q + 2 + q^{-1})\,(N_{1,1} - N_2) \qquad ; \qquad N_{3,1}^2 = N_{3,1} + (q + 1 + q^{-1})\, N_{2,1,1} - (q + 2 + q^{-1})\, N_{2,2}.$$

Let us denote by $\mathscr{A}_{\infty,q}$ the subspace of $\mathscr{C}_{\infty,q}$ whose elements are *finite* linear combinations of generic norms; this is in fact a subalgebra, which we call the **Hecke-Ivanov-Kerov algebra** since it plays the same role for Iwahori-Hecke algebras as $\mathscr{A}_\infty$ for symmetric group algebras.

[iv] Unfortunately, we did not succeed in proving this result with adequate filtrations on $\mathscr{D}_{\infty,q}$ or $\mathscr{D}'_{\infty,q}$.

5 Completion of partitions and symmetric functions

The proof of Theorem 1 and of the last part of Theorem 7 relies now on a rather elementary property of the transition matrices $M2E$ and $E2M$. By convention, we set $e_{\lambda\uparrow n} = 0$ if $|\lambda| > n$, and $m_{\lambda\to n} = 0$ if $|\lambda| + \ell(\lambda) > n$. Then:

Proposition 8 *There exists polynomials $P_{\lambda\mu}(n) \in \mathbb{Q}[n]$ and $Q_{\lambda\mu}(n) \in \mathbb{Q}[n]$ such that*

$$\forall \lambda, n, \qquad m_{\lambda\to n} = \sum_{\mu' \leq_d \lambda} P_{\lambda\mu}(n)\; e_{\mu\uparrow n} \quad \text{and} \quad e_{\lambda\uparrow n} = \sum_{\mu \leq_d \lambda'} Q_{\lambda\mu}(n)\; m_{\mu\to n},$$

where $\mu \leq_d \lambda$ is the domination relation on partitions.

This fact follows from the study of the Kotska matrix elements $K_{\lambda,\mu\to n}$, see [Mac95, §1.6, in particular the example 4. (c)]. It can also be shown directly by expanding $e_{\lambda\uparrow n}$ on a sufficient number of variables and collecting the monomials; this simpler proof explains the appearance of binomial coefficients $\binom{n}{k}$. For instance,

$$\begin{aligned} m_{2,1\to n} &= e_{2,1\uparrow n} - 3\, e_{3\uparrow n} - (n-3)\, e_{1,1\uparrow n} + (2n-8)\, e_{2\uparrow n} + (2n-5)\, e_{1\uparrow n} - n(n-4)\, e_{\uparrow n}, \\ e_{2,1\uparrow n} &= \frac{n(n-1)(n-2)}{2}\, m_{\to n} + \frac{(n-2)(3n-7)}{2}\, m_{1\to n} + (3n-10)\, m_{1,1\to n} + 3\, m_{1,1,1\to n} \\ &\quad + (n-3)\, m_{2\to n} + m_{2,1\to n}. \end{aligned}$$

In the following, $N_{\lambda,n} = N_{\lambda\uparrow n}$ if $|\lambda| \leq n$, and 0 otherwise. Because of the existence of the projective limits M_λ, we know that $N_{\lambda,n} * N_{\mu,n} = \sum_\nu g^\nu_{\lambda\mu}\, N_{\nu,n}$, where the sum is not restricted. But on the other hand, by using Proposition 5 and the second identity in Proposition 8, one sees that

$$N_{\lambda,n} * N_{\mu,n} = \sum_{|\rho|\leq|\lambda|,\ |\sigma|\leq|\mu|} h^{\rho\sigma}_{\lambda\mu}(n)\;\; \Gamma_{\rho,n} * \Gamma_{\sigma,n}, \quad \text{with the } h^{\rho\sigma}_{\lambda\mu}(n) \in \mathbb{Q}[n,q,q^{-1}].$$

Because of the result of Francis and Wang, the latter sum may be written as $\sum_{|\tau|\leq|\lambda|+|\mu|} i^\tau_{\lambda\mu}(n)\, \Gamma_{\tau,n}$, and by using the first identity of Proposition 8, one has finally

$$N_{\lambda,n} * N_{\mu,n} = \sum_{|\nu|\leq|\lambda|+|\mu|} j^\nu_{\lambda\mu}(n)\, N_{\nu,n}, \quad \text{with the } j^\nu_{\lambda\mu}(n) \in \mathbb{Q}[n,q,q^{-1}].$$

From this, it can be shown that the first sum $\sum_\nu g^\nu_{\lambda\mu}\, N_{\nu,n}$ is in fact restricted on partitions $|\nu|$ such that $|\nu| \leq |\lambda| + |\mu|$, and because the projections separate the vectors of $\mathscr{D}'_{\infty,q}$, this implies that $M_\lambda * M_\mu = \sum_{|\nu|\leq|\lambda|+|\mu|} g^\nu_{\lambda\mu}\, M_\nu$, so the last part of Theorem 7 is proved. Finally, by reversing the argument, one sees that the $a^\nu_{\lambda\mu}(n,q)$ are in $\mathbb{Q}[n](q)$:

$$\begin{aligned} \Gamma_{\lambda,n} * \Gamma_{\mu,n} &= (q/(q-1))^{|\lambda|+|\mu|} \sum_{\rho,\sigma} P_{\lambda\rho}(n)\, P_{\mu\sigma}(n)\;\, N_{\rho,n} * N_{\sigma,n} \\ &= (q/(q-1))^{|\lambda|+|\mu|} \sum_{\rho,\sigma,\tau} P_{\lambda\rho}(n)\, P_{\mu\sigma}(n)\, g^\tau_{\rho\sigma}\, N_{\tau,n} \\ &= \sum_{\rho,\sigma,\tau,\nu} (q/(q-1))^{|\lambda|+|\mu|-|\nu|}\, P_{\lambda\rho}(n)\, P_{\mu\sigma}(n)\, g^\tau_{\rho\sigma}(q)\, Q_{\tau\nu}(n)\, \Gamma_{\nu,n} = \sum_\nu a^\nu_{\lambda\mu}(n,q)\, \Gamma_{\nu,n} \end{aligned}$$

with $a^{\nu}_{\lambda\mu}(n,q) = (q/(q-1))^{|\lambda|+|\mu|-|\nu|}\,(P^{\otimes 2}(n)\,g(q)\,Q(n))^{\nu}_{\lambda\mu}$ in tensor notation. And since the Γ_λ are known to be defined over $\mathbb{Z}[q,q^{-1}]$, the coefficients $a^{\nu}_{\lambda\mu}(n,q) \in \mathbb{Q}[n](q)$ are in fact[(v)] in $\mathbb{Q}[n,q,q^{-1}]$. Using this technique, one can for instance show that

$$(\Gamma_{(1),n})^2 = \frac{n(n-1)}{2}\,q\,\Gamma_{(0),n} + (n-1)\,(q-1)\,\Gamma_{(1),n} + (q+q^{-1})\,\Gamma_{(1,1),n} + (q+1+q^{-1})\,\Gamma_{(2),n},$$

and this is because $m_{1\to n} = e_{1\uparrow n} - n\,e_{\uparrow n}$ and $e_{1\uparrow n} = n\,m_{\to n} + m_{1\to n}$. Let us conclude by two remarks. First, the reader may have noticed that we did not construct generic conjugacy classes $F_\lambda \in \mathscr{A}_{\infty,q}$ such that $\mathrm{pr}_{\infty,n}(F_\lambda) = \Gamma_{\lambda,n}$; since the Geck-Rouquier elements themselves are difficult to describe, we had little hope to obtain simple generic versions of these Γ_λ. Secondly, the Ivanov-Kerov projective limits of other group algebras — *e.g.*, the algebras of the finite reductive Lie groups $\mathrm{GL}(n,\mathbb{F}_q)$, $\mathrm{U}(n,\mathbb{F}_{q^2})$, etc. — have not yet been studied. It seems to be an interesting open question.

References

[FH59] H. Farahat and G. Higman. The centers of symmetric group rings. *Proc. Roy. Soc. London (A)*, 250:212–221, 1959.

[Fra99] A. Francis. The minimal basis for the centre of an Iwahori-Hecke algebra. *J. Algebra*, 221:1–28, 1999.

[FW09] A. Francis and W. Wang. The centers of Iwahori-Hecke algebras are filtered. *Representation Theory, Comtemporary Mathematics*, 478:29–38, 2009.

[GR97] M. Geck and R. Rouquier. Centers and simple modules for Iwahori-Hecke algebras. In *Finite reductive groups (Luminy, 1994)*, volume 141 of *Progr. Math.*, pages 251–272. Birkhaüser, Boston, 1997.

[IK99] V. Ivanov and S. Kerov. The algebra of conjugacy classes in symmetric groups, and partial permutations. In *Representation Theory, Dynamical Systems, Combinatorial and Algorithmical Methods III*, volume 256 of *Zapiski Nauchnyh Seminarov POMI*, pages 95–120, 1999. English translation available at `arXiv:math/0302203v1 [math.CO]`.

[Jon90] L. Jones. Centers of generic Hecke algebras. *Trans. Amer. Math. Soc.*, 317:361–392, 1990.

[Las06] A. Lascoux. The Hecke algebra and structure constants of the ring of symmetric polynomials, 2006. Available at `arXiv:math/0602379 [math.CO]`.

[Mac95] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. Oxford University Press, 2nd edition, 1995.

[Mat99] A. Mathas. *Iwahori-Hecke algebras and Schur algebras of the symmetric group*, volume 15 of *University Lecture Series*. Amer. Math. Soc., 1999.

(v) They are even in $\mathbb{Q}_{\mathbb{Z}}[n] \otimes \mathbb{Z}[q,q^{-1}]$, where $\mathbb{Q}_{\mathbb{Z}}[n]$ is the $\mathbb{Z}$-module of polynomials with rational coefficients and integer values on integers; indeed, the matrices $M2E$ and $E2M$ have integer entries. It is well known that $\mathbb{Q}_{\mathbb{Z}}[n]$ is spanned over $\mathbb{Z}$ by the binomials $\binom{n}{k}$.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 802–808

An algorithm which generates linear extensions for a generalized Young diagram with uniform probability

Kento NAKADA and Shuji OKAMURA

K. NAKADA : Wakkanai Hokusei Gakuen University, Faculty of Integrated Media. nakada@wakhok.ac.jp
S. OKAMURA : Osaka Prefectural College of Technology. okamura@ipc.osaka-pct.ac.jp

Abstract. The purpose of this paper is to present an algorithm which generates linear extensions for a generalized Young diagram, in the sense of D. Peterson and R. A. Proctor, with uniform probability. This gives a proof of a D. Peterson's hook formula for the number of reduced decompositions of a given minuscule elements.

Résumé. Le but de ce papier est présenter un algorithme qui produit des extensions linéaires pour un Young diagramme généralisé dans le sens de D. Peterson et R. A. Proctor, avec probabilité constante. Cela donne une preuve de la hook formule d'un D. Peterson pour le nombre de décompositions réduites d'un éléments minuscules donné.

Keywords: Generalized Young diagrams, Algorithm, linear extension, Kac-Moody Lie algebra

1 Introduction

In [3], C. Greene, A. Nijenhuis, and H. S. Wilf constructed an algorithm which generates standard tableaux for a given Young diagram with uniform probability. This provides a proof of the hook formula [2] for the number of the standard tableaux of a Young diagram, which is originally due to J. S. Frame, G. de B. Robinson, and R. M. Thrall.

As a generalization of the result of [3], the second author constructed an algorithm, in his master's thesis [9], which generates standard tableau of a given generalized Young diagram. Here, a "generalized Young diagram" is one in the sense of D. Peterson and R. A. Proctor. Similary, this result provides a proof of the hook formula for the number of the standard tableaux of a generalized Young diagram. The purpose of this paper is to present the following theorem:

Theorem 1.1 *Let λ be a finite pre-dominant integral weight over a simply-laced Kac-Moody Lie algebra. Let L be a linear extension of the diagram $\mathrm{D}(\lambda)$ of λ. Then the algorithm* A *for $\mathrm{D}(\lambda)$ generates L with the probability:*

$$\mathrm{Prob}_{\mathrm{D}(\lambda)}(L) = \frac{\prod_{\beta\in\mathrm{D}(\lambda)} \mathrm{ht}\,(\beta)}{(\#\mathrm{D}(\lambda))!}.$$

Here, λ is a certain integral weight (see section 5), $\mathrm{D}(\lambda)$ a certain set of positive real roots determined by λ (section 5), linear extension is a certain sequence of elements of $\mathrm{D}(\lambda)$ (section 2 and 5). $\mathrm{Prob}_{\mathrm{D}(\lambda)}(L)$ the probability we get L by the algorithm A for a diagram $\mathrm{D}(\lambda)$ (section 2), and $\mathrm{ht}\,(\beta)$ denotes the height of β.

2 An algorithm for a graph $(\Gamma;\rightarrow)$

Let $\Gamma=(\Gamma;\rightarrow)$ be a finite simple directed acyclic graph, where $\rightarrow$ denotes the adjacency relation of Γ.

Definition 2.1 *Put $d:=\#\Gamma$. A bijection $L:\{1,\cdots,d\}\longrightarrow\Gamma$ is said to be a* linear extension of $(\Gamma;\rightarrow)$ *if:*

$$L(k)\rightarrow L(l) \textit{ implies } k>l, \quad k,l\in\{1,\cdots,d\}.$$

The set of linear extensions of $(\Gamma;\rightarrow)$ *is denoted by* $\mathcal{L}\,(\Gamma;\rightarrow)$.

For a given $v\in\Gamma$, we define a set $\mathrm{H}_\Gamma\,(v)^+$ by:

$$\mathrm{H}_\Gamma\,(v)^+ := \left\{\, v'\in\Gamma \,\middle|\, v\rightarrow v' \,\right\}.$$

For a given Γ, we call the following algorithm the *algorithm* A *for* Γ:

GNW1. Set $i:=0$ and set $\Gamma_0:=\Gamma$.

GNW2. (Now Γ_i has $d-i$ nodes.) Set $j:=1$ and pick a node $v_1\in\Gamma_i$ with the probability $1/(d-i)$.

GNW3. If $\#\mathrm{H}_{\Gamma_i}\,(v_j)^+\neq 0$, pick a node $v_{j+1}\in\mathrm{H}_{\Gamma_i}\,(v_j)^+$ with the probability $1/\#\mathrm{H}_{\Gamma_i}\,(v_j)^+$. If not, go to GNW5.

GNW4. Set $j:=j+1$ and return to GNW3.

GNW5. (Now $\#\mathrm{H}_{\Gamma_i}\,(v_j)^+=0$.) Set $L(i+1):=v_j$ and set $\Gamma_{i+1}:=\Gamma_i\setminus v_j$ (the graph deleted v_j from Γ_i).

GNW6. Set $i:=i+1$. If $i<d$, return to GNW2; if $i=d$, terminate.

Then, by the definition of the algorithm A for Γ, the map $L:i\mapsto L(i)$ generated above is a linear extension of Γ. We denote by $\mathrm{Prob}_\Gamma(L)$ the probability we get $L\in\mathcal{L}\,(\Gamma)$ by the algorithm A.

3 Case of Young diagrams

When we draw a Young diagram, we use *nodes* instead of *cells* like FIGURE 3.1(left) below:

Definition 3.1 *We equip the set $\mathbb{Y}:=\mathbb{N}\times\mathbb{N}$ with the partial order:*

$$(i,j)\leq(i',j') \Longleftrightarrow i\geq i' \ \text{ and } \ j\geq j'.$$

A finite order filter Y of $\mathbb{Y}$ is called a Young diagram.

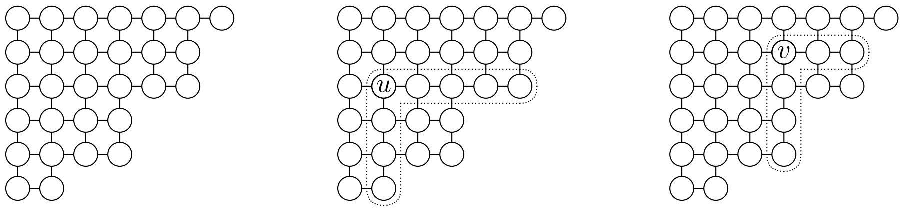


Fig. 3.1: a Young diagram and Hooks of u and v

Definition 3.2 *Let Y be a Young diagram. Let $v = (i, j) \in Y$. We define the subset $\mathrm{H}_Y(v)$ of Y by:*

$$\mathrm{Arm}(v) := \left\{ (i', j') \in Y \,\middle|\, i = i' \text{ and } j < j' \right\}.$$
$$\mathrm{Leg}(v) := \left\{ (i', j') \in Y \,\middle|\, i < i' \text{ and } j = j' \right\}.$$

$$\mathrm{H}_Y(v) := \{v\} \sqcup \mathrm{Arm}(v) \sqcup \mathrm{Leg}(v).$$
$$\mathrm{H}_Y(v)^+ := \mathrm{Arm}(v) \sqcup \mathrm{Leg}(v).$$

The set $\mathrm{H}_Y(v)$ is called the hook of $v \in Y$ *(see FIGURE 3.1(right)).*

For $v, v' \in Y$, we define a relation $v \to v'$ by $v' \in \mathrm{H}_Y(v)^+$. Then $(Y; \to)$ is a finite simple directed acyclic graph.

Then we have the following theorem:

Theorem 3.3 (C. Greene, A. Nijenhuis, and H. S. Wilf [3]) *Let $(Y; \to)$ be a graph defined above for a Young diagram Y. Let $L \in \mathcal{L}(Y; \to)$. Then the algorithm* A *for $(Y; \to)$ generates L with the probability*

$$\mathrm{Prob}_{(Y;\leq)}(L) = \frac{\prod_{v \in Y} \#\mathrm{H}_Y(v)}{\#Y!}. \tag{3.1}$$

Since the right hand side of (3.1) is independent from the choice of $L \in \mathcal{L}(Y; \to)$, we have:

Corollary 3.4 *Let $(Y; \to)$ be a graph for a Young diagram Y. Then we have:*

$$\#\mathcal{L}(Y; \to) = \frac{\#Y!}{\prod_{v \in Y} \#\mathrm{H}_Y(v)}.$$

This gives a proof of hook length formula [2] for the number of standard tableaux for a Young diagram.

4 Case of shifted Young diagrams

Definition 4.1 *We equip the $\mathbb{S} := \left\{ (i, j) \in \mathbb{N} \times \mathbb{N} \,\middle|\, i \leq j \right\}$ with the partial order:*

$$(i, j) \leq (i', j') \Longleftrightarrow i \geq i' \text{ and } j \geq j'.$$

A finite order filter S of $\mathbb{S}$ is called a shifted Young diagram.

Definition 4.2 *Let S be a shifted Young diagram. Let $v = (i, j) \in S$. We define the subset $\mathrm{H}_S(v)$ of S by:*

$$\begin{aligned}
\mathrm{Arm}_S(v) &:= \left\{ (i', j') \in S \,\middle|\, i = i' \text{ and } j < j' \right\}. \\
\mathrm{Leg}_S(v) &:= \left\{ (i', j') \in S \,\middle|\, i < i' \text{ and } j = j' \right\}. \\
\mathrm{Tail}_S(v) &:= \left\{ (i', j') \in S \,\middle|\, j + 1 = i' \text{ and } j < j' \right\}.
\end{aligned}$$

$$\begin{aligned}
\mathrm{H}_S(v) &:= \{v\} \sqcup \mathrm{Arm}_S(v) \sqcup \mathrm{Leg}_S(v) \sqcup \mathrm{Tail}_S(v). \\
\mathrm{H}_S(v)^+ &:= \mathrm{Arm}_S(v) \sqcup \mathrm{Leg}_S(v) \sqcup \mathrm{Tail}_S(v).
\end{aligned}$$

The set $\mathrm{H}_S(v)$ is called the hook of $v \in S$ *(see FIGURE 4.1).*

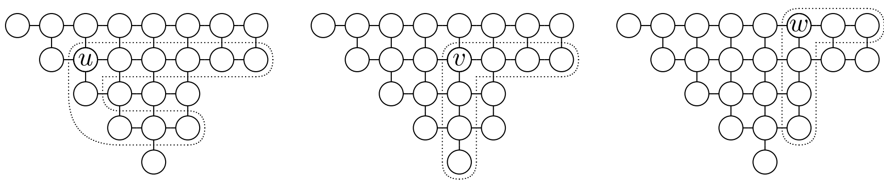


Fig. 4.1: Hooks of u, v , and w.

For $v, v' \in Y$, we define a relation $v \to v'$ by $v' \in \mathrm{H}_Y(v)^+$. Then $(Y; \to)$ is a finite simple directed acyclic graph.

Then we have the following theorem:

Theorem 4.3 (B. E. Sagan [11]) *Let $(S; \to)$ be a graph defined above for a shifted Young diagram S. Let $L \in \mathcal{L}(S; \to)$. Then the algorithm* A *for $(S; \to)$ generates L with the probability*

$$\mathrm{Prob}_{(Y;\leq)}(L) = \frac{\prod_{v \in S} \#\mathrm{H}_S(v)}{\#S!}. \tag{4.1}$$

Since the right hand side of (4.1) is independent from the choice of $L \in \mathcal{L}(S; \to)$, we have:

Corollary 4.4 *Let $(S; \to)$ be a graph for a shifted Young diagram S. Then we have:*

$$\#\mathcal{L}(S; \to) = \frac{\#S!}{\prod_{v \in S} \#\mathrm{H}_S(v)}.$$

This gives a proof of hook length formula [12] for the number of standard tableaux for a shifted Young diagram.

5 General case

In this section, we fix a simply-laced Kac-Moody Lie algebra $\mathfrak{g}$ with a simple root system $\Pi = \left\{ \alpha_i \,\middle|\, \in I \right\}$. For all undefined terminology in this section, we refer the reader to [4] [5].

Definition 5.1 *An integral weight λ is said to be* pre-dominant *if:*

$$\langle \lambda, \beta^\vee \rangle \geq -1 \quad \textit{for each } \beta^\vee \in \Phi_+^\vee,$$

where $\Phi_+^\vee$ denotes the set of positive real coroots. The set of pre-dominant integral weights is denoted by $P_{\geq -1}$. For $\lambda \in P_{\geq -1}$, we define the set $\mathrm{D}(\lambda)$ by:

$$\mathrm{D}(\lambda) := \left\{ \beta \in \Phi_+ \,\middle|\, \langle \lambda, \beta^\vee \rangle = -1 \right\}.$$

The set $\mathrm{D}(\lambda)$ is called the diagram of λ. *If $\#\mathrm{D}(\lambda) < \infty$, then λ is called* finite.

5.1 Hooks

Definition 5.2 *Let $\lambda \in P_{\geq -1}$ and $\beta \in \mathrm{D}(\lambda)$. We define the set $\mathrm{H}_\lambda(\beta)$ by:*

$$\mathrm{H}_\lambda(\beta) := \mathrm{D}(\lambda) \cap \Phi(s_\beta),$$
$$\mathrm{H}_\lambda(\beta)^+ := \mathrm{H}_\lambda(\beta) \setminus \{\beta\}.$$

where $\Phi(s_\beta)$ denotes the inversion set of the reflection corresponding to β:

$$\Phi(s_\beta) = \left\{ \gamma \in \Phi_+ \,\middle|\, s_\beta(\gamma) < 0 \right\}.$$

Proposition 5.3 (see [6],[8]) *Let $\lambda \in P_{\geq -1}$ be finite and $\beta^\vee \in \mathrm{D}(\lambda)$. Then we have:*

1. $\#\mathrm{H}_\lambda(\beta) = \mathrm{ht}(\beta)$.
2. *If $\gamma \in \mathrm{H}_\lambda(\beta)$, then $\gamma \leq \beta$.*

By proposition 5.3 (2), defining $\beta \to \gamma$ by $\gamma \in \mathrm{H}_\lambda(\beta)^+$, the graph $(\mathrm{D}(\lambda); \to)$ is acyclic.

5.2 Main Theorem and Corollaries

Theorem 5.4 (see [8][9]) *Let $\lambda \in P_{\geq -1}$ be finite. Let $L \in \mathcal{L}(\mathrm{D}(\lambda); \to)$. Then the algorithm* A *for $(\mathrm{D}(\lambda); \to)$ generates L with the probability*

$$\mathrm{Prob}_{(\mathrm{D}(\lambda);\to)}(L) = \frac{\prod_{\beta \in \mathrm{D}(\lambda)} \mathrm{ht}(\beta)}{\#\mathrm{D}(\lambda)!}. \tag{5.1}$$

Remark 5.5 *The original statement of theorem 5.4 was proved by the second author [9]. The proof in [9] was done by case-by case argument. On the other hand, the proof in [8] is given by an application of the very special case of the colored hook formula [6].*

Since the right hand side of (5.1) is independent from the choice of $L \in \mathcal{L}(\mathrm{D}(\lambda); \to)$, we have:

Corollary 5.6 *Let $\lambda \in P_{\geq -1}$ be finite. Then we have:*

$$\#\mathcal{L}(\mathrm{D}(\lambda); \to) = \frac{\#\mathrm{D}(\lambda)}{\prod_{\beta \in \mathrm{D}(\lambda)} \mathrm{ht}(\beta)}.$$

Corollary 5.6 gives a proof of Peterson's hook formula for the number of reduced decompositions of minuscule [1][6] element, in simply-laced case. Another proof of Peterson's hook formula is given in [6].

Remark 5.7 *The finite pre-dominant integral weights λ are identified with the minuscule elements w [6]. And, we have $\mathrm{D}(\lambda) = \Phi(w)$ [6]. Furthermore, the linear extensions of $\mathrm{D}(\lambda)$ are identified with the reduced decompositions of w [6] by the following one-to-one correspondence:*

$$\mathrm{Red}(w) \ni (s_{i_1}, s_{i_2}, \cdots, s_{i_d}) \longleftrightarrow L \in \mathcal{L}(\mathrm{D}(\lambda)), \qquad L(k) = s_{i_1} \cdots s_{i_{k-1}}(\alpha_{i_k}) \in \mathrm{D}(\lambda)\ (k = 1, \cdots d),$$

where $\mathrm{Red}(w)$ denotes the set of reduced decompositions of w, $d = \ell(w)$ the length of w. Hence, corollary 5.6 is equivalent with the Peterson's hook formula:

$$\#\mathrm{Red}(w) = \frac{\ell(w)!}{\prod_{\beta \in \Phi(w)} \mathrm{ht}(\beta)}.$$

Remark 5.8 *A Young diagram is realized as a diagram for some pre-dominant integral weight over a root system of type A. Similarly, a shifted Young diagram is realized as a diagram for some pre-dominant integral weight over a root system of type D.*

There are 15 classes of generalized Young diagrams (over simply-laced Kac-Moody Lie algebras). We note that many of them are realized over root systems of indefinite types (see [10]).

Remark 5.9 *The first author has also succeeded in proving Theorem 1.1 in the case of a root system of type B by a certain similar algorithm [7].*

References

[1] J. B. Carrell, *Vector fields, flag varieties and Schubert calculus,* Proc. Hyderabad Conference on Algebraic Groups (ed. S.Ramanan), Manoj Prakashan, Madras, 1991.

[2] J. S. Frame, G. de B. Robinson, and R. M. Thrall, *The hook graphs of symmetric group,* Canad. J. Math. **6** (1954),316-325.

[3] C. Greene, A. Nijenhuis, and H. S. Wilf, *A probabilistic proof of a formula for the number of Young tableaux of a given shape*, Adv. in Math. **31** (1979), 104-109.

[4] V. G. Kac, "Infinite Dimentional Lie Algebras," Cambridge Univ. Press, Cambridge, UK, 1990.

[5] R. V. Moody and A. Pianzola, "Lie Algebras With Triangular Decompositions," Canadian Mathematical Society Series of Monograph and Advanced Text, 1995.

[6] K. Nakada, *Colored hook formula for a generalized Young diagram,* Osaka J. of Math. **Vol.45 No.4** (2008), 1085-1120.

[7] K. Nakada, *Another proof of hook formula for a shifted Young diagram*, in preparation.

[8] K. Nakada and S. Okamura, *Uniform generation of standard tableaux of a generalized Young diagram*, preprint.

[9] S. Okamura, *An algorithm which generates a random standard tableau on a generalized Young diagram* (in Japanese), master's thesis, Osaka university, 2003.

[10] R. A. Proctor, *Dynkin diagram classification of λ-minuscule Bruhat lattices and of d-complete posets,* J.Algebraic Combin. **9** (1999), 61-94.

[11] B. E. Sagan, *On selecting a random shifted Young tableaux*, J. Algorithm **1** (1980), 213-234.

[12] R. M. Thrall, *A combinatorial problem,* Mich.Math.J. **1** (1952), 81-88.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 809–820

On γ-vectors satisfying the Kruskal-Katona inequalities

E. Nevo[1†] and T. K. Petersen[2]

[1] *Department of Mathematics, Cornell University, Ithaca USA*
[2] *Department of Mathematical Sciences, DePaul University, Chicago USA*

Abstract. We present examples of flag homology spheres whose γ-vectors satisfy the Kruskal-Katona inequalities. This includes several families of well-studied simplicial complexes, including Coxeter complexes and the simplicial complexes dual to the associahedron and to the cyclohedron. In these cases, we construct explicit flag simplicial complexes whose f-vectors are the γ-vectors in question, and so a result of Frohmader shows that the γ-vectors satisfy not only the Kruskal-Katona inequalities but also the stronger Frankl-Füredi-Kalai inequalities. In another direction, we show that if a flag $(d-1)$-sphere has at most $2d+3$ vertices its γ-vector satisfies the Frankl-Füredi-Kalai inequalities. We conjecture that if Δ is a flag homology sphere then $\gamma(\Delta)$ satisfies the Kruskal-Katona, and further, the Frankl-Füredi-Kalai inequalities. This conjecture is a significant refinement of Gal's conjecture, which asserts that such γ-vectors are nonnegative.

Résumé. Nous présentons des exemples de sphères d'homologie flag dont γ-vecteurs satisfaire les inégalités de Kruskal-Katona. Cela comprend plusieurs familles de bien étudiés simplicial complexes, y compris les complexes de Coxeter et les complexes simpliciaux dual de l'associahedron et à la cyclohedron. Dans ces cas, nous construisons explicite flag simplicial complexes dont f-vecteurs sont les γ-vecteurs en question, et ainsi de suite de Frohmader montre que le γ-vecteurs de satisfaire non seulement les inégalités de Kruskal-Katona mais aussi la plus fortes inégalités Frankl-Füredi-Kalai. Dans une autre direction, nous montrons que, si un flag $(d-1)$-sphère a au plus $2d+3$ ses sommets γ-vecteur satisfait aux inégalités de Frankl-Füredi-Kalai. Nous conjecture que, si Δ est une sphère d'homologie flag alors $\gamma(\Delta)$ satisfait aux inégalités de Kruskal-Katona, en outre, les de Frankl-Füredi-Kalai. Cette conjecture est un raffinement significative de la conjecture de Gal, qui affirme que ces γ-vecteurs sont nonnégatifs.

Keywords: simplicial complex, Coxeter complex, associahedron, Gal's conjecture, γ-vector

1 Introduction

In [Ga] Gal gave counterexamples to the real-root conjecture for flag spheres and conjectured a weaker statement which still implies the Charney-Davis conjecture. The conjecture is phrased in terms of the so-called γ-*vector*.

Conjecture 1.1 (Gal) *[Ga, Conjecture 2.1.7] If Δ is a flag homology sphere then $\gamma(\Delta)$ is nonnegative.*

† Partially supported by NSF Award DMS-0757828

This conjecture is known to hold for the order complex of a Gorenstein* poset [Kar], all Coxeter complexes (see [Ste], and references therein), and for the (dual simplicial complexes of the) "chordal nestohedra" of [PoRWi]—a class containing the associahedron, permutahedron, and other well-studied polytopes.

If Δ has a nonnegative γ-vector, one may ask what these nonnegative integers count. In certain cases (the type A Coxeter complex, say), the γ-vector has a very explicit combinatorial description. We will exploit such descriptions to show that not only are these numbers nonnegative, but they satisfy certain non-trivial inequalities known as the *Kruskal-Katona inequalities*. Put another way, such a γ-vector is the f-vector of a simplicial complex. Our main result is the following.

Theorem 1.2 *The γ-vector of Δ satisfies the Kruskal-Katona inequalities for each of the following classes of flag spheres:*

(a) Δ is a Coxeter complex.

(b) Δ is the simplicial complex dual to an associahedron.

(c) Δ is the simplicial complex dual to a cyclohedron (type B associahedron).

Note that the type A Coxeter complex is dual to the permutahedron, and for types B and D there is a similarly defined polytope—the "Coxeterhedron" of Reiner and Ziegler [RZ].

We prove Theorem 1.2 by constructing, for each such Δ, a simplicial complex whose faces correspond to the combinatorial objects enumerated by $\gamma(\Delta)$.

In a different direction, we are also able to show that if Δ is a flag sphere with few vertices relative to its dimension, then its γ-vector satisfies the Kruskal-Katona inequalities.

Theorem 1.3 *Let Δ be a $(d-1)$-dimensional flag homology sphere with at most $2d+3$ vertices, i.e., with $\gamma_1(\Delta) \leq 3$. Then $\gamma(\Delta)$ satisfies the Kruskal-Katona inequalities. Moreover, all possible γ-polynomials with $\gamma_1 \leq 3$ that satisfy the Kruskal-Katona inequalities, except for $1 + 3t + 3t^2$, occur as $\gamma(\Delta; t)$ for some flag sphere Δ.*

The proof of Theorem 1.3 can be found in [NPe]. It characterizes the structure of such flag spheres.

Computer evidence suggests that Theorems 1.2 and 1.3 may be enlarged significantly. We make the following strengthening of Gal's conjecture.

Conjecture 1.4 *If Δ is a flag homology sphere then $\gamma(\Delta)$ satisfies the Kruskal-Katona inequalities.*

This conjecture is true, but not sharp, for flag homology 3- (or 4-) spheres. Indeed, Gal showed that $0 \leq \gamma_2(\Delta) \leq \gamma_1(\Delta)^2/4$ must hold for flag homology 3- (or 4-) spheres [Ga], which implies the Kruskal-Katona inequality $\gamma_2(\Delta) \leq \binom{\gamma_1(\Delta)}{2}$. Our stronger Conjecture 5.3 is sharp for flag homology spheres of dimension at most 4.

In Section 2 we review some key definitions. Section 3 collects some known results describing the combinatorial objects enumerated by the γ-vectors of Theorem 1.2. Section 4 constructs simplicial complexes based on these combinatorial objects and proves Theorem 1.2. Finally, Section 5 describes a strengthening of Theorem 1.2 by showing that under the same hypotheses the stronger Frankl-Füredi-Kalai inequalities hold for the γ-vector. A stronger companion to Conjecture 1.4 is also presented, namely Conjecture 5.3.

This paper is an abridged version of [NPe]. Full definitions and proofs can be found there.

2 Terminology

We assume the reader has a basic familiarity with abstract simplicial complexes.

We say that Δ is *flag* if all the minimal subsets of V which are not in Δ have size 2; equivalently $F \in \Delta$ if and only if all the *edges* of F (two element subsets) are in Δ.

The *f-polynomial* of a $(d-1)$-dimensional simplicial complex Δ is the generating function for the dimensions of the faces of the complex:

$$f(\Delta; t) := \sum_{F \in \Delta} t^{\dim F + 1} = \sum_{0 \leq i \leq d} f_i(\Delta) t^i.$$

The *f-vector*

$$f(\Delta) := (f_0, f_1, \ldots, f_d)$$

is the sequence of coefficients of the f-polynomial.

The *h-polynomial* of Δ is a transformation of the f-polynomial:

$$h(\Delta; t) := (1-t)^d f(\Delta; t/(1-t)) = \sum_{0 \leq i \leq d} h_i(\Delta) t^i,$$

and the *h-vector* is the corresponding sequence of coefficients,

$$h(\Delta) := (h_0, h_1, \ldots, h_d).$$

Though they contain the same information, often the h-polynomial is easier to work with than the f-polynomial. For instance, if Δ is a homology sphere, then the *Dehn-Sommerville relations* guarantee that the h-vector is symmetric, i.e., $h_i = h_{d-i}$ for all $0 \leq i \leq d$.

When referring to the f- or h-polynomial of a simple polytope, we mean the f- or h-polynomial of the boundary complex of its dual. So, for instance, we refer to the h-vector of the type A Coxeter complex and the permutahedron interchangeably.

Whenever a polynomial of degree d has symmetric integer coefficients, it has an integer expansion in the basis $\{t^i(1+t)^{d-2i} : 0 \leq i \leq d/2\}$. Specifically, if Δ is a $(d-1)$-dimensional homology sphere then there exist integers $\gamma_i(\Delta)$ such that

$$h(\Delta; t) = \sum_{0 \leq i \leq d/2} \gamma_i(\Delta) t^i (1+t)^{d-2i}.$$

We refer to the sequence $\gamma(\Delta) := (\gamma_0, \gamma_1, \ldots)$ as the *γ-vector* of Δ, and the corresponding generating function $\gamma(\Delta; t) = \sum \gamma_i t^i$ is the *γ-polynomial*. Our goal is to show that under the hypotheses of Theorems 1.2 and 1.3 the γ-vector for Δ is seen to be the f-vector for some other simplicial complex.

A result of Schützenberger, Kruskal and Katona (all independently), characterizes the f-vectors of simplicial complexes. (See [Sta, Ch. II.2].) By convention we call the conditions characterizing these f-vectors the *Kruskal-Katona inequalities*.

We will use the Kruskal-Katona inequalities directly for Theorem 1.3 and for checking the Coxeter complexes of exceptional type in part (a) of Theorem 1.2. (See Table 1.) For the remainder of Theorem 1.2 we construct explicit simplicial complexes with the desired f-vectors.

W	$\gamma(W)$
E_6	$(1, 1266, 7104, 3104)$
E_7	$(1, 17628, 221808, 282176)$
E_8	$(1, 881744, 23045856, 63613184, 17111296)$
F_4	$(1, 232, 208)$
G_2	$(1, 8)$
H_3	$(1, 56)$
H_4	$(1, 2632, 3856)$
$I_2(m)$	$(1, 2m-4)$

Tab. 1: The γ-vectors for finite Coxeter complexes of exceptional type.

3 Combinatorial descriptions of γ-nonnegativity

Here we provide combinatorial descriptions (mostly already known) for the γ-vectors of the complexes described in Theorem 1.2.

3.1 Type A Coxeter complex

We begin by describing the combinatorial objects enumerated by the γ-vector of the type A_{n-1} Coxeter complex, or equivalently, the permutahedron. (For the reader looking for more background on the Coxeter complex itself, we refer to [H, Section 1.15]; for the permutahedron see [Z, Example 0.10].)

Recall that a *descent* of a permutation $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ is a position $i \in [n-1]$ such that $w_i > w_{i+1}$. A *peak* (resp. *valley*) is a position $i \in [2, n-1]$ such that $w_{i-1} < w_i > w_{i+1}$ (resp. $w_{i-1} > w_i < w_{i+1}$). We let $des(w)$ denote the number of descents of w, and we let $peak(w)$ denote the number of peaks. It is well known that the h-polynomial of the type A_{n-1} Coxeter complex is expressed as:

$$h(A_{n-1}; t) = \sum_{w \in \mathfrak{S}_n} t^{des(w)}.$$

Foata and Schützenberger were the first to demonstrate the γ-nonnegativity of this polynomial (better known as the *Eulerian polynomial*), showing $h(A_{n-1}; t) = \sum \gamma_i t^i (1+t)^{n-1-2i}$, where γ_i = the number of equivalence classes of permutations of n with $i+1$ peaks [FoSch]. (Two permutations are in the same equivalence class if they have the same sequence of values at their peaks and valleys.) See also Shapiro, Woan, and Getu [ShWoGe] and, in a broader context, Brändén [B] and Stembridge [Ste].

Following Postnikov, Reiner, and Williams [PoRWi], we choose the following set of representatives for these classes:

$$\widehat{\mathfrak{S}}_n = \{w \in \mathfrak{S}_n : w_{n-1} < w_n, \text{ and if } w_{i-1} > w_i \text{ then } w_i < w_{i+1}\}.$$

In other words, $\widehat{\mathfrak{S}}_n$ is the set of permutations w with no double descents and no final descent, or those for which $des(w) = peak(0w0) - 1$. We now phrase the γ-nonnegativity of the type A_{n-1} Coxeter complex in this language.

Theorem 3.1 (Foata-Schützenberger) *[FoSch, Théorème 5.6] The h-polynomial of the type A_{n-1} Coxeter complex can be expressed as follows:*

$$h(A_{n-1};t) = \sum_{w \in \widehat{\mathfrak{S}}_n} t^{des(w)}(1+t)^{n-1-2des(w)}.$$

We now can state precisely that the type A_{n-1} Coxeter complex (permutahedron) has $\gamma(A_{n-1}) = (\gamma_0, \gamma_1, \dots, \gamma_{\lfloor \frac{n-1}{2} \rfloor})$, where

$$\gamma_i(A_{n-1}) = |\{w \in \widehat{\mathfrak{S}}_n : des(w) = i\}|.$$

The permutahedron is an example of a *chordal nestohedron*. Following [PoRWi], a chordal nestohedron $P_{\mathcal{B}}$ is characterized by its *building set*, $\mathcal{B}$. Each building set $\mathcal{B}$ on $[n]$ has associated to it a set of $\mathcal{B}$-*permutations*, $\mathfrak{S}_n(\mathcal{B}) \subset \mathfrak{S}_n$, and we similarly define $\widehat{\mathfrak{S}}_n(\mathcal{B}) = \mathfrak{S}_n(\mathcal{B}) \cap \widehat{\mathfrak{S}}_n$. See [PoRWi] for details. The following is a main result of Postnikov, Reiner, and Williams [PoRWi].

Theorem 3.2 (Postnikov, Reiner, Williams) *[PoRWi, Theorem 11.6] If $\mathcal{B}$ is a connected chordal building set on $[n]$, then*

$$h(P_{\mathcal{B}};t) = \sum_{w \in \widehat{\mathfrak{S}}_n(\mathcal{B})} t^{des(w)}(1+t)^{n-1-2des(w)}.$$

Thus, for a chordal nestohedron, $\gamma_i(P_{\mathcal{B}}) = |\{w \in \widehat{\mathfrak{S}}_n(\mathcal{B}) : des(w) = i\}|$.

3.2 Type B Coxeter complex

We now turn our attention to the type B_n Coxeter complex. The framework of [PoRWi] no longer applies, so we must discuss a new, if similar, combinatorial model.

In type B_n, the γ-vector is given by $\gamma_i = 4^i$ times the number of permutations w of $\mathfrak{S}_n$ such that $peak(0w) = i$. See Petersen [Pe] and Stembridge [Ste]. We define the set of *decorated permutations* Dec_n as follows. A decorated permutation $\mathbf{w} \in Dec_n$ is a permutation $w \in \mathfrak{S}_n$ with bars following the peak positions (with $w_0 = 0$). Moreover these bars come in four colors: $\{| = |^0, |^1, |^2, |^3\}$. Thus for each $w \in \mathfrak{S}_n$ we have $4^{peak(0w)}$ decorated permutations in Dec_n. For example, Dec_9 includes elements such as

$$4|238|^1 76519, \quad 4|^3 238|^2 76519, \quad 25|137|^1 69|^2 84.$$

(Note that $\widehat{\mathfrak{S}}_n \subset Dec_n$.) Let $peak(\mathbf{w}) = peak(0w)$ denote the number of bars in $\mathbf{w}$. In this context we have the following result.

Theorem 3.3 (Petersen) *[Pe, Proposition 4.15] The h-polynomial of the type B_n Coxeter complex can be expressed as follows:*

$$h(B_n;t) = \sum_{\mathbf{w} \in Dec_n} t^{peak(\mathbf{w})}(1+t)^{n-2peak(\mathbf{w})}.$$

Thus,

$$\gamma_i(B_n) = |\{\mathbf{w} \in Dec_n : peak(\mathbf{w}) = i\}|.$$

3.3 Type D Coxeter complex

We now describe how to view the elements enumerated by the γ-vector of the type D Coxeter complex in terms of a subset of decorated permutations. Define a subset $Dec_n^D \subset Dec_n$ as follows:

$$Dec_n^D = \{\mathbf{w} = w_1 \cdots |^{c_1} w_{i_1} \cdots |^{c_2} \cdots \in Dec_n \text{ such that } w_1 < w_2 < w_3, \text{ or, } \\ \text{both } \max\{w_1, w_2, w_3\} \neq w_3 \text{ and } c_1 \in \{0,1\}\}.$$

In other words, we remove from Dec_n all elements whose underlying permutations have $w_2 < w_1 < w_3$, then for what remains we dictate that bars in the first or second positions can only come in one of two colors. Stembridge [Ste] gives an expression for the h-polynomial of the type D_n Coxeter complex, which we now phrase in the following manner.

Theorem 3.4 (Stembridge) *[Ste, Corollary A.5]. The h-polynomial of the type D_n Coxeter complex can be expressed as follows:*

$$h(D_n; t) = \sum_{\mathbf{w} \in Dec_n^D} t^{peak(\mathbf{w})}(1+t)^{n-2peak(\mathbf{w})}.$$

Thus,

$$\gamma_i(D_n) = |\{\mathbf{w} \in Dec_n^D : peak(\mathbf{w}) = i\}|.$$

3.4 The associahedron

The associahedron $Assoc_n$ is an example of a chordal nestohedron, so Theorem 3.2 applies. Following [PoRWi, Section 10.2], the $\mathcal{B}$-permutations of $Assoc_n$ are precisely the 312-avoiding permutations. Let $\mathfrak{S}_n(312)$ denote the set of all $w \in \mathfrak{S}_n$ such that there is no triple $i < j < k$ with $w_j < w_k < w_i$. Then we have:

$$h(Assoc_n; t) = \sum_{w \in \widehat{\mathfrak{S}}_n(312)} t^{des(w)}(1+t)^{n-1-2des(w)},$$

where $\widehat{\mathfrak{S}}_n(312) = \mathfrak{S}_n(312) \cap \widehat{\mathfrak{S}}_n$. Hence,

$$\gamma_i(Assoc_n) = |\{w \in \widehat{\mathfrak{S}}_n(312) : des(w) = i\}|.$$

3.5 The cyclohedron

The cyclohedron Cyc_n, or type B associahedron, is a nestohedron, though not a chordal nestohedron and hence Theorem 3.2 does not apply. Its γ-vector can be explicitly computed from its h-vector as described in [PoRWi, Proposition 11.15]. We have $\gamma_i(Cyc_n) = \binom{n}{i,i,n-2i}$. Define

$$P_n = \{(L, R) \subseteq [n] \times [n] : |L| = |R|, L \cap R = \emptyset\}.$$

It is helpful to think of elements of P_n as follows. For $\sigma = (L, R)$ with $|L| = |R| = k$, write σ as a $k \times 2$ array with the elements of L written in increasing order in the first column, the elements of R in increasing order in the second column. That is, if $L = \{l_1 < \cdots < l_k\}$ and $R = \{r_1 < \cdots < r_k\}$, we write

$$\sigma = \begin{pmatrix} l_1 & r_1 \\ \vdots & \vdots \\ l_k & r_k \end{pmatrix}.$$

For $\sigma \in P_n$, let $\rho(\sigma) = |L| = |R|$. Then we can write

$$h(Cyc_n; t) = \sum_{\sigma \in P_n} t^{\rho(\sigma)}(1+t)^{n-2\rho(\sigma)}.$$

Thus,

$$\gamma_i(Cyc_n) = |\{\sigma \in P_n : \rho(\sigma) = i\}|.$$

4 The Γ-complexes

We will now describe simplicial complexes whose f-vectors are the γ-vectors described in Section 3.

4.1 Coxeter complexes

Notice that if

$$\mathbf{w} = w_1|^{c_1} \cdots |^{c_{i-1}} w_i|^{c_i} w_{i+1}|^{c_{i+1}} \cdots |^{c_{l-1}} w_l,$$

is a decorated permutation, then each word $w_i = w_{i,1} \dots w_{i,k}$ has some j such that:

$$w_{i,1} > w_{i,2} > \cdots > w_{i,j} > w_{i,j+1} < w_{i,j+2} < \cdots < w_{i,k}.$$

We say w_i is a *down-up word*. We call $\grave{w}_i = w_{i,1} \cdots w_{i,j}$ the *decreasing part* of w_i and $\acute{w}_i = w_{i,j+1} \cdots w_{i,k}$ the *increasing part* of w_i. Note that the decreasing part may be empty, whereas the increasing part is nonempty if $i \neq l$. Also, the rightmost block of $\mathbf{w}$ may be strictly decreasing (in which case $w_l = \grave{w}_l$) and the leftmost block is always increasing, even if it is a singleton.

Define the vertex set

$$V_{Dec_n} := \{\mathbf{v} \in Dec_n : peak(\mathbf{v}) = 1\}.$$

The adjacency of two such vertices is defined as follows. Let

$$\mathbf{u} = \acute{u}_1|^c \grave{u}_2 \acute{u}_2$$

and

$$\mathbf{v} = \acute{v}_1|^d \grave{v}_2 \acute{v}_2$$

be two vertices with $|\acute{u}_1| < |\acute{v}_1|$. We define $\mathbf{u}$ and $\mathbf{v}$ to be adjacent if and only if there is an element $\mathbf{w} \in Dec_n$ such that

$$\mathbf{w} = \acute{u}_1|^c \grave{u}_2 \acute{a}|^d \grave{v}_2 \acute{v}_2,$$

where $\acute{a}$ is the letters of $\acute{u}_2 \cap \acute{v}_1$ written in increasing order. Such an element $\mathbf{w}$ exists if, as sets:

- $\acute{u}_1 \cup \grave{u}_2 \subset \acute{v}_1$ ($\Leftrightarrow \grave{v}_2 \cup \acute{v}_2 \subset \acute{u}_2$),
- $\min \acute{u}_2 \cap \acute{v}_1 < \min \grave{u}_2$, and
- $\max \acute{u}_2 \cap \acute{v}_1 > \max \grave{v}_2$. (Note that $\acute{u}_2 \cap \acute{v}_1$ is nonempty by the first condition.)

Definition 4.1 *Let $\Gamma(Dec_n)$ be the collection of all subsets F of V_{Dec_n} such that every two distinct vertices in F are adjacent.*

Note that by definition $\Gamma(Dec_n)$ is a flag complex. It remains to show that the faces of $\Gamma(Dec_n)$ correspond to decorated permutations.

Let $\phi : Dec_n \to \Gamma(Dec_n)$ be the map defined as follows. If

$$\mathbf{w} = w_1|^{c_1} \cdots |^{c_{i-1}} w_i|^{c_i} w_{i+1}|^{c_{i+1}} \cdots |^{c_{l-1}} w_l,$$

then

$$\phi(\mathbf{w}) = \{w_1|^{c_1} \grave{w}_2 \acute{b}_1, \ldots, \acute{a}_i|^{c_i} \grave{w}_{i+1} \acute{b}_i, \ldots, \acute{a}_{l-1}|^{c_{l-1}} \grave{w}_l \acute{b}_{l-1}\},$$

where $\acute{a}_i$ is the set of letters to the left of $\grave{w}_{i+1}$ in $\mathbf{w}$ written in increasing order and $\acute{b}_i$ is the set of letters to the right of $\grave{w}_{i+1}$ in $\mathbf{w}$ written in increasing order.

Proposition 4.2 *The map ϕ is a bijection between faces of $\Gamma(Dec_n)$ and decorated permutations in Dec_n.*

The proof of Proposition 4.2 is not difficult; it can be found in [NPe].

We now make explicit how to realize Dec_n as the face poset of $\Gamma(Dec_n)$. We say $\mathbf{w}$ covers $\mathbf{u}$ if and only if $\mathbf{u}$ can be obtained from $\mathbf{w}$ by removing a bar $|^{c_i}$ and reordering the word $w_i w_{i+1} = \grave{w}_i \acute{w}_i w_{i+1}$ as a down-up word $\grave{w}_i a$ where a is the word formed by writing the letters of $\acute{w}_i w_{i+1}$ in increasing order. Then $(Dec_n, \leq)$ is a poset graded by number of bars and we have the following result.

Proposition 4.3 *The map ϕ is an isomorphism of graded posets from $(Dec_n, \leq)$ to $(\Gamma(Dec_n), \subseteq)$.*

We now claim that the γ-objects for the type A_{n-1} and type D_n Coxeter complexes form flag subcomplexes of $\Gamma(Dec_n)$. Let $S \in \{\widehat{\mathfrak{S}}_n, Dec_n^D\}$. To show $\Gamma(S)$ is a subcomplex, by Proposition 4.3 it suffices to show that $(S, \leq)$ is a lower ideal in $(Dec_n, \leq)$. To show that $\Gamma(S)$ is flag, we show that it is the flag complex generated by the elements of S with exactly one bar. Both facts are straightforward to verify for either choice of S.

Proposition 4.4 *For $S \in \{\widehat{\mathfrak{S}}_n, Dec_n^D\}$ the image $\Gamma(S) := \phi(S)$ is a flag subcomplex of $\Gamma(Dec_n)$.*

In light of the results of Sections 3.1, 3.2, and 3.3, and because the dimension of faces corresponds to the number of bars, we have the following result, which, along with Table 1 implies part (a) of Theorem 1.2.

Corollary 4.5 *We have:*

1. $\gamma(A_{n-1}) = f(\Gamma(\widehat{\mathfrak{S}}_n))$,
2. $\gamma(B_n) = f(\Gamma(Dec_n)$, *and*
3. $\gamma(D_n) = f(\Gamma(Dec_n^D)$.

In particular, the γ-vectors of the type A_{n-1}, B_n, and D_n Coxeter complexes satisfy the Kruskal-Katona inequalities.

Remark 4.6 *In view of Theorem 3.2, we can observe that if $\mathcal{B}$ is a connected chordal building set such that $(\widehat{\mathfrak{S}}_n(\mathcal{B}), \leq)$ is a lower ideal in $(Dec_n, \leq)$, then a result such as Corollary 4.5 applies. That is, we would have $\gamma(P_{\mathcal{B}}) = f(\phi(\widehat{\mathfrak{S}}_n(\mathcal{B})))$. In particular, we would like to apply such an approach to the γ-vector of the associahedron. However, $\widehat{\mathfrak{S}}_n(312)$ is not generally a lower ideal in Dec_n. For example, with $n = 5$, we have $w = 3|14|25 > 3|1245 = u$. While w is 312-avoiding, u is clearly not.*

4.2 The associahedron

First we give a useful characterization of the set $\widehat{\mathfrak{S}}_n(312)$.

Observation 4.7 *If $w \in \widehat{\mathfrak{S}}_n(312)$, it has the form*

$$w = \acute{a}_1\ j_1 i_1\ \acute{a}_2\ j_2 i_2\ \cdots\ \acute{a}_k\ j_k i_k\ \acute{a}_{k+1}, \tag{1}$$

where:

- $j_1 < \cdots < j_k$,
- $j_s > i_s$ *for all* s*, and*
- $\acute{a}_s$ *is the word formed by the letters of* $\{r \in [n] \setminus \{i_1, j_1, \dots, i_k, j_k\} : j_{s-1} < r < j_s\}$ *(with* $j_0 = 0$*,* $j_{k+1} = n+1$*) written in increasing order.*

In particular, since w has no double descents and no final descent, we see that $\acute{a}_{k+1}$ is always nonempty and $w_n = n$. We refer to (i_s, j_s) as a *descent pair* of w.

Given distinct integers a, b, c, d with $a < b$ and $c < d$, we say the pairs (a, b) and (c, d) are *crossing* if either of the following statements are true:

- $a < c < b < d$ or
- $c < a < d < b$.

Otherwise, we say the pairs are *noncrossing*. For example, $(1, 5)$ and $(4, 7)$ are crossing, whereas both the pairs $(1, 5)$ and $(2, 4)$ and the pairs $(1, 5)$ and $(6, 7)$ are noncrossing.

Define the vertex set

$$V_{\widehat{\mathfrak{S}}_n(312)} := \{(a, b) : 1 \leq a < b \leq n-1\}.$$

Definition 4.8 *Let $\Gamma(\widehat{\mathfrak{S}}_n(312))$ be the collection of subsets F of $V_{\widehat{\mathfrak{S}}_n(312)}$ such that every two distinct vertices in F are noncrossing.*

By definition $\Gamma(\widehat{\mathfrak{S}}_n(312))$ is a flag simplicial complex, and so the task remains to show that the faces of the complex correspond to the elements of $\widehat{\mathfrak{S}}_n(312)$.

Define a map $\pi : \widehat{\mathfrak{S}}_n(312) \to \Gamma(\widehat{\mathfrak{S}}_n(312))$ as follows:

$$\pi(w) = \{(w_{i+1}, w_i) : w_i > w_{i+1}\}.$$

Suppose w is as in (1). We claim that the descent pairs (i_s, j_s) and (i_t, j_t) (with $j_s < j_t$, say) are noncrossing. Indeed, if $i_s < i_t < j_s < j_t$, then the subword $j_s i_s i_t$ forms the pattern 312. Therefore (and because $w_n = n$) we see the map $\pi(w) = \{(i_1, j_1), \dots, (i_k, j_k)\}$ is well-defined. Using Observation 4.7, the following is straightforward to prove.

Proposition 4.9 *The map π is a bijection between faces of $\Gamma(\widehat{\mathfrak{S}}_n(312))$ and $\widehat{\mathfrak{S}}_n(312)$.*

By construction, we have $|\pi(w)| = des(w)$, and therefore the results of Section 3.4 imply the following result, proving part (b) of Theorem 1.2.

Corollary 4.10 *We have:*

$$\gamma(Assoc_n) = f(\Gamma(\widehat{\mathfrak{S}}_n(312))).$$

In particular, the γ-vector of the associahedron satisfies the Kruskal-Katona inequalities.

Remark 4.11 *It is well known that the h-vector of the associahedron has a combinatorial interpretation given by* noncrossing partitions. *Simion and Ullmann [SiU] give a particular decomposition of the lattice of noncrossing partitions that can be used to describe $\gamma(Assoc_n)$ in (essentially) the same way.*

4.3 The cyclohedron

For the cyclohedron, let

$$V_{P_n} := \{(l,r) \in [n] \times [n] : l \neq r\}.$$

Two vertices (l_1, r_1) and (l_2, r_2) are adjacent if and only if:

- l_1, l_2, r_1, r_2 are distinct and
- $l_1 < l_2$ if and only if $r_1 < r_2$.

Define $\Gamma(P_n)$ to be the flag complex whose faces F are all subsets of V_{P_n} such that every two distinct vertices in F are adjacent.

We let $\psi : P_n \to \Gamma(P_n)$ be defined as follows. If

$$\sigma = \begin{pmatrix} l_1 & r_1 \\ \vdots & \vdots \\ l_k & r_k \end{pmatrix}$$

is an element of P_n, then $\psi(\sigma)$ is simply the set of rows of σ:

$$\psi(\sigma) = \{(l_1, r_1), \ldots, (l_k, r_k)\}.$$

Clearly this map is invertible, for we can list a set of pairwise adjacent vertices in increasing order (by l_i or by r_i) to obtain an element of P_n. We have the following.

Proposition 4.12 *The map ψ is a bijection between faces of $\Gamma(P_n)$ and the elements of P_n.*

We are now able to complete the proof of Theorem 1.2, as the following implies part (c).

Corollary 4.13 *We have*

$$\gamma(Cyc_n) = f(\Gamma(P_n)).$$

In particular, the γ-vector of the cyclohedron satisfies the Kruskal-Katona inequalities.

5 Stronger inequalities

A $(d-1)$-dimensional simplicial complex Δ on a vertex set V is *balanced* if there is a coloring of its vertices $c : V \to [d]$ such that for every face $F \in \Delta$ the restriction map $c : F \to [d]$ is injective. That is, every face has distinctly colored vertices.

Frohmader [Fro] proved that the f-vectors of flag complexes form a (proper) subset of the f-vectors of balanced complexes. (This was conjectured earlier by Eckhoff and Kalai, independently.) Further, a characterization of the f-vectors of balanced complexes is known [FraFüKal], yielding stronger upper bounds on f_{i+1} in terms of f_i than the Kruskal-Katona inequalities, namely the *Frankl-Füredi-Kalai inequalities*. For example, a balanced 1-dimensional complex is a bipartite graph, hence satisfies $f_2 \leq f_1^2/4$, while the complete graph has $f_2 = \binom{f_1}{2}$. See [FraFüKal] for the general description of the Frankl-Füredi-Kalai inequalities.

Because the Γ-complexes of Section 4 are flag complexes, Frohmader's result shows that the γ-vectors of Theorem 1.2 satisfy the Frankl-Füredi-Kalai inequalities. The same is easily verified for the γ-vectors given by Theorem 1.3 and in Table 1 for the exceptional Coxeter complexes. We obtain the following strengthening of Theorem 1.2.

Theorem 5.1 *The γ-vector of Δ satisfies the Frankl-Füredi-Kalai inequalities for each of the following classes of flag spheres:*

(a) *Δ is a Coxeter complex.*

(b) *Δ is the simplicial complex dual to an associahedron.*

(c) *Δ is the simplicial complex dual to a cyclohedron.*

(d) *Δ has $\gamma_1(\Delta) \leq 3$.*

Remark 5.2 *The complexes $\Gamma(S)$ where $S \in \{Dec_n, \widehat{\mathfrak{S}}_n, Dec_n^D\}$ are balanced. The color of a vertex v with a peak at position i is $\lceil \frac{i}{2} \rceil$.*

Similarly this suggests the following strengthening of Conjecture 1.4.

Conjecture 5.3 *If Δ is a flag homology sphere then $\gamma(\Delta)$ satisfies the Frankl-Füredi-Kalai inequalities.*

As mentioned in the Introduction, this conjecture is true for flag homology spheres of dimension at most 4. We do not have a counterexample to the following possible strengthening of this conjecture.

Problem 5.4 *If Δ is a flag homology sphere, then $\gamma(\Delta)$ is the f-vector of a flag complex.*

References

[B] P. Brändén, Sign-graded posets, unimodality of W-polynomials and the Charney-Davis conjecture, *Electron. J. Combin.* **11** (2004/06), Research Paper 9, 15pp.

[FraFüKal] P. Frankl, Z. Füredi and G. Kalai, Shadows of colored complexes, *Math. Scand.* **63** (1988), 169–178.

[Fro] A. Frohmader, Face vectors of flag complexes, *Israel J. Math.* **164** (2008), 153–164.

[FoSch] D. Foata and M.-P. Schützenberger, "Théorie géométrique des polynômes eulériens," Lecture Notes in Mathematics, Vol. 138, Springer-Verlag, Berlin, 1970.

[Ga] S. R. Gal, Real root conjecture fails for five- and higher-dimensional spheres, *Discrete Comput. Geom.* **34** (2005), 269–284.

[H] J. E. Humphreys, "Reflection groups and Coxeter groups," Cambridge Univ. Press, Cambridge, 1990.

[Kar] K. Karu, The cd-index of fans and posets, *Compos. Math.* **142** (2006), 701–718.

[M] R. Meshulam, Domination numbers and homology, *J. Combin. Theory Ser. A* **102** (2003), 321–330.

[N] E. Nevo, Higher minors and Van Kampen's obstruction, *Math. Scand.* **101** (2007), 161–176.

[NPe] E. Nevo and T. K. Petersen, On γ-vectors satisfying the Kruskal-Katona inequalities, arXiv: 0909.0694.

[Pe] T. K. Petersen, Enriched P-partitions and peak algebras, *Adv. Math.* **209** (2007), 561–610.

[PoRWi] A. Postnikov, V. Reiner, and L. Williams, Faces of generalized permutohedra, *Doc. Math.* **13** (2008) 207–273.

[RZ] V. Reiner and G. Ziegler, Coxeter-associahedra, *Mathematika* **41** (1994), 364–393.

[ShWoGe] L. W. Shapiro, W. J. Woan, S. Getu, Runs, slides and moments, *SIAM J. Algebraic Discrete Methods* **4** (1983), 459–466.

[SiU] R. Simion and D. Ullman, On the structure of the lattice of noncrossing partitions, *Discrete Math.* **98** (1991), 193–206.

[Sta] R. P. Stanley, "Combinatorics and Commutative Algebra" (2nd ed.), Birkhäuser, Boston, 1996.

[Ste] J. R. Stembridge, Coxeter cones and their h-vectors, Adv. Math. 217 (2008) 1935–1961.

[Z] G. M. Ziegler, "Lectures on polytopes," Graduate Texts in Mathematics, **152**, Springer-Verlag, New York, 1995.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 821–832

On formulas for moments of the Wishart distributions as weighted generating functions of matchings

Yasuhide NUMATA[1,3] and Satoshi KURIKI[2,3]

[1] *Department of Mathematical Informatics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.*
[2] *The Institute of Statistical Mathematics, 10-3 Midoricho, Tachikawa, Tokyo 190-8562, Japan.*
[3] *Japan Science and Technology Agency (JST), CREST.*

Abstract. We consider the real and complex noncentral Wishart distributions. The moments of these distributions are shown to be expressed as weighted generating functions of graphs associated with the Wishart distributions. We give some bijections between sets of graphs related to moments of the real Wishart distribution and the complex noncentral Wishart distribution. By means of the bijections, we see that calculating these moments of a certain class the real Wishart distribution boils down to calculations for the case of complex Wishart distributions.

Résumé. Nous considérons les lois Wishart non-centrale réel et complexe. Les moments sont décrits comme fonctions génératrices de graphes associées avex les lois Wishart. Nous donnons bijections entre ensembles de graphes relatifs aux moments des lois Wishart non-centrale réel et complexe. Au moyen de la bijections, nous voyons que le calcul des moments d'une certaine classe la loi Wishart réel deviennent le calcul de moments de loi Wishart complexes.

Keywords: generating funtion; Hafnian; matching; moments formula; Wishart distribution.

1 Introduction

First we recall the Wishart distributions which originate from the paper by Wishart [18]. Let $X_1 = (x_{i1})_{1\le i\le p}$, $X_2 = (x_{i2})_{1\le i\le p}, \dots, X_\nu = (x_{i\nu})_{1\le i\le p}$ be p-dimensional random column vectors distributed independently according to the normal (Gauss) distribution $N_p(\mu_1,\Sigma),\dots,N_p(\mu_\nu,\Sigma)$ with mean vectors $\mu_1 = (\mu_{i1})_{1\le i\le p},\dots,\mu_\nu = (\mu_{i\nu})_{1\le i\le p}$ (respectively) and a common covariance matrix $\Sigma = (\sigma_{ij})$. The distribution of a $p\times p$ symmetric random matrix $W = (w_{ij})_{1\le i,j\le p}$ defined by $w_{ij} = \sum_{t=1}^{\nu} x_{it}x_{jt}$ is the real noncentral Wishart distribution $W_p(\nu,\Sigma,\Delta)$, where $\Delta = (\delta_{ij})_{1\le ij\le p}$ is the mean square matrix defined by $\delta_{ij} = \sum_{t=1}^{\nu}\mu_{it}\mu_{jt}$. The Wishart distribution for $\Delta = 0$ is said to be central and is denoted by $W_p(\nu,\Sigma)$.

The matrix $\Omega = \Sigma^{-1}\Delta$ is called the noncentrality matrix. It is usually used instead of Δ to parameterize the Wishart distribution. However, in this paper, we use (ν,Σ,Δ) for simplicity in describing formulas. The complex Wishart distribution $CW_p(\nu,\Sigma,\Delta)$ is defined as the distribution of some $p\times p$ Hermitian random matrix constructed from random vectors distributed independently according to the complex normal (complex Gauss) distributions.

The moment generating function of the real Wishart distribution is given by

$$\mathbb{E}[e^{\operatorname{tr}(\Theta W)}] = \det(I - 2\Theta\Sigma)^{-\frac{\nu}{2}} e^{-\frac{1}{2}\operatorname{tr}(I-2\Theta\Sigma)^{-1}\Theta\Delta},$$

where Θ is a $p \times p$ symmetric parameter matrix [13]. Similarly, the moment generating function of the complex Wishart distribution is given as follows:

$$\mathbb{E}[e^{\operatorname{tr}(\Theta W)}] = \det(I - \Theta\Sigma)^{-\nu} e^{-\operatorname{tr}(I-\Theta\Sigma)^{-1}\Theta\Delta},$$

where Θ is a $p \times p$ Hermitian parameter matrix. See also [2] for the central case. Our first objective is to describe the moments $\mathbb{E}[w_{i_1,i_2} w_{i_3,i_4} \cdots w_{i_{2n-1},i_{2n}}]$ of the Wishart distributions of general degrees in explicit forms. Since the Wishart distribution is one of the most important distributions, it has been studied by many researchers not only in the field of mathematical statistics but also in other fields (e.g., [1, 11]). Its moments have been well studied, and methods to calculate the moments in the central cases have been developed by Lu and Richards [10]; Graczyk, Letac, and Massam [3, 4]; Vere-Jones [16]; and many other authors. In particular, Graczyk, Letac and Massam [3, 4] developed a formula for the moments using the representation theory of symmetric group. More recently, Letac and Massam [9] introduced a method to calculate the moments of the noncentral Wishart distributions. In this paper, we introduce another formula for the moments of Wishart distribution; in our formula, the moments are described as special values of the weighted generating function of matchings of graphs. Calculation of the moments boils down to enumeration of graphs via our formulas. As an application of our formulas, we construct some correspondences between some sets of graphs, which implies several identities of moments.

The organization of this paper is as follows. In Section 2, we introduce some notations for the graphs. In Section 3, we define the generating functions of matchings and give the main formulas, which are an extension of Takemura [15] dealing with the central case. In Section 4.1, we give some correspondences between directed and undirected graphs, which implies equations between the moments of the complex Wishart distribution and the moments of the real Wishart distribution for some special parameter. In Sections 4.2 and 4.3, we consider the Wishart distribution with some degenerated parameters. We see that the calculation of its moments is reduced to enumerating graphs satisfying some conditions.

A part of this paper is taken from our previous paper [8]. Please see the paper for the omitted proofs.

2 Notation of graphs

In this paper, we consider both undirected and directed graphs. For $l \in \mathbb{Z}$, we define $\dot{l}$ and $\ddot{l}$ by $\dot{l} = 2l - 1$ and $\ddot{l} = 2l$. Let us fix $n \in \mathbb{Z}_{>0}$. We also fix sets V and V' as follows: $V = [n] = \{ 1, \dots, n \}$, $\dot{V} = [\dot{n}] = \{ \dot{1}, \dots, \dot{n} \}$, $\ddot{V} = [\ddot{n}] = \{ \ddot{1}, \dots, \ddot{n} \}$, and $V' = \dot{V} \amalg \ddot{V} = [\dot{n}] \amalg [\ddot{n}] = [2n]$. We use V and V' as the sets of vertices of directed and undirected graphs, respectively.

First we consider undirected graphs. For $v \neq w$, the undirected edge between v and w is denoted by $\{v, w\} = \{w, v\}$. We do not consider undirected self loops, i.e., $\{v, v\}$. For sets W' and U' of vertices, we define sets $K'_{U'}$ and $K'_{W',U'}$ of undirected edges by $K'_{W',U'} = \{ \{w, u\} \mid w \in W', u \in U', w \neq u \}$, $K'_{U'} = K'_{U',U'} = \{ \{v, u\} \mid v \neq u \in V' \}$. We call a pair $G' = (V', E')$ of a finite set V' and a subset $E' \subset K'_{V'}$ an undirected graph. For an undirected graph $G' = (V', E')$, we define $\operatorname{vertex}(E')$ by $\operatorname{vertex}(E') = \{ v \in V' \mid \{v, u\} \in E' \text{ for some } u \in V' \}$. Let (V', K') be an undirected graph. We call a subset $E' \subset K'$ a matching in (V', K') if no two edges in E' share a common vertex. We define $\mathcal{M}'(V', K')$ to be the set of matchings in (V', K') and $\mathcal{M}'(V')$ to be the set $\mathcal{M}'(V', K'_{V'})$ of matchings in the complete graph $(V', K'_{V'})$. A matching E' in (V', K') is said to be perfect if $\mathrm{vertex}(E') = V'$. We define $\mathcal{P}'(V', K')$ to be the set of perfect matchings in (V', K') and $\mathcal{P}'(V')$ by $\mathcal{P}'(V') = \mathcal{P}'(V', K'_{V'})$.

Next we consider directed graphs. A directed edge from v to w is denoted by (v, u). For $v \neq u$, $(v, u) \neq (u, v)$. In the case of directed graphs, we also consider a directed self loop (v, v). For a directed edge $e = (v, u)$, we respectively call v and u a starting and end points of e. For sets W and U of vertices, we define sets K_U and $K_{W,U}$ of directed edges by $K_{W,U} = \{ (v, u) \mid v \in W, u \in U \}$, $K_U = K_{U,U} = \{ (v, u) \mid v, u \in U \}$. We call a pair $G = (V, E)$ of a finite set V and a subset $E \subset K_V$ a directed graph. For a directed graph $G = (V, E)$, we define $\mathrm{start}(E)$ and $\mathrm{end}(E)$ by

$$\mathrm{start}(E) = \{ v \in V \mid\ (v, u) \in E \text{ for some } u \in V \},$$
$$\mathrm{end}(E) = \{ u \in V \mid\ (v, u) \in E \text{ for some } v \in V \}.$$

For a directed graph (V, K), we call a subset $E \subset K$ a matching in (V, K) if $v \neq v'$ and $u \neq u'$ for any two distinct directed edges (v, u) and $(v', u') \in E$. We define $\mathcal{M}(V, K)$ to be the set of matchings in (V, K) and $\mathcal{M}(V)$ by $\mathcal{M}(V) = \mathcal{M}(V, K_V)$. A matching E in (V, E) is said to be perfect if $\mathrm{start}(E) = V$ and $\mathrm{end}(E) = V$. We define $\mathcal{P}(V, K)$ to be the set of perfect matchings in (V, K) and $\mathcal{P}(V)$ by $\mathcal{P}(V) = \mathcal{P}(V, K_V)$.

Remark 2.1 *We can identify a directed graph (V, E) with a bipartite graph $(\dot{V}, \ddot{V}, \{ \{\dot{v}, \ddot{u}\} \mid (v, u) \in E \})$. i.e., a graph whose edges connect a vertex in $\dot{V}$ to a vertex in $\ddot{V}$. Via this identification, an element in $\mathcal{M}(V, K)$ is identified with a matching in the bipartite graph. In this sense, we call an element in $\mathcal{M}(V, K)$ a matching in (V, K).*

3 Weighted generating functions and moments

First we consider undirected graphs to describe the moments of real Wishart distributions. For an undirected graph (V', E') and variables $\boldsymbol{x} = (x_{i,j})$, we define the weight monomial $\boldsymbol{x}^{E'}$ by $\boldsymbol{x}^{E'} = \prod_{\{v,u\} \in E'} x_{v,u}$. If $x_{v,u} = x_{u,v}$ for $v, u \in V'$, then the weight monomial $\boldsymbol{x}^{E'}$ is well-defined. We define E'_0 to be $\{ \{\dot{1}, \ddot{1}\}, \ldots, \{\dot{n}, \ddot{n}\} \} \subset K'_{\dot{V},\ddot{V}}$. For $E' \in \mathcal{M}'(V')$, we define $\check{E}'$ and $\mathrm{len}(E')$ by

$$\check{E}' = \left\{ \{v, u\} \in K'_{V' \setminus \mathrm{vertex}(E')} \;\middle|\; \text{There exists a chain between } v \text{ and } u \text{ in } E' \cup E'_0. \right\} \subset K'_{V'}$$
$$\mathrm{len}(E') = (\text{the number of connected components in } (V', E' \cup E'_0)) - \left|\check{E}'\right|.$$

Remark 3.1 *For $E' \in \mathcal{M}'(V')$, $\check{E}'$ can be defined as a subset of $K'_{V'}$ satisfying the following conditions:*

- $\check{E}' \in \mathcal{M}'(V')$,
- $\check{E}' \cap E' = \emptyset$,
- $\check{E}' \cup E' \in \mathcal{P}'(E')$,
- *The number of connected components in $(V', E' \cup E'_0)$ equals the number of connected components in $(V', \check{E}' \cup E' \cup E'_0)$.*

Remark 3.2 *For $E' \in \mathcal{M}'(V')$, let us consider the undirected graph $(V', E' \amalg E'_0)$ with multiple edges. The connected components of $(V', E' \amalg E'_0)$ are chains and cycles without chords. The number of cycles in $(V', E' \amalg E'_0)$ equals $\operatorname{len}(E')$. The vertices $V' \setminus \operatorname{vertex}(E')$ which do not appear in E' are terminals of chains in $(V', E' \amalg E'_0)$. The set of pairs of terminals of chains in $(V', E' \amalg E'_0)$ equals $\check{E}'$.*

Definition 3.3 *For a set $K' \subset K'_{V'}$ of undirected edges, we define polynomials $\Phi'_{K'}$ and $\Psi'_{K'}$ by*

$$\Phi'_{K'}(t, \boldsymbol{x}, \boldsymbol{y}) = \sum_{E' \in \mathcal{M}'(V', K')} t^{\operatorname{len}(E')} \boldsymbol{x}^{E'} \boldsymbol{y}^{\check{E}'}, \qquad \Psi'_{K'}(t, \boldsymbol{x}) = \sum_{E' \in \mathcal{P}'(V', K')} t^{\operatorname{len}(E')} \boldsymbol{x}^{E'}.$$

We also respectively define $\Phi'(t, \boldsymbol{x}, \boldsymbol{y})$ and $\Psi'(t, \boldsymbol{x})$ to be $\Phi'_{K'_{V'}}(t, \boldsymbol{x}, \boldsymbol{y})$ and $\Psi'_{K'_{V'}}(t, \boldsymbol{x})$.

Remark 3.4 *By definition, $\Psi'_{K'}(t, \boldsymbol{x}) = \Phi'_{K'}(t, \boldsymbol{x}, 0)$ for each $K' \subset K'_{V'}$.*

We have the following formula that describes the moments of the real noncentral Wishart distribution as the special values of the weighted generating function.

Theorem 3.5 *Let $W = (w_{i,j}) \sim W_p(\nu, \Sigma, \Delta)$, namely, let W be a random matrix distributed according to the real noncentral Wishart distribution $W_p(\nu, \Sigma, \Delta)$. Then*

$$\mathbb{E}[w_{1,2} w_{3,4} \cdots w_{2n-1,2n}] = \mathbb{E}[w_{\dot{1},\ddot{1}} w_{\dot{2},\ddot{2}} \cdots w_{\dot{n},\ddot{n}}] = \Phi'(t, \boldsymbol{x}, \boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{u,v},\ y_{u,v}=\delta_{u,v}} = \Phi'(\nu, \Sigma, \Delta).$$

Corollary 3.6 *For $W \sim W_p(\nu, \Sigma, \Delta)$*

$$\mathbb{E}[w_{i_1,i_2} w_{i_3,i_4} \cdots w_{i_{2n-1},i_{2n}}] = \mathbb{E}[w_{i_{\dot{1}},i_{\ddot{1}}} w_{i_{\dot{2}},i_{\ddot{2}}} \cdots w_{i_{\dot{n}},i_{\ddot{n}}}] = \Phi'(t, \boldsymbol{x}, \boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{i_u,i_v},\ y_{u,v}=\delta_{i_u,i_v}}.$$

In the case where $\Delta = 0$, $W_p(\nu, \Sigma, 0)$ is called the real central Wishart distribution and is denoted by $W_p(\nu, \Sigma)$. It follows from Remark 3.4 that the moments of the central real Wishart distribution are written as special values of Ψ'.

Corollary 3.7 *For $W = (w_{i,j}) \sim W_p(\nu, \Sigma)$*

$$\mathbb{E}[w_{1,2} w_{3,4} \cdots w_{2n-1,2n}] = \mathbb{E}[w_{\dot{1},\ddot{1}} w_{\dot{2},\ddot{2}} \cdots w_{\dot{n},\ddot{n}}] = \Psi'(t, \boldsymbol{x})\Big|_{t=\nu,\ x_{u,v}=\sigma_{u,v}} = \Psi'(\nu, \Sigma).$$

Corollary 3.8 *For $W = (w_{i,j}) \sim W_p(\nu, \Sigma)$*

$$\mathbb{E}[w_{i_1,i_2} w_{i_3,i_4} \cdots w_{i_{2n-1},i_{2n}}] = \mathbb{E}[w_{i_{\dot{1}},i_{\ddot{1}}} w_{i_{\dot{2}},i_{\ddot{2}}} \cdots w_{i_{\dot{n}},i_{\ddot{n}}}] = \Psi'(t, \boldsymbol{x})\Big|_{t=\nu,\ x_{u,v}=\sigma_{i_u,i_v}}.$$

Next we consider directed graphs to describe the moments of complex Wishart distributions. For a directed graph (V, E) and variables $\boldsymbol{x} = (x_{i,j})$, we define the weight monomial $\boldsymbol{x}^E$ by $\boldsymbol{x}^E = \prod_{(v,u) \in E} x_{v,u}$. Let $E \in \mathcal{M}(V)$. The pair (V, E) is a directed graph whose connected components are directed chains and directed cycles without chords. We define $\operatorname{len}(E)$ to be the number of cycles (and self loops) in (V, E). The vertices $V \setminus \operatorname{start}(E)$ are the endpoints of the chains in (V, E), while the vertices $V \setminus \operatorname{end}(E)$ are the start points of the chains in (V, E). We define $\check{E}$ by

$$\check{E} = \left\{ (v, u) \in K_{V \setminus \operatorname{start}(E), V \setminus \operatorname{end}(E)} \;\middle|\; \text{There exists a chain from } u \text{ to } v \text{ in } E. \right\} \subset K_V.$$

Remark 3.9 *For $E \in \mathcal{M}(V)$, $\check{E}$ can be defined as a subset of K_V satisfying the following conditions:*

- $\check{E} \in \mathcal{M}(V)$,
- $\check{E} \cap E = \emptyset$,
- $\check{E} \cup E \in \mathcal{P}(E)$,
- *The number of connected components in (V, E) equals the number of connected components in $(V, \check{E} \cup E)$.*

Remark 3.10 *For $E \in \mathcal{M}(V)$, we can also define $\operatorname{len}(E)$ by*

$$\operatorname{len}(E) = (\textit{the number of connected components in } (V, E)) - \left|\check{E}\right|.$$

Remark 3.11 *We can identify $E \in \mathcal{P}(V)$ with the element σ_E of the symmetric group S_n such that $\sigma_E(i) = j$ for each $(i, j) \in E$. For each E, $\operatorname{len}(E)$ is the number of cycles of σ_E.*

Definition 3.12 *For a set $K \subset K_V$ of directed edges, we define polynomials Φ_K and Ψ_K by*

$$\Phi_K(t, \boldsymbol{x}, \boldsymbol{y}) = \sum_{E \in \mathcal{M}(V,K)} t^{\operatorname{len}(E)} \boldsymbol{x}^E \boldsymbol{y}^{\check{E}}, \qquad \Psi_K(t, \boldsymbol{x}) = \sum_{E \in \mathcal{P}(V,K)} t^{\operatorname{len}(E)} \boldsymbol{x}^E.$$

We also respectively define $\Phi(t, \boldsymbol{x}, \boldsymbol{y})$ and $\Psi(t, \boldsymbol{x})$ to be $\Phi_{K_V}(t, \boldsymbol{x}, \boldsymbol{y})$ and $\Psi_{K_V}(t, \boldsymbol{x})$.

Remark 3.13 *By definition, $\Psi_K(t, \boldsymbol{x}) = \Phi_K(t, \boldsymbol{x}, 0)$ for each $K \subset K_V$.*

We describe the moments of complex Wishart distributions as special values of the generating functions.

Theorem 3.14 *Let $W = (w_{i,j})$ be a random matrix distributed according to the complex noncentral Wishart distribution $CW_p(\nu, \Sigma, \Delta)$. Then*

$$\mathbb{E}[w_{1,2} w_{3,4} \cdots w_{2n-1,2n}] = \mathbb{E}[w_{\dot{1},\ddot{1}} w_{\dot{2},\ddot{2}} \cdots w_{\dot{n},\ddot{n}}] = \Phi(t, \boldsymbol{x}, \boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{\dot{u},\ddot{v}},\ y_{u,v}=\delta_{\dot{u},\ddot{v}}}.$$

Corollary 3.15 *For $W = (w_{i,j}) \sim CW_p(\nu, \Sigma, \Delta)$*

$$\mathbb{E}[w_{i_1,i_2} w_{i_3,i_4} \cdots w_{i_{2n-1},i_{2n}}] = \mathbb{E}[w_{i_{\dot{1}},i_{\ddot{1}}} w_{i_{\dot{2}},i_{\ddot{2}}} \cdots w_{i_{\dot{n}},i_{\ddot{n}}}] = \Phi(t, \boldsymbol{x}, \boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{i_{\dot{u}},i_{\ddot{v}}},\ y_{u,v}=\delta_{i_{\dot{u}},i_{\ddot{v}}}}.$$

By substituting 0 for Δ in the theorem, we have the following formula for the central complex case.

Corollary 3.16 *For $W = (w_{i,j}) \sim CW_p(\nu, \Sigma)$*

$$\mathbb{E}[w_{1,2} w_{3,4} \cdots w_{2n-1,2n}] = \mathbb{E}[w_{\dot{1},\ddot{1}} w_{\dot{2},\ddot{2}} \cdots w_{\dot{n},\ddot{n}}] = \Psi(t, \boldsymbol{x})\Big|_{t=\nu,\ x_{uv}=\sigma_{\dot{u},\ddot{v}}}.$$

Corollary 3.17 *For $W = (w_{i,j}) \sim CW_p(\nu, \Sigma)$*

$$\mathbb{E}[w_{i_1,i_2} w_{i_3,i_4} \cdots w_{i_{2n-1},i_{2n}}] = \mathbb{E}[w_{i_{\dot{1}},i_{\ddot{1}}} w_{i_{\dot{2}},i_{\ddot{2}}} \cdots w_{i_{\dot{n}},i_{\ddot{n}}}] = \Psi(t, \boldsymbol{x})\Big|_{t=\nu,\ x_{u,v}=\sigma_{i_{\dot{u}},i_{\ddot{v}}}}.$$

Remark 3.18 *For a square matrix $A = (a_{ij})$, the α-determinant (or α-permanent) is defined by*

$$\det_\alpha(A) = \sum_{\sigma \in S_n} \alpha^{n-\mathrm{len}(\sigma)} a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}.$$

This polynomial is an α-analogue of both the determinant and the permanent. Equivalently, the α-determinant is nothing but the ordinary determinant and permanent for $\alpha = -1$ and 1, respectively. (See also [16, 17].) Through the identification in Remark 3.10, we have $\alpha^n \Psi(\alpha^{-1}, A) = \det_\alpha(A)$. Moreover, by Corollary 3.16, the moments of the complex central Wishart distribution are expressed by α-determinants.

In [12], Matsumoto introduced the α-Pfaffian, which is defined by

$$\mathrm{pf}_\alpha(A) = \sum_{E' \in \mathcal{P}'(V')} (-\alpha)^{n-\mathrm{len}(E')} \operatorname{sgn}(E') A^{E'}$$

for a skew-symmetric matrix A, where $\operatorname{sgn}(E')A^{E'}$ is defined to be $\operatorname{sgn}(x) a_{x_{\dot{1}},x_{\ddot{1}}} \cdots a_{x_{\dot{n}},x_{\ddot{n}}}$ for $x \in S_{2n}$ such that $E' = \{ \{x_{\dot{1}}, x_{\ddot{1}}\}, \ldots, \{x_{\dot{n}}, x_{\ddot{n}}\} \}$. Since A is skew symmetric, $\operatorname{sgn}(E')A^{E'}$ is independent from choices of $x \in S_{2n}$. The α-Pfaffian is an analogue of the Pfaffian. Equivalently, in the case when $\alpha = -1$, α-Pfaffian $\mathrm{pf}_{-1}(A)$ is nothing but the ordinary Pfaffian $\mathrm{pf}(A)$, i.e., $\sum \operatorname{sgn}(x) a_{x_{\dot{1}} x_{\ddot{1}}} \cdots a_{x_{\dot{n}} x_{\ddot{n}}}$.

Let us define the polynomial $\mathrm{hf}_\alpha(A)$ by

$$\mathrm{hf}_\alpha(B) = \sum_{E' \in \mathcal{P}'(V')} \alpha^{n-\mathrm{len}(E')} B^{E'}$$

for a symmetric matrix B. The polynomial $\mathrm{hf}_\alpha(B)$ is an α-analogue of the Hafnian. Equivalently, $\mathrm{hf}_\alpha(B)$ is the ordinary Hafnian $\mathrm{hf}(B)$, i.e., $\sum b_{x_{\dot{1}} x_{\ddot{1}}} \cdots b_{x_{\dot{n}} x_{\ddot{n}}}$, for $\alpha = 1$. By definition, $\mathrm{hf}_\alpha(B) = \alpha^n \Psi'(\alpha^{-1}, B)$. In this sense, the moments of the real central Wishart distributions are expressed by α-Hafnians.

4 Application

4.1 Relation between real and complex cases

There exist bijections between directed graphs and undirected graphs which preserve the weight monomials in some special cases. These bijections induce equations between weighted generating functions of matchings. From the equations, we can obtain some formulas for the moments of complex and real Wishart distributions with special parameters.

4.1.1 Prototypical case

As in Remark 2.1, there exists a correspondence between directed graphs and bipartite graphs.

Lemma 4.1 *The map $\mathcal{M}(V, K_V) \ni (u, v) \mapsto \{\dot{u}, \ddot{v}\} \in \mathcal{M}'(V', K'_{\dot{V},\ddot{V}})$ is a bijection. The map induces the bijection $\mathcal{P}(V, K_V) \ni (u, v) \mapsto \{\dot{u}, \ddot{v}\} \in \mathcal{P}'(V', K'_{\dot{V},\ddot{V}})$.*

These bijections imply $\Phi_{K_V}(t,\boldsymbol{x},\boldsymbol{y}) = \Phi'_{K'_{\dot V,\ddot V}}(t,\boldsymbol{x},\boldsymbol{y})$ and $\Psi_{K_V}(t,\boldsymbol{x}) = \Psi'_{K'_{\dot V,\ddot V}}(t,\boldsymbol{x})$. If $\Sigma' = (\sigma'_{uv})$ and $\Delta' = (\delta'_{uv})$ satisfy $\sigma'_{u,v} = 0$ and $\delta'_{u,v} = 0$ for $\{u,v\} \in K'_{\dot V} \cup K'_{\ddot V}$, then

$$\Phi'(t,\boldsymbol{x},\boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma'_{u,v},\ y_{u,v}=\delta'_{u,v}} = \Phi'_{K'_{\dot V,\ddot V}}(t,\boldsymbol{x},\boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma'_{u,v},\ y_{u,v}=\delta'_{u,v}}.$$

If $\Sigma = (\sigma_{uv})$ and $\Delta = (\delta_{uv})$ satisfy $\sigma_{\dot u,\ddot v} = \sigma'_{u,v}$ and $\delta_{\dot u,\ddot v} = \delta'_{u,v}$ for $u, v \in V$, then the equation implies

$$\Phi(t,\boldsymbol{x},\boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{\dot u,\ddot v},\ y_{u,v}=\delta_{\dot u,\ddot v}} = \Phi'(t,\boldsymbol{x},\boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{u,v},\ y_{u,v}=\delta_{u,v}}.$$

Hence we have the following:

Propsition 4.2 *Let $\Sigma' = (\sigma'_{u,v})$, $\Delta' = (\delta'_{u,v})$, $\Sigma = (\sigma_{u,v})$ and $\Delta = (\delta_{u,v})$ satisfy $\sigma'_{u,v} = 0$, $\delta'_{u,v} = 0$ for $\{u,v\} \in K'_{\dot V} \cup K'_{\ddot V}$, and $\sigma_{\dot u,\ddot v} = \sigma'_{\dot u,\ddot v}$, $\delta_{\dot u,\ddot v} = \delta'_{\dot u,\ddot v}$ for $u, v \in V$. For $W = (w_{u,v}) \sim CW_p(\nu,\Sigma,\Delta)$ and $W' = (w'_{u,v}) \sim W_p(\nu,\Sigma',\Delta')$, $\mathbb{E}[w_{\dot 1,\ddot 1}\cdots w_{\dot n,\ddot n}] = \mathbb{E}[w'_{\dot 1,\ddot 1}\cdots w'_{\dot n,\ddot n}]$.*

4.1.2 Central case

Next we consider the central Wishart distribution. In this case, we may consider only perfect matchings. We define $\widetilde{\mathcal{P}}'(V')$ and $\widetilde{\mathcal{P}}(V)$ by

$$\widetilde{\mathcal{P}}'(V') = \left\{ (E',\omega') \;\middle|\; {\scriptstyle E'\in\mathcal{P}'(V'), \atop \omega':\ \{\text{ cycles in } (V',\, E \amalg E_0)\} \to \{\pm 1\}} \right\}, \qquad \widetilde{\mathcal{P}}(V) = \left\{ (E,\omega) \;\middle|\; {\scriptstyle E\in\mathcal{P}(V), \atop \omega:\ E\to\{\pm 1\}} \right\}.$$

Lemma 4.3 *There exists a bijection between $\widetilde{\mathcal{P}}'(V')$ and $\widetilde{\mathcal{P}}(V)$.*

We shall give a bijection ψ between $\widetilde{\mathcal{P}}'(V')$ and $\widetilde{\mathcal{P}}(V)$ in Section 4.1.4. The bijection preserves the weight monomials, equivalently, $t^{\mathrm{len}(E)}\boldsymbol{x}^E = t^{\mathrm{len}(E')}(\boldsymbol{x}')^{E'}$ for elements $E' \in \mathcal{P}'(V)$ corresponding to $E \in \mathcal{P}(V)$, in the case when $\boldsymbol{x} = (x_{u,v})$ and $\boldsymbol{x}' = (x'_{u,v})$ satisfy $x'_{u',v'} = x_{u,v}$ for any $u, v \in V$ and any $\{u',v'\} \in K_{\{\dot u\ddot v\},\{\dot u\ddot v\}}$. Hence Proposition 4.4 follows from the following equations:

$$2^n\Psi(t,\boldsymbol{x}) = 2^n \sum_{E\in\mathcal{P}(V)} t^{\mathrm{len}(E)}\boldsymbol{x}^E = \sum_{E\in\widetilde{\mathcal{P}}(V)} t^{\mathrm{len}(E)}\boldsymbol{x}^E,$$

$$\Psi'(2t,\boldsymbol{x}) = 2^n \sum_{E'\in\mathcal{P}'(V')} (2t)^{\mathrm{len}(E')}\boldsymbol{x}^{E'} = \sum_{E'\in\widetilde{\mathcal{P}}(V')} t^{\mathrm{len}(E')}\boldsymbol{x}^{E'}.$$

Propsition 4.4 *Let $\Sigma = (\sigma_{u,v})$ and $\Sigma' = (\sigma'_{u,v})$ satisfy $\sigma'_{u',v'} = \sigma_{u,v}$ for any $u, v \in V$ and any $\{u',v'\} \in K_{\{\dot u\ddot v\},\{\dot u\ddot v\}}$. Then*

$$2^n\Psi(t,\boldsymbol{x})\Big|_{t=\nu,\ x_{u,v}=\sigma_{\dot u,\ddot v}} = \Psi'(2t,\boldsymbol{x})\Big|_{t=\nu,\ x_{u,v}=\sigma'_{u,v}}.$$

Corollary 4.5 *Let $\Sigma = (\sigma_{u,v})$ and $\Sigma' = (\sigma'_{u,v})$ satisfy $\sigma'_{u',v'} = \sigma_{u,v}$ for any $u, v \in V$ and any $\{u',v'\} \in K_{\{\dot u\ddot v\},\{\dot u\ddot v\}}$. For $W = (w_{u,v}) \sim CW_p(\nu,\Sigma)$ and $W' = (w'_{u,v}) \sim W_p(\nu,\Sigma')$, $\mathbb{E}[w_{\dot 1,\ddot 1}\cdots w_{\dot n,\ddot n}] = \mathbb{E}[w'_{\dot 1,\ddot 1}\cdots w'_{\dot n,\ddot n}]$.*

4.1.3 Noncentral case

Next we consider the noncentral Wishart distribution. In this case, we consider all matchings in complete graphs. We define $\widetilde{\mathcal{M}}(V)$ and $\widetilde{\mathcal{M}}'(V')$ by

$$\widetilde{\mathcal{M}}'(V') = \left\{ (E', \omega') \,\middle|\, \substack{E' \in \mathcal{M}'(V'),\\ \omega' \colon \{\, \text{cycles in } (V', E' \cup E'_0) \,\} \to \{\, \pm 1 \,\}} \right\}, \quad \widetilde{\mathcal{M}}(V) = \left\{ (E, \omega) \,\middle|\, \substack{E \in \mathcal{M}(V),\\ \omega \colon E \to \{\, \pm 1 \,\}} \right\}.$$

Lemma 4.6 *There exists a bijection between $\widetilde{\mathcal{M}}(V)$ and $\widetilde{\mathcal{M}}'(V')$.*

We shall give a bijection between $\widetilde{\mathcal{P}}'(V')$ and $\widetilde{\mathcal{P}}(V)$ in Section 4.1.4. The bijection preserves the weight monomial in a special case. Hence Proposition 4.7 follows from the following equations:

$$\Phi'(2t, \boldsymbol{x}, \boldsymbol{y}) = \sum_{E' \in \mathcal{M}'(V')} (2t)^{\operatorname{len}(E')} \boldsymbol{x}^{E'} \boldsymbol{y}^{\check{E}'} = \sum_{(E', \omega') \in \widetilde{\mathcal{M}}'(V')} t^{\operatorname{len}(E')} \boldsymbol{x}^{E'} \boldsymbol{y}^{\check{E}'},$$

$$\Phi(t, 2\boldsymbol{x}, \boldsymbol{y}) = \sum_{E \in \mathcal{M}(V)} t^{\operatorname{len}(E)} (2\boldsymbol{x})^{E} \boldsymbol{y}^{\check{E}} = \sum_{(E, \omega) \in \widetilde{\mathcal{M}}(V)} t^{\operatorname{len}(E)} \boldsymbol{x}^{E} \boldsymbol{y}^{\check{E}},$$

where $2\boldsymbol{x} = (2x_{u,v})$

Propsition 4.7 *Let $\Sigma = (\sigma_{u,v})$, $\Delta = (\delta_{u,v})$, $\Sigma' = (\sigma'_{u,v})$ and $\Delta' = (\delta'_{u,v})$ satisfy $\sigma'_{u',v'} = \sigma_{u,v}$ $\delta'_{u',v'} = \delta_{u,v}$ for any $u, v \in V$ and any $\{u', v'\} \in K_{\{\, \dot{u}\ddot{v} \,\},\{\, \dot{u}\ddot{v} \,\}}$. Then*

$$\Psi(t, 2\boldsymbol{x}, \boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{\dot{u},\ddot{v}},\ y_{u,v}=\delta_{\dot{u},\ddot{v}}} = \Psi'(2t, \boldsymbol{x}, \boldsymbol{y})\Big|_{t=\nu,\ x_{u,v}=\sigma_{u,v},\ y_{u,v}=\delta_{u,v}}.$$

Corollary 4.8 *Let $\Sigma = (\sigma_{u,v})$, $\Delta = (\delta_{u,v})$, $\Sigma' = (\sigma'_{u,v})$ and $\Delta' = (\delta'_{u,v})$ satisfy $\sigma'_{u',v'} = \sigma_{u,v}$ $\delta'_{u',v'} = \delta_{u,v}$ for any $u, v \in V$ and any $\{u', v'\} \in K_{\{\, \dot{u}\ddot{v} \,\},\{\, \dot{u}\ddot{v} \,\}}$. For $W = (w_{u,v}) \sim CW_p(\nu, 2\Sigma, \Delta)$ and $W' = (w'_{u,v}) \sim W_p(2\nu, \Sigma', \Delta')$, $\mathbb{E}[w_{\dot{1},\ddot{1}} \cdots w_{\dot{n},\ddot{n}}] = \mathbb{E}[w'_{\dot{1},\ddot{1}} \cdots w'_{\dot{n},\ddot{n}}]$.*

4.1.4 Construction of Bijections

Here we construct bijections to prove Lemmas 4.3 and 4.6. First we construct a bijection ψ from $\widetilde{\mathcal{P}}(V)$ to $\widetilde{\mathcal{P}}'(V')$. To define the bijection, we define the following map. For $(E, \omega) \in \widetilde{\mathcal{P}}(V)$, let $h_{E,\omega}$ and $h'_{E,\omega}$ be maps from V to V' defined by

$$h_{E,\omega}(v) = \begin{cases} \dot{v} & \text{if } \omega((u,v)) = 1 \text{ for some } (u,v) \in E, \\ \ddot{v} & \text{otherwise}, \end{cases}$$

$$h'_{E,\omega}(v) = \begin{cases} \ddot{v} & \text{if } \omega((u,v)) = 1 \text{ for some } (u,v) \in E, \\ \dot{v} & \text{otherwise}. \end{cases}$$

Remark 4.9 *For $(E, \omega) \in \widetilde{\mathcal{P}}(V)$ and $v \in V$, $\{h_{E,\omega}(v), h'_{E,\omega}(v)\} \in E'_0$.*

First we construct $E' \in \mathcal{P}'(V')$ for each $(E, \omega) \in \widetilde{\mathcal{P}}(V)$. For each $(E, \omega) \in \widetilde{\mathcal{P}}(V)$, we define a surjection $\psi_{E,\omega} \colon E \to K'_{V'}$, and then we define E' to be the image $\psi_{E,\omega}(E)$. Let (v_1, v_2), (v_2, v_3),…, (v_{k-1}, v_k),

(v_k, v_1) be a cycle in E such that $v_1 = \min\{ v_1, \dots, v_k \}$. For each directed edge in the cycle, we define $\psi_{E,\omega}$ by

$$\begin{aligned} \psi_{E,\omega}((v_1, v_2)) &= \{\dot{v}_1, h_{E,\omega}(v_2)\}, \\ \psi_{E,\omega}((v_i, v_{i+1})) &= \{h'_{E,\omega}(v_{i+1}), h_{E,\omega}(v_{i+1})\} \quad (\text{for } i = 2, \dots, k-1), \\ \psi_{E,\omega}((v_k, v_1)) &= \{h'_{E,\omega}(v_k), \ddot{v}_1\}. \end{aligned}$$

Then the image of the cycle forms a cycle C' in the undirected graph $E' \amalg E'_0$. For the cycle C', we define $\omega'(C')$ to be $\omega((v_k, v_1))$.

Remark 4.10 *It is easy to construct the inverse map of ψ, which implies that ψ is bijective.*

Remark 4.11 *This correspondence ψ is equivalent to the one in [4], which is described in more algebraic terms. Let S_{2m} be the $2m$-th symmetric group, and let $B_m = S_m \wr \mathbb{Z}/2\mathbb{Z}$ the hyperoctehedral group, i.e., the subgroup of the permutations $\pi \in S_{2m}$ such that $|\pi(\dot{n}) - \pi(\ddot{n})| = 1$ for all $n = 1, \dots, m$. For $gB_m \in S_{2m}/B_m$, we can define E_{gB_m} by $E_{gB_m} = \{ \{g(\dot{n}), g(\ddot{n})\} \mid n = 1, \dots m \} \in \mathcal{P}(V)$, and we can identify elements $gB_m \in S_{2m}/B_m$ with perfect matchings in (V, K_V). Through this identification, the correspondence in Section 4 of [4] is equivalent to ours.*

Next we construct a bijection φ from $\widetilde{\mathcal{M}}(V)$ to $\widetilde{\mathcal{M}}'(V')$. For each $(E, \omega) \in \widetilde{\mathcal{M}}(V)$, we shall define $(E', \omega') \in \widetilde{\mathcal{M}}(V)$. For each cycle in E, we construct undirected edges and ω' in the same manner as ψ. To define the undirected edges corresponding to chains, we define $t_{E,\omega}$ and $t'_{E,\omega}$ by

$$t_{E,\omega}(v) = \begin{cases} \dot{v} & \text{if } \omega((u,v)) = 1 \text{ for some } (v,u) \in E, \\ \ddot{v} & \text{otherwise,} \end{cases}$$

$$t'_{E,\omega}(v) = \begin{cases} \ddot{v} & \text{if } \omega((u,v)) = 1 \text{ for some } (v,u) \in E, \\ \dot{v} & \text{otherwise.} \end{cases}$$

Let (v_1, v_2), (v_2, v_3),…,(v_{k-1}, v_k) be a maximal chain in E. If $v_1 < v_k$, then we define $\varphi_{E,\omega}$ by

$$\begin{aligned} \varphi_{E,\omega}((v_1, v_2)) &= \{\dot{v}_1, h_{E,\omega}(v_2)\}, \\ \varphi_{E,\omega}((v_i, v_{i+1})) &= \{h'_{E,\omega}(v_i), h_{E,\omega}(v_{i+1})\} \quad (\text{for } i = 2, \dots, k-1). \end{aligned}$$

If $v_1 > v_k$, then we define $\varphi_{E,\omega}$ by

$$\begin{aligned} \varphi_{E,\omega}((v_{k-1}, v_k)) &= \{t_{E,\omega}(k_{k-1}), \ddot{v}_k\} \\ \varphi_{E,\omega}((v_{i-1}, v_i)) &= \{t_{E,\omega}(v_{i-1}), t'_{E,\omega}(v_i)\} \quad (\text{for } i = k-1, \dots, 2). \end{aligned}$$

Then the image of the maximal chains forms a maximal chain in the undirected graph $E' \amalg E'_0$.

Remark 4.12 *It is easy to construct the inverse map of φ, which implies φ is bijective.*

4.2 Noncentral chi-square distribution

In this section, we consider the Wishart distributions for a special parameter, which is linked with the noncentral chi-square distributions.

Propsition 4.13 *Let $\sigma_{u,v} = \sigma$, $\delta_{u,v} = \delta$. For $W = (w_{u,v}) \sim CW_p(\nu, \Sigma, \Delta)$,*

$$\mathbb{E}[w_{\dot{1},\ddot{1}} w_{\dot{2},\ddot{2}} \cdots w_{\dot{n},\ddot{n}}] = \sum_{m=0}^{n} \sum_{l} g_{lmn} \nu^l \sigma^m \delta^{n-m},$$

where g_{lmn} is the number of $E \subset K_V$ such that $\operatorname{len}(E) = l$, $|E| = m$ and $\left|\check{E}\right| = n - m$.

Corollary 4.14 *For the noncentral complex chi-square distribution $\chi^2_\nu(\delta)$ with ν degrees of freedom and the noncentrality parameter δ, its n-th moment $\mathbb{E}(w^n)$ is given as follows:*

$$\mathbb{E}[w^n] = \sum_{m=0}^{n} \sum_{l} g_{lmn} \nu^l \delta^{n-m}.$$

In the case where we add a new directed edge whose staring point is a fixed vertex, we have just one choice of end-points that increase the number of cycles. Hence we obtain Lemma 4.15.

Lemma 4.15 *Let $0 \leq m \leq n$. Then the generating function $G_{mn}(t)$ of g_{lmn} with respect to the number l of cycles satisfies*

$$G_{mn}(t) = \sum_{l \geq 0} g_{lmn} t^l = \binom{n}{m} \prod_{i=1}^{m} (t + n - i).$$

We also obtain the following corollary, which is well-known expression for the noncentral chi-square distribution (e.g. [5])

Corollary 4.16 *For the n-th moment $\mathbb{E}(w^n)$ of the noncentral chi-square distribution $\chi^2_\nu(\delta)$ with ν degrees of freedom and the noncentrality parameter δ,*

$$\mathbb{E}[w^n] = \sum_{m=0}^{n} \sum_{l} g_{lmn} \nu^l \delta^{n-m} = \sum_{m=0}^{n} G_{mn}(\nu) \delta^{n-m} = \sum_{m=0}^{n} \binom{n}{m} \delta^{n-m} \prod_{i=1}^{m} (\nu + n - i).$$

Remark 4.17 *The numbers $s_n(m,l)$ defined by the following generating function are called the noncentral Stirling numbers of the first kind:*

$$\sum_{l} s_n(m,l) t^l = \prod_{i=1}^{m} (t + n - i).$$

If $m = n$, then $s_n(m,l)$ is the Stirling number of the first kind. Lemma 4.15 implies that $g_{lmn} = \binom{n}{m} s_n(m,l)$. Equivalently, we can explicitly describe the moments of the noncentral chi-square distribution $\chi^2_\nu(\delta)$ with the noncentral Stirling numbers. Koutras pointed out that moments of some noncentral distributions are described with the noncentral Stirling numbers of the first kind [7].

Next consider the real case.

Propsition 4.18 *Let $\sigma_{u,v} = \sigma$, $\delta_{u,v} = \delta$. For $W = (w_{u,v}) \sim W_p(\nu, \Sigma, \Delta)$,*

$$\mathbb{E}[w_{\dot{1},\ddot{1}} w_{\dot{2},\ddot{2}} \cdots w_{\dot{n},\ddot{n}}] = \sum_{m=0}^{n} \sum_{l} g'_{lmn} \nu^l \sigma^m \delta^{n-m},$$

where g'_{lmn} is the number of $E' \subset K'_{V'}$ such that $\operatorname{len}(E') = l$, $|E'| = m$ and $\left|\check{E}'\right| = n - m$.

Corollary 4.19 *For the noncentral chi-square distribution $\chi^2_\nu(\delta)$ with ν degrees of freedom and the noncentrality parameter δ, its n-th moment $\mathbb{E}(w^n)$ is given as:*

$$\mathbb{E}[w^n] = \sum_{m=0}^{n} \sum_{l} g'_{lmn} \nu^l \delta^{n-m},$$

where g'_{lmn} is the number of $E' \subset K'_{V'}$ such that $\operatorname{len}(E') = l$, $|E'| = m$ and $\left|\check{E}'\right| = n - m$.

Proposition 4.7 and Lemma 4.15 imply Lemma 4.20.

Lemma 4.20 *Let $0 \le m \le n$, $n \le 0$. Then the generating function $G'_{mn}(t)$ of g'_{lmn} with respect to the number l of cycles satisfies*

$$G'_{mn}(t) = \sum_{l \ge 0} g'_{lmn} t^l = \binom{n}{m} \prod_{i=1}^{m} (t + 2(n-i)).$$

Corollary 4.21 *For the n-th moment $\mathbb{E}(w^n)$ of the noncentral chi-square distribution $\chi^2_\nu(\delta)$ with ν degrees of freedom and the noncentrality parameter δ,*

$$\mathbb{E}[w^n] = \sum_{m=0}^{n} \sum_{l} g'_{lmn} \nu^l \delta^{n-m} = \sum_{m=0}^{n} G'_{mn}(\nu) \delta^{n-m} = \sum_{m=0}^{n} \binom{n}{m} \delta^{n-m} \prod_{i=1}^{m} (\nu + 2(n-i)).$$

4.3 Bivariate chi-square distribution

We can explicitly describe the moments of Wishart distributions by enumerating the matchings satisfying some conditions. For example, in Proposition 4.23, we obtain the description of the moments of the bivariate real chi-square distribution, which was introduced by Kibble [6]. The formulas imply formulas for the complex distribution by Proposition 4.7. See [8] for details and other applications.

Propsition 4.22 *Let $\Sigma = (\sigma_{uv})$ and $\Delta = (\delta_{uv})$ satisfy*

$$\sigma_{u,v} = \begin{cases} 1 & (u, v \le 2b \text{ or } 2b+1 \le u, v), \\ \rho & (\text{otherwise}), \end{cases} \qquad \delta_{u,v} = 0.$$

For a random matrix $W = (w_{u,v}) \sim W_{b+c}(\nu, \Sigma, \Delta)$,

$$\begin{aligned} &\mathbb{E}[w_{\dot{1},\ddot{1}} \cdots w_{\dot{b},\ddot{b}} \cdot w_{(b\dot{+}1),(b\ddot{+}1)} \cdots w_{(b\dot{+}c),(b\ddot{+}c)}] \\ &= \sum_{a=0}^{\min(b,c)} \rho^{2a} \frac{2^a b! c!}{(b-a)!(c-a)!a!} \prod_{i=1}^{a} (\nu + a(a-i)) \prod_{i=1}^{b-a} (\nu + a(b-i)) \prod_{i=1}^{c-a} (\nu + a(c-i)). \end{aligned}$$

Propsition 4.23 *Let $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, and $W = (w_{u,v}) \sim W_2(\nu, \Sigma)$. For $b, c \in \mathbb{Z}_{\ge 0}$,*

$$\mathbb{E}[w_{1,1}^b w_{2,2}^c] = \sum_{a=0}^{\min(b,c)} \rho^{2a} \frac{2^a b! c!}{(b-a)!(c-a)!a!} \prod_{i=1}^{a} (\nu + 2(a-i)) \prod_{i=1}^{b-a} (\nu + 2(b-i)) \prod_{i=1}^{c-a} (\nu + 2(c-i)).$$

Remark 4.24 *In [14], Nadarajah and Kotz derived another expression for $\mathbb{E}[w_{1,1}^b w_{2,2}^c]$ with the Jacobi polynomials.*

References

[1] Bai, Z. D. (1999). Methodologies in spectral analysis of large dimensional random matrices, A review. *Statist. Sinica*, **9**, 611–677.

[2] Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (An introduction). *Ann. Math. Statist.*, **34**, 152–177.

[3] Graczyk, P., Letac, G. and Massam, H. (2003). The complex Wishart distribution and the symmetric groups. *Ann. Statist.*, **31**, 287–309.

[4] Graczyk, P., Letac, G. and Massam, H. (2005). The hyperoctahedral group, symmetric group representations and the moments of the real Wishart distribution. *J. Theor. Probab.*, **18**, 1–42.

[5] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, Vol. 2, 2nd ed. Wiley-Interscience.

[6] Kibble, W. F. (1941). A two-variate gamma type distribution. *Sankhya*, *5A*, 137–150.

[7] Koutras, M. (1982). Noncentral Stirling numbers and some applications. *Discrete Math.*, **42**, 73–89.

[8] Kuriki, S. and Numata, Y. (2010). Graph representations for moments of noncentral Wishart distributions and their applications. *Annals of the Institute of Statistical Mathematics*, **62** 4, 645–672.

[9] Letac, G. and Massam, H. (2008). The noncentral Wishart as an exponential family, and its moments. *J. Multivariate Anal.*, **99**, 1393–1417.

[10] Lu, I-L. and Richards, D. St. P. (2001). MacMahon's master theorem, representation theory, and moments of Wishart distributions. *Adv. Appl. Math.*, **27**, 531–547.

[11] Maiwald, D. and Kraus, D. (2000). Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices. *IEE Proc.-Radar, Sonar Navig*, **147**, 162–168.

[12] Matsumoto, S. (2005) α-Pfaffian, pfaffian point process and shifted Schur measure. *Linear Algebra and its Applications*, **403**, 369–398.

[13] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. John Wiley & Sons

[14] Nadarajah, S. and Kotz, S. (2006). Product moments of Kibble's bivariate gamma distribution. *Circuits Systems Signal Process.*, **25**, 567–570.

[15] Takemura, A. (1991). *Foundations of Multivariate Statistical Inference* (in Japanese). Kyoritsu Shuppan.

[16] Vere-Jones, D. (1988). A generalization of permanents and determinants. *Linear Algebra Appl.*, **111**, 119–124.

[17] Vere-Jones, D. (1997). Alpha-permanents and their applications to multivariate gamma, negative binomial and ordinary binomial distributions. *New Zealand J. Math.*, **26**, 125–149.

[18] Wishart, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, **20A**, 32–52.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 833–840

Bruhat order, rationally smooth Schubert varieties, and hyperplane arrangements

Suho Oh[1†] and Hwanchul Yoo

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139

Abstract. We link Schubert varieties in the generalized flag manifolds with hyperplane arrangements. For an element of a Weyl group, we construct a certain graphical hyperplane arrangement. We show that the generating function for regions of this arrangement coincides with the Poincaré polynomial of the corresponding Schubert variety if and only if the Schubert variety is rationally smooth.

Résumé. Nous relions des variétés de Schubert dans le variété flag généralisée avec des arrangements des hyperplans. Pour un élément dún groupe de Weyl, nous construisons un certain arrangement graphique des hyperplans. Nous montrons que la fonction génératrice pour les régions de cet arrangement coincide avec le polynome de Poincaré de la variété de Schubert correspondante si et seulement si la variété de Schubert est rationnellement lisse.

Keywords: Bruhat order, Schubert Variety, Rational Smoothness, Palidromic, Hyperplanes, Coxeter arrangement

1 Introduction

For an element of a Weyl group $w \in W$, let $P_w(q) := \sum_{u \leq w} q^{\ell(u)}$, where the sum is over all elements $u \in W$ below w in the (strong) Bruhat order. Geometrically, the polynomial $P_w(q)$ is the Poincaré polynomial of the Schubert variety $X_w = BwB/B$ in the flag manifold G/B.

The *inversion hyperplane arrangement* $\mathcal{A}_w$ is defined as the collection of hyperplanes corresponding to all inversions of w. Let $R_w(q) := \sum_r q^{d(r_0,r)}$ be the generating function that counts regions r of the arrangement $\mathcal{A}_w$ according to the distance $d(r_0, r)$ from the fixed initial region r_0.

The main result of the paper is the claim that $P_w(q) = R_w(q)$ if and only if the Schubert variety X_w is rationally smooth. We have previously given an elementary combinatorial proof for Type A case of this problem in Oh et al. (2008).

According to the criterion of Peterson and Carrell (1994), the Schubert variety X_w is rationally smooth if and only if the Poincaré polynomial $P_w(q)$ is palindromic, that is $P_w(q) = q^{\ell(w)} P_w(q^{-1})$. If w is not rationally smooth then the polynomial $P_w(q)$ is not palindromic, but the polynomial $R_w(q)$ is always palindromic. So $P_w(q) \neq R_w(q)$ in this case. Hence it is enough to show that $P_w(q) = R_w(q)$ when w is rationally smooth. Our proof is purely combinatorial, combining basics of Weyl groups with a result from Billey and Postnikov (2005).

[†]S.O. was supported in part by Samsung Scholarship.

2 Rational smoothness of Schubert varieties and Inversion hyperplane arrangement

In this section we will explain how rational smoothness can be expressed by conditions on the lower Bruhat interval. We will also define the Inversion hyperplane arrangement. In this paper, unless stated otherwise, we refer to the strong Bruhat order.

Let G be a semisimple simply-connected complex Lie group, B a Borel subgroup and $\mathfrak{h}$ the corresponding Cartan subalgebra. Let W be the corresponding Weyl group, $\Delta \subset \mathfrak{h}^*$ be the set of roots and $\Pi \subset \Delta$ be the set of simple roots. The choice of simple roots determines the set of positive roots. We will write $\alpha > 0$ for $\alpha \in \Delta$ being a positive root. Following the conventions of Björner and Brenti (2005), let S be the set of simple reflections and $T := \{wsw^{-1} : s \in S, w \in W\}$ be the set of reflections. Set $\Pi = \{\alpha_1, \cdots, \alpha_n\}$, $S = \{s_1, \cdots, s_n\}$ and index them properly so that s_i and α_i corresponds to the same node of the Dynkin diagram for $1 \leq i \leq n$. Then there is a bijection between T and Δ by matching ws_iw^{-1} with $w(\alpha_i)$. Then ws_iw^{-1} is exactly the reflection that reflects by the hyperplane corresponding to the root $w(\alpha_i)$.

We have the following definitions as in Björner and Brenti (2005):

$$T_L(w) := \{t \in T : \ell(tw) < \ell(w)\},$$

$$T_R(w) := \{t \in T : \ell(wt) < \ell(w)\},$$

$$D_L(w) := T_L(w) \cap S,$$

$$D_R(w) := T_R(w) \cap S.$$

They are called the left(right) ***associated reflections*** of w and left(right) ***descent set*** of w. In this paper, we concentrate on lower Bruhat intervals in W, $[\mathrm{id}, w] := \{u \in S_n \mid u \leq w\}$. They are related to *Schubert varieties* $X_w = \overline{BwB/B}$ inside the generalized flag manifold G/B. The *Poincaré polynomial* of the Schubert variety X_w is the rank generating function for the interval $[\mathrm{id}, w]$, e.g., see Billey et al. (2000):

$$P_w(q) = \sum_{u \leq w} q^{\ell(u)}.$$

For convenience, we will say that $P_w(q)$ is the Poincaré polynomial of w. And we will say that w is rationally smooth if X_w is rationally smooth. Due to Carrell and Peterson, one can check whether the rational locus of a Schubert variety is smooth or not by studying $P_w(q)$. Let us denote a polynomial $f(q) = a_0 + a_1\, q + \cdots + a_d\, q^d$ as *palindromic* if $f(q) = q^d f(q^{-1})$, i.e., $a_i = a_{d-i}$ for $i = 0, \ldots, d$.

Theorem 1 (Carrell-Peterson Carrell (1994), see also (Billey et al., 2000, Sect. 6.2)) *For any element of a Weyl group $w \in W$, the Schubert variety X_w is rationally smooth if and only if the Poincaré polynomial $P_w(q)$ is palindromic.*

For each $w \in W$, we will be comparing this polynomial $P_w(q)$ with another polynomial, that comes from an associated hyperplane arrangement. To assign a hyperplane arrangement to each $w \in W$, we first need the definition of the inversion set of w. The inversion set Δ_w of w is defined as the following:

$$\Delta_w := \{\alpha | \alpha \in \Delta, \alpha > 0, w(\alpha) < 0\}.$$

For type A case, this gives the usual definition of an inversion set for permutations. Let us define the arrangement $\mathcal{A}_w$ as the collection of hyperplanes $\alpha(x) = 0$ for all roots $\alpha \in \Delta_w$. Let r_0 be the ***fundamental chamber*** of $\mathcal{A}_w$, the chamber that contains the points satisfying $\alpha(x) > 0$ for all $\alpha \in \Delta_w$. Then we can define a polynomial from this $\mathcal{A}_w$:

$$R_w(q) := \sum_r q^{d(r_0,r)},$$

where the sum is over all chambers of the arrangement $\mathcal{A}_w$ and $d(r_0, r)$ is the number of hyperplanes separating r_0 and r. Our goal in this paper is to show that $R_w(q) = P_w(q)$ whenever $P_w(q)$ is palindromic.

Remark 2 *We have $P_w(q) = P_{w^{-1}}(q)$ and $R_w(q) = R_{w^{-1}}(q)$ by definition. Whenever we use this fact, we will call this the* **duality** *of $P_w(q)$ and $R_w(q)$.*

Given an arrangement $\mathcal{A}_w$ and its subarrangement $\mathcal{A}'$, let c be a chamber of $\mathcal{A}'$. Then a ***chamber graph*** of c with respect to $\mathcal{A}_w$ is defined as a directed graph $G = (V, E)$ where

- The vertex set V consists of vertices representing each chambers of $\mathcal{A}_w$ contained in c,
- we have an edge directed from vertex representing chamber c_1 to a vertex representing chamber c_2 if c_1 and c_2 are adjacent and $d(r_0, c_1) + 1 = d(r_0, c_2)$.

We will say that $\mathcal{A}_w$ is ***uniform*** with respect to $\mathcal{A}'$ if for all chambers of $\mathcal{A}'$, chamber graphs with respect to $\mathcal{A}_w$ are isomorphic. One can easily see that if $\mathcal{A}_u$ is a subarrangement of $\mathcal{A}_w$ and $\mathcal{A}_w$ is uniform with respect to $\mathcal{A}_u$, then $R_w(q)$ is divided by $R_u(q)$.

3 Parabolic Decomposition

In this section, we introduce a theorem of Billey and Postnikov (2005) regarding parabolic decomposition that will serve as a key tool in our proof. Let's first recall the definition of the parabolic decomposition. Given a Weyl group W, fix a subset J of simple roots. Denote W_J to be the parabolic subgroup generated by simple reflections of J. Let W^J be the set of minimal length coset representatives of $W_J \backslash W$. Then it is a well-known fact that every $w \in W$ has a unique parabolic decomposition $w = uv$ where $u \in W_J, v \in W^J$ and $\ell(w) = \ell(u) + \ell(v)$.

Lemma 3 *(van den Hombergh (1974)) For any $w \in W$ and subset J of simple roots, W_J has a unique maximal element below w.*

We will denote the maximal element of W_J below w as $m(w, J)$.

Theorem 4 *(Billey and Postnikov (2005)) Let J be any subset of simple roots. Assume $w \in W$ has parabolic decomposition $w = uv$ with $u \in W_J$ and $v \in W^J$ and furthermore, $u = m(w, J)$. Then*

$$P_w(t) = P_u(t)P_v^{W^J}(t)$$

where $P_v^{W^J} = \sum_{z \in W^J, z \leq v} t^{\ell(z)}$ is the Poincaré polynomial for v in the quotient.

This decomposition is very useful in the sense that it allows us to factor the Poincaré polynomials. We will say that $J = \Pi \setminus \{\alpha\}$ is ***leaf-removed*** if α corresponds to a leaf in the Dynkin diagram of Π.

The following theorem of Billey and Postnikov (2005) tells us that we only need to look at maximal leaf-removed parabolic subgroups for our purpose.

Theorem 5 *(Billey and Postnikov (2005)) Let $w \in W$ be a rationally smooth element. Then there exists a maximal proper subset $J = \Pi \setminus \{\alpha\}$ of simple roots, such that*

1. *we have a decomposition of w or w^{-1} as in Theorem 4,*
2. *α corresponds to a leaf in the Dynkin diagram of W.*

We will call the parabolic decompositions that satisfies the conditions of the above theorem as ***BP***-decompositions. For Weyl groups of type A,B and D, there is a stronger result by Billey:

Lemma 6 *(Billey (1998)) Let W be a Weyl group of type A,B or D. Let $w \in W$ be a rationally smooth element. If w is not the longest element of W, then there exists a BP-decomposition of w or w^{-1} with respect to J such that $P^J(v)$ is of the form $q^l + q^{l-1} + \cdots + q + 1$, where l is the length of v.*

If v satisfies the conditions of the above lemma, we will say that v is a ***chain element*** of W^J. Using the fact that Dynkin diagrams of type A or D are simply-laced, it is easy to deduce the following result from the above lemma.

Corollary 7 *Let W be a Weyl group of type A or D. If $w \in W$ is rationally smooth then there exists a BP-decomposition of w or w^{-1} with respect to $J = \Pi \setminus \{\alpha\}$ such that v is the longest element of $W_I^{I \cap J}$ for some $I \subset \Pi$ containing α.*

Using computers, we have found a nice property of palindromic intervals in maximal parabolic quotient groups of type E.

Proposition 8 *Let W be a Weyl group of type A,D and E and let $J = \Pi \setminus \{\alpha\}$, where α corresponds to a leaf in the Dynkin diagram. Then, v has palindromic lower interval in W^J if and only if there exists a subset I of Π containing α such that v is the longest element in $W_I^{I \cap J}$.*

Let's look at an example. Choose D_6 to be our choice of Weyl group and label the simple roots $\Pi = \{\alpha_1, \cdots, \alpha_6\}$ so that the labels match the corresponding nodes in the Dynkin diagram 1. If we set $J = \Pi \setminus \{\alpha_1\}$, then the list of $v \in W^J$ such that the lower interval in W^J being palindromic is:

$$id, s_1, s_1s_2, s_1s_2s_3, s_1s_2s_3s_4, s_1s_2s_3s_4s_5, s_1s_2s_3s_4s_6, s_1s_2s_3s_4s_5s_6s_4s_3s_2s_1.$$

Each of them are the longest elements of $W_I^{I \cap J}$, where I is the set of simple reflections appearing in v. One can see that the set of nodes I is connected inside the Dynkin diagram of D_6.

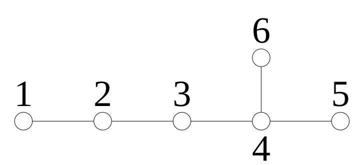


Fig. 1: Dynkin diagram of D_6

Now we will study how $R_w(q)$ behaves with respect to the BP-decomposition . Using the notations of Proposition 8, our first step is to prove that every reflection formed by simple reflections in $I \cap J$ is in $T_R(u)$. We need the following lemma to prove it:

Lemma 9 *Let $w \in W$ be a rationally smooth element and $w = uv$ be a BP-decomposition. Then every simple reflection in J appearing in the reduced word of v is a right descent of u.*

Proof: Multiplying $t \in T_L(w)$ to w corresponds to deleting one simple reflection in a certain reduced word of w. If we delete every simple reflection appearing in v but one in J, then the resulting element is in W_J and is below w. Hence by maximality of u, it is below u. □

Actually, we can state much more about u in terms of simple reflections of J appearing in v.

Lemma 10 *Let $w = uv$ be a BP-decomposition with respect to J. Let I be the subset of Π that appears in the reduced word of v. Then every reflection formed by simple reflections in $I \cap J$ is a right inversion reflection of u. In fact, there is a minimal length decomposition $u = u' u_{I\cap J}$ where $u_{I\cap J}$ is the longest element of $W_{I\cap J}$.*

Proof: Take the parabolic decomposition of u under the right quotient by $W_{I\cap J}$. Say, $u = u' u_{I\cap J}$. Then u' is the minimal length representative of u in $W/W_{I\cap J}$. For any simple reflection $s \in I \cap J$, the minimal length representative of us in $W/W_{I\cap J}$ is still u', hence the parabolic decomposition of us is $us = u'(u_{I\cap J} s)$. Since s is a right descent of u by Lemma 9, s is a right descent of $u_{I\cap J}$. Therefore $u_{I\cap J}$ is the longest element in $W_{I\cap J}$. The rest follows from this. □

The above lemma tells us that for each rationally smooth $w \in W$, we can decompose w or w^{-1} to $u' u_{I\cap J} v$ where uv is the BP-decomposition with respect to J, $u = u' u_{I\cap J}$ and $u_{I\cap J}$ is the longest element of $W_{I\cap J}$. Recall that we denote by Δ_w the inversion set of $w \in W$. For $I \subset \Pi$, we will denote Δ_I the set of roots of W_I. We have a decomposition

$$\Delta_w = \Delta_{u'} \sqcup u' \Delta_{u_{I\cap J}} {u'}^{-1} \sqcup u \Delta_v u^{-1}.$$

One can see that $\Delta_{u_{I\cap J}} = \Delta_{I\cap J}$ and $\Delta_v \subseteq \Delta_I \setminus \Delta_{I\cap J}$. And this tells us that $u' \Delta_{u_{I\cap J}} {u'}^{-1} = u \Delta_{I\cap J} u^{-1}$. By duality, let's assume we have decomposed some rationally smooth w as above. Let $\mathcal{A}_1, \mathcal{A}_0, \mathcal{A}_2$ denote the hyperplane arrangement coming from $u^{-1} \Delta_{u'} u, \Delta_{I\cap J}, \Delta_v$. We can study $\mathcal{A} := \mathcal{A}_1 \sqcup \mathcal{A}_0 \sqcup \mathcal{A}_2$ instead of looking at $\mathcal{A}_w$.

Lemma 11 *Let c be some chamber inside $\mathcal{A}_1 \sqcup \mathcal{A}_0$. Let c' be the chamber of $\mathcal{A}_0$ that contains c. Then the chamber graph of c with respect to $\mathcal{A}$ is isomorphic to the chamber graph of c' with respect to $\mathcal{A}_0 \sqcup \mathcal{A}_2$.*

Proof: Let c_1 and c_2 be two different chambers of $\mathcal{A}$ contained in c. They are separated by a hyperplane in $\mathcal{A}_2$. Let $c'_1(c'_2)$ be the chamber of $\mathcal{A}_0 \sqcup \mathcal{A}_2$ that contains $c_1(c_2)$. c'_1 and c'_2 are different chambers since they are separated by the hyperplane that separates c_1 and c_2. If c_1 and c_2 are adjacent, then c'_1 and c'_2 are adjacent. If c'_1 and c'_2 are adjacent but c_1 and c_2 are not, that means there is a hyperplane of $\mathcal{A}_1$ that separates c_1 and c_2. But that contradicts the fact that c_1 and c_2 are both contained in the same chamber of $\mathcal{A}_1 \sqcup \mathcal{A}_0$. So c_1 and c_2 are adjacent if and only if c'_1 and c'_2 are. From the fact that the distance from the fundamental chamber is equal to the number of hyperplanes that separate the chamber from the fundamental chamber, we see that the direction of the corresponding edges in the chamber graphs are the same.

Hence it is enough to show that the number of chambers of $\mathcal{A}$ in c equals number of chambers of $\mathcal{A}_0 \sqcup \mathcal{A}_2$ in c'. And this follows from showing that any chamber of $\mathcal{A}_0 \sqcup \mathcal{A}_2$ shares a common interior point with a chamber of $\mathcal{A}_0 \sqcup \mathcal{A}_1$ as long as they are contained in the same chamber of $\mathcal{A}_0$. To show this, we may include additional hyperplanes to $\mathcal{A}_1$ and $\mathcal{A}_2$. So let $\mathcal{A}_2$ be the hyperplane arrangement of $\Delta_I \setminus \Delta_{I\cap J}$ and $\mathcal{A}_1$ hyperplane arrangement of $\Delta \setminus \Delta_I$. Now $\mathcal{A}$ is just the Coxeter arrangement of W, and each chamber of $\mathcal{A}$ is indexed by $w \in W$.

We have a parabolic decomposition of W by $W_{I\cap J} W_I^{I\cap J} W^I$. Fixing a chamber c of $\mathcal{A}_0$ corresponds to fixing an element of $W_{I\cap J}$. In c, fixing a chamber $x(y)$ of $\mathcal{A}_0 \sqcup \mathcal{A}_2(\mathcal{A}_0 \sqcup \mathcal{A}_1)$ corresponds to fixing an element of $W_I^{I\cap J}(W^I)$. So given any such chamber x and y, we can find a chamber of $\mathcal{A}$ contained in them. This concludes the argument.

□

Corollary 12 *In the above decomposition, if $\mathcal{A}_{w'}$ is uniform with respect to $\mathcal{A}_u$ and v is the longest element of W_I^J, then $R_u(q) = P_u(q)$ implies $R_w(q) = P_w(q)$.*

Proof: If v is the longest element of W_I^J, then $w' := u_{I\cap J} v$ is the longest element of W_I. Then it is obvious that $\mathcal{A}_{w'}$ is uniform with respect to $\mathcal{A}_u$. Now it follows from above lemma that $R_w(q)/R_u(q) = R_{u_{I\cap J}v}(q)/R_{u_{I\cap J}}(q)$. Since we also know that the right hand side equals $P_v^{W^J}(q)$, $R_u(q) = P_u(q)$ implies $R_w(q) = P_w(q)$. □

In the next section, we will use the above lemma and corollary to prove the main theorem for type A,B,D and E.

4 The main Proof

In this section, we prove the main theorem. Type G case is trivial and omitted, type F case is done with a computer and is omitted in this extended abstract. For type A,D and E, the proof is very easy using Proposition 8 and Corollary 12.

Proposition 13 *Let W be a Weyl group of type A,D or E. Let w be a rationally smooth element. Then $R_w(q) = P_w(q)$.*

Proof: Decompose w or w^{-1} as in the remark preceding Lemma 11. By applying Proposition 8, we see that v is the longest element of W_I^J. Now we can apply Corollary 12. So we can replace w with some rationally smooth u that is contained in some Weyl group of type A,D or E with strictly smaller rank. Now the result follows from an obvious induction argument. □

For type B, we will use Lemma 6 and Lemma 11. Let's denote $\Pi = \{\alpha_0 = x_1, \alpha_1 = x_2 - x_1, \cdots, \alpha_n = x_{n+1} - x_n\}$. We will be studying $W^{\Pi\setminus\{\alpha_0\}}$ and $W^{\Pi\setminus\{\alpha_n\}}$. In both of them, if in the reduced word of v there is an adjacent commuting letters, then v is not a chain element. So when $J = \Pi \setminus \{\alpha_0\}$, the chain elements are

$$id, s_0, s_0 s_1, s_0 s_1 s_2, \cdots, s_0 s_1 \ldots s_n, s_0 s_1 s_0.$$

And when $J = \Pi \setminus \{\alpha_n\}$, the chain elements are

$$id, s_n, s_n s_{n-1}, \cdots, s_n s_{n-1} \ldots s_1 s_0, s_n s_{n-1} \ldots s_1 s_0 s_1, \cdots, s_n s_{n-1} \ldots s_1 s_0 s_1 \ldots s_{n-1} s_n.$$

Proposition 14 *Let W be a Weyl group of type B. Let w be a rationally smooth element. Then $R_w(q) = P_w(q)$.*

Proof: By Lemma 6, we may assume w or w^{-1} decomposes to uv where $u \in W_J$, $v \in W^J$, J is leaf-removed and v is a chain-element. Let's first show that when u is the longest element of W_J, then $\mathcal{A}_w$ is uniform with respect to $\mathcal{A}_u$ and $R_w(q) = P_w(q)$. Instead of looking at hyperplane arrangement coming from $\Delta_w = \Delta_u \sqcup u\Delta_v u^{-1}$, we can look at the hyperplane arrangement coming from $u^{-1}\Delta_w u = \Delta_u \sqcup \Delta_v$. So $\mathcal{A}_u$ consists of hyperplanes coming from Δ_u and $\mathcal{A}_v$ consists of hyperplanes coming from Δ_v.

When $J = \Pi \setminus \{\alpha_0\}$, we have $\Delta_v \subset \{x_1, \cdots, x_n\}$ and $|\Delta_v| = \ell(v)$. Choosing a chamber in $\mathcal{A}_u$ is equivalent to giving a total ordering on $\{x_1, \cdots, x_n\}$. Choosing a chamber in $\mathcal{A}_v$ is equivalent to assigning signs to roots of Δ_v. Given any total ordering on $\{x_1, \cdots, x_n\}$, there is a unique way to assign t number of $+$'s and $|v| - t$ number of $-$'s to Δ_v so that it is compatible with the total order on Δ_v. This tells us that $\mathcal{A}_w$ is uniform with respect to $\mathcal{A}_u$ and $R_w(q) = R_u(q)(1 + q + \cdots + q^{|v|}) = R_u(q)P_v^{W^J}(q)$. When $J = \Pi \setminus \{\alpha_n\}$, the proof is pretty much similar and is omitted.

Now let's return to the general case. Using Lemma 11 and above argument, we can replace w with some rationally smooth u that is contained in some Weyl group of type A or B with strictly smaller rank. Then the result follows from an obvious induction argument.

$\square$

5 Further remarks

As in Oh et al. (2008), our proof of the main theorem is based on a recurrence relation. It would be interesting to give a proof based on a bijection between elements of $[id, w]$ and regions of $\mathcal{A}_w$.

The statement of our main theorem can be extended to Coxeter groups. Although we don't have Schubert varieties for Coxeter groups, the Poincaré polynomial $P_w(q)$ can still be defined as the rank generating function of the interval $[id, w]$.

Conjecture 15 *Let W be any Coxeter group. Then $[id, w]$ is palindromic if and only if $P_w(q) = R_w(q)$.*

Our proof for the Weyl group case relied heavily on Theorem 5 and Proposition 8. Described a bit roughly, the former helps us to find the recurrence for $P_w(q)$ and the latter helps us to find the recurrence for $R_w(q)$. So the key would be to extending these two statements. For Proposition 8, it is easy to see that one direction holds for all Weyl groups. We give a slightly weakened statement that seems to hold for all Weyl groups.

Conjecture 16 *Let W be a Weyl group and let J be a maximal proper subset of the simple roots. Then, v has palindromic lower interval in W^J if and only if the interval is isomorphic to a maximal parabolic quotient of some Weyl group.*

Let's look at an example for the above conjecture. Choose F_4 to be our choice of Weyl group and label the simple roots $\Pi = \{\alpha_1, \cdots, \alpha_4\}$ so that the labels match the corresponding nodes in the Dynkin diagram 2. If we set $J = \Pi \setminus \{\alpha_4\}$, then the list of $v \in W^J$ such that the lower interval in W^J being palindromic is:

$$id, s_4, s_4s_3, s_4s_3s_2, s_4s_3s_2s_1, s_4s_3s_2s_3, s_4s_3s_2s_3s_4, s_4s_3s_2s_3s_1s_2s_3s_4s_3s_2s_3s_1s_2s_3s_4.$$

Those that do not correspond to longest elements of $W_I^{I \cap J}$ for some $I \subset \Pi$ are $s_4s_3s_2, s_4s_3s_2s_3$ and $s_4s_3s_2s_1$. But in these cases, hasse diagram of $[id, v]$ in W^J is a chain. So we can say that $[id, v]$ in W^J is isomorphic to a maximal parabolic quotient of a Weyl group of type A in these cases.

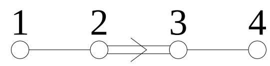


Fig. 2: Dynkin diagram of F_4

One nice property that $R_w(q)$ has is that it is always palindromic regarthless of the rational smoothness of w. And this is a property that intersection homology Poncaré polynomial $IP_w(q)$ also has.
So it would be interesting to compare these two polynomials.

Acknowledgements

We would like to thank Alexander Postnikov and Sara Billey for useful remarks.

References

S. C. Billey. Pattern avoidance and rational smoothness of schubert varieties. *Adv. Math*, 139:141–156, 1998.

S. C. Billey and A. Postnikov. Smoothness of schubert varieties via patterns in root subsystems. *Advances in Applied Mathematics*, 34:447–466, 2005.

S. C. Billey, , and V. Lakshmibai. *Singular Loci of Schubert Varieties*. Birkhäuser, 2000.

A. Björner and F. Brenti. *Combinatorics of Coxeter Groups*. Springer, 2005.

J. B. Carrell. The bruhat graph of a coxeter group, a conjecture of deodhar, and rational smoothness of schubert varieties. *Proceedings of Symposia in Pure Math*, 56:53–61, 1994.

S. Oh, A. Postnikov, and H. Yoo. Bruhat order, smooth schubert varieties, and hyperplane arrangement. *J.Combin.Theory Ser.A*, 115:1156–1166, 2008.

A. van den Hombergh. About the automorphisms of the bruhat-ordering in a coxeter group. *Indag. math*, 36:125–131, 1974.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 841–850

Bijective enumeration of permutations starting with a longest increasing subsequence

Greta Panova[1]

[1]*Harvard University*

Abstract. We prove a formula for the number of permutations in S_n such that their first $n-k$ entries are increasing and their longest increasing subsequence has length $n-k$. This formula first appeared as a consequence of character polynomial calculations in recent work of Adriano Garsia and Alain Goupil. We give two 'elementary' bijective proofs of this result and of its q-analogue, one proof using the RSK correspondence and one only permutations.

Résumé. Nous prouvons une formule pour le nombre des permutations dans S_n dont les prémiers $n-k$ entrées sont croissantes et dont la plus longue sous-súite croissante est de longeur $n-k$. Cette formule est d'abord apparue en conséquence de calculs sur les polynômes caractères des travaux récents de Adriano Garsia et Alain Goupil. Nous donnons deux preuves bijectifs 'élementaires' de cet résultat et de son q-analogue, une preuve employant le corréspondance RSK et une autre n'employant que les permutations.

Keywords: permutations, longest increasing subsequence, q-analogue, major index, RSK

1 Introduction

In [2], Adriano Garsia and Alain Goupil derived as a consequence of character polynomial calculations a simple formula for the enumeration of certain permutations. In his talk at the MIT Combinatorics Seminar [1], Garsia offered a \$100 award for an 'elementary' proof of this formula. We give such a proof of this formula and its q-analogue.

Let $\Pi_{n,k} = \{w \in S_n | w_1 < w_2 < \cdots < w_{n-k}, \mathrm{is}(w) = n-k\}$, the set of all permutations w in S_n, such that their first $n-k$ entries form an increasing sequence and the longest increasing sequence of w has length $n-k$; where we denote by $\mathrm{is}(w)$ the maximal length of an increasing subsequence of w.

The formula in question is the following theorem originally proven by A. Garsia and A.Goupil [2].

Theorem 1 *If $n \geq 2k$, the number of permutations in $\Pi_{n,k}$ is given by*

$$\#\Pi_{n,k} = \sum_{r=0}^{k} (-1)^{k-r} \binom{k}{r} \frac{n!}{(n-r)!}.$$

This formula has a $q-$analogue, also due to Garsia and Goupil.

Theorem 2 *For permutations in $\Pi_{n,k}$, if $n \geq 2k$, we have that*

$$\sum_{w \in \Pi_{n,k}} q^{\mathrm{maj}(w^{-1})} = \sum_{r=0}^{k} (-1)^{k-r} \binom{k}{r} [n]_q \cdots [n-r+1]_q,$$

where $\mathrm{maj}(\sigma) = \sum_{i|\sigma_i > \sigma_{i+1}} i$ denotes the major index of a permutation and $[n]_q = \frac{1-q^n}{1-q}$.

In this paper we will exhibit several bijections which will prove the above theorems. We will first define certain sets of permutations and pairs of tableaux which come in these bijections. We will then construct a relatively simple bijection showing a recurrence relation for the numbers $\#\Pi_{n,k}$. Using ideas from this bijection we will then construct a bijection proving Theorems 1 and 2 directly. We will also show a bijective proof which uses only permutations.

2 A few simpler sets and definitions.

Let rsk denote the RSK correspondence between permutations and pairs of tableaux [4], i.e. $\mathrm{rsk}(w) = (P, Q)$, where $w \in S_n$ and P and Q are standard Young tableaux (SYT) on $[n]$ and of the same shape, with P the insertion tableau and Q the recording tableau of w.

Let $C_{n,s} = \{w \in S_n | w_1 < w_2 < \cdots < w_{n-s}\}$ be the set of all permutations on $[n]$ with their first $n-s$ entries forming an increasing sequence. A permutation in $C_{n,s}$ is bijectively determined by the choice of the first $n-s$ elements from $[n]$ in $\binom{n}{s}$ ways and the arrangement of the remaining s in $s!$ ways, so

$$\#C_{n,s} = \binom{n}{s} s! = \frac{n!}{(n-s)!}.$$

Let $C^{\mathrm{rsk}}_{n,s} = \mathrm{rsk}(C_{n,s})$. Its elements are precisely the pairs of same-shape SYTs (P, Q) such that the first row of Q starts with $1, 2, \ldots, n-s$: the first $n-s$ elements are increasing and so will be inserted in this order in the first row, thereby recording their positions $1, 2, \ldots, n-s$ in Q in the first row also.

Let also $\Pi^{\mathrm{rsk}}_{n,s} = \mathrm{rsk}(\Pi_{n,s})$. Its elements are pairs of SYTs (P, Q), such that, as with $C^{\mathrm{rsk}}_{n,s}$, the first row of Q starts with $1, 2, \ldots, n-s$. By a theorem of Schensted, the length of the first row in P and Q is the length of the longest increasing subsequence of w, which is $n-s$ in the case of $\Pi_{n,s}$, so the first row of Q is exactly $1, 2, \ldots, n-s$. That is, $\Pi^{\mathrm{rsk}}_{n,s}$ is the set of pairs of same-shape SYTs (P, Q), such that the first row of Q has length $n-s$ and elements $1, 2, \ldots, n-s$.

Finally, let $D_{n,k,s}$ be the set of pairs of same-shape tableaux (P, Q), where P is an SYT on $[n]$ and Q is a tableau filled with $[n]$, with first row $1, 2, \ldots, n-k, a_1, \ldots, a_s, b_1, \ldots$ where $a_1 > a_2 > \cdots > a_s$, $b_1 < b_2 < \cdots$ and the remaining elements of Q are increasing in rows and down columns. Thus Q without its first row is an SYT. Notice that when $s = 0$ we just have $D_{n,k,0} = C^{\mathrm{rsk}}_{n,k}$.

The three sets of pairs we defined are determined by their Q tableaux as shown below.

1	$\cdots$	$n-s$	$\cdots$
	$\cdots$		
$\vdots$			

Q, for $(P, Q) \in C^{\mathrm{rsk}}_{n,s}$

1	$\cdots$	$n-s$
	$\cdots$	
$\vdots$		

Q, for $(P, Q) \in \Pi^{\mathrm{rsk}}_{n,s}$

1	$\cdots$	$n-k$	a_1	$\cdots$	a_s	$\cdots$
	$\cdots$					
$\vdots$						

Q, for $(P, Q) \in D_{n,k,s}$

3 A bijection.

We will exhibit a simple bijection, which will give us a recurrence relation for the numbers $\#\Pi_{n,k}$ equivalent to Theorem 1.

We should remark that while the recurrence can be inverted to give the inclusion-exclusion form of Theorem 1, the bijection itself does not succumb to direct inversion. However, the ideas of this bijection will lead us to discover the necessary construction for Theorem 1.

Proposition 1 *The number of permutations in $\Pi_{n,k}$ satisfies the following recurrence:*

$$\sum_{s=0}^{k} \binom{k}{s} \#\Pi_{n,s} = \binom{n}{k} k!$$

Proof:

Let $C^{\mathrm{rsk}}_{n,k,s}$ with $s \leq k$ be the set of pairs of same-shape tableaux (P, Q), such that the length of their first rows is $n-k+s$ and the first row of Q starts with $1, 2, \ldots, n-k$; clearly $C^{\mathrm{rsk}}_{n,k,s} \subset C^{\mathrm{rsk}}_{n,k}$. We have that

$$\bigcup_{s=0}^{k} C^{\mathrm{rsk}}_{n,k,s} = C^{\mathrm{rsk}}_{n,k}, \tag{1}$$

as $C^{\mathrm{rsk}}_{n,k}$ consists of the pairs (P, Q) with Q's first row starting with $1, \ldots, n-k$ and if $n-k+s$ is this first row's length then $(P, Q) \in C^{\mathrm{rsk}}_{n,k,s}$.

There is a bijection $C^{\mathrm{rsk}}_{n,k,s} \leftrightarrow \Pi^{\mathrm{rsk}}_{n,k-s} \times \binom{[n-k+1,\ldots,n]}{s}$ given as follows. If $(P, Q) \in C^{\mathrm{rsk}}_{n,k,s}$ and the first row of Q is $1, 2, \ldots, n-k, b_1, \ldots, b_s$, let

$$f : [n-k+1, \ldots, n] \setminus \{b_1, \ldots, b_s\} \to [n-k+s+1, \ldots, n]$$

be the order-preserving map. Let then Q' be the tableau obtained from Q by replacing every entry b not in the first row with $f(b)$ and the first row with $1, 2, \ldots, n-k, n-k+1, \ldots, n-k+s$. Then Q' is an SYT, since f is order-preserving and so the rows and columns are still increasing, first row included as its elements are smaller than any element below it. Then the bijection in question is $(P, Q) \leftrightarrow (P, Q', b_1, \ldots, b_s)$. Conversely, if $b_1, \ldots, b_s \in [n-k+1, \ldots, n]$ (in increasing order) and $(P, Q') \in \Pi^{\mathrm{rsk}}_{n,k-s}$, then replace all entries b below the first row of Q' with $f^{-1}(b)$ and the the first row of Q' with $1, 2, \ldots, n-k, b_1, \ldots, b_s$. We end up with a tableau Q, which is an SYT because: the entries below the first row preserve their order under f; and, since they are at most $k \leq n-k$, they are all below the first $n-k$ entries of the first row of Q (which are $1, 2, \ldots, n-k$, and thus smaller).

So we have that $\#C^{\mathrm{rsk}}_{n,k,s} = \binom{k}{s} \#\Pi^{\mathrm{rsk}}_{n,k-s}$ and substituting this into (1) gives us the statement of the lemma. $\square$

4 Proofs of the theorems.

We will prove Theorems 1 and 2 by exhibiting an inclusion-exclusion relation between the sets $\Pi^{\mathrm{rsk}}_{n,k}$ and $D_{n,k,s}$ for $s = 0, 1, \ldots, k$.

Proof of Theorem 1:

First of all, if $n \geq 2k$ we have a bijection $D_{n,k,s} \leftrightarrow C^{\mathrm{rsk}}_{n,k-s} \times \binom{[n-k+1,\dots,n]}{s}$, where the correspondence is $(P,Q) \leftrightarrow (P,Q') \times \{a_1,\dots,a_s\}$ given as follows.

Consider the order-preserving bijection

$$f : [n-k+1,\dots,n] \setminus \{a_1,\dots,a_s\} \to [n-k+s+1,\dots,n].$$

Then Q' is obtained from Q by replacing $a_1,\dots,a_s$ in the first row with $n-k+1,\dots,n-k+s$ and every other element b in Q, $b > n-k$ and $\neq a_i$, with $f(b)$. The first $n-k$ elements in the first row remain $1,2,\dots,n-k$. Since f is order-preserving, Q' without its first row remains an SYT (the inequalities within rows and columns are preserved). Since also $n-k \geq k$, we have that the second row of Q (and Q') has length less than or equal to k and hence $n-k$, so since the elements above the second row are among $1,2,\dots,n-k$ they are smaller than any element in the second row (which are all from $[n-k+1,\dots,n]$). Also, the remaining first row of Q' is increasing since it starts with $1,2,\dots,n-k,n-k+1,\dots,n-k+s$ and its remaining elements are in $[n-k+s+1,\dots,n]$ and are increasing because f is order-preserving. Hence Q' is an SYT with first row starting with $1,\dots,n-k+s$, so $(P,Q') \in C^{\mathrm{rsk}}_{n,k-s}$.

Conversely, if $(P,Q') \in C^{\mathrm{rsk}}_{n,k-s}$ and $\{a_1,\dots,a_s\} \in [n-k+1,\dots,n]$ with $a_1 > a_2 \cdots > a_s$, then we obtain Q from Q' by replacing $n-k+1,\dots,n=k+s$ with $a_1,\dots,a_s$ and the remaining elements $b > n-k$ with $f^{-1}(b)$, again preserving their order, and so $(P,Q) \in D_{n,k,s}$.

Hence, in particular,

$$\#D_{n,k,s} = \binom{k}{s} \#C^{\mathrm{rsk}}_{n,k-s} = \binom{k}{s} \#C_{n,k-s} = \binom{k}{s} \frac{n!}{(n-k+s)!}. \tag{2}$$

We have that $\Pi^{\mathrm{rsk}}_{n,k} \subset C^{\mathrm{rsk}}_{n,k}$ since $\Pi_{n,k} \subset C_{n,k}$. Then $C^{\mathrm{rsk}}_{n,k} \setminus \Pi^{\mathrm{rsk}}_{n,k}$ is the set of pairs of SYTs (P,Q) for which the first row of Q is $1,2,\dots,n-k,a_1,\dots$ for at least one a_1. So $E_{n,k,1} = C^{\mathrm{rsk}}_{n,k} \setminus \Pi^{\mathrm{rsk}}_{n,k}$ is then a subset of $D_{n,k,1}$. The remaining elements in $D_{n,k,1}$, that is $E_{n,k,2} = D_{n,k,1} \setminus E_{n,k,1}$, would be exactly the ones for which Q is not an SYT, which can happen only when the first row of Q is $1,2,\dots,n-k,a_1 > a_2,\dots$. These are now a subset of $D_{n,k,2}$ and by the same argument, we haven't included the pairs for which the first row of Q is $1,2,\dots,n-k,a_1 > a_2 > a_3,\dots$, which are now in $D_{n,k,3}$. Continuing in this way, if $E_{n,k,l+1} = D_{n,k,l} \setminus E_{n,k,l}$, we have that $E_{n,k,l}$ is the set of $(P,Q) \in D_{n,k,l}$, such that the first row of Q is $1,2,,\dots,n-k,a_1 > \cdots > a_l < \cdots$. Then $E_{n,k,l+1}$ is the subset of $D_{n,k,l}$, for which the element after a_l is smaller than a_l and so $E_{n,k,l+1} \subset D_{n,k,l+1}$. Finally, $E_{n,k,k} = D_{n,k,k}$ and $E_{n,k,k+1} = \varnothing$. We then have

$$\Pi^{\mathrm{rsk}}_{n,k} = C^{\mathrm{rsk}}_{n,k} \setminus (D_{n,k,1} \setminus (D_{n,k,2} \setminus \cdots \setminus (D_{n,k,k-1} \setminus D_{n,k,k}))), \tag{3}$$

or in terms of number of elements, applying (2), we get

$$\begin{aligned} \#\Pi_{n,k} = \#\Pi^{\mathrm{rsk}}_{n,k} &= \frac{n!}{(n-k)!} - \binom{k}{1}\frac{n!}{(n-k+1)!} + \binom{k}{2}\frac{n!}{(n-k+2)!} + \dots \\ &= \sum_{i=0}^{k} (-1)^i \binom{k}{i} \frac{n!}{(n-k+i)!}, \end{aligned}$$

which is what we needed to prove. $\square$

Theorem 2 will follow directly from (3) after we prove the following lemma.

Lemma 1 *We have that*

$$\sum_{w \in C_{n,s}} q^{\mathrm{maj}(w^{-1})} = [n]_q \dots [n-s+1]_q,$$

where $\mathrm{maj}(\sigma) = \sum_{i:\sigma_i > \sigma_{i+1}} i$ *denotes the major index of* σ.

Proof:

Let P be the poset on $[n]$ consisting of a chain $1, \dots, n-s$ and the single points $n-s+1, \dots, n$. Then $\sigma \in C_{n,s}$ if and only if $\sigma^{-1} \in \mathcal{L}(P)$, *i.e.* $\sigma : P \to [n]$ is a linear extension of P. We have that

$$\sum_{w \in C_{n,s}} q^{\mathrm{maj}(w^{-1})} = \sum_{w \in \mathcal{L}(P)} q^{\mathrm{maj}(w)};$$

denote this expression by $W_P(q)$. By theorem 4.5.8 from [3] on P-partitions, we have that

$$W_P(q) = G_P(q)(1-q)\dots(1-q^n), \tag{4}$$

where $G_P(q) = \sum_{m \geq 0} a(m) q^m$ with $a(m)$ denoting the number of P-partitions of m. That is, $a(m)$ is the number of order-reversing maps $\tau : P \to \mathbb{N}$, such that $\sum_{i \in P} \tau(i) = m$. In our particular case, these correspond to sequences $\tau(1), \tau(2), \dots$, whose sum is m and whose first $n-s$ elements are non-increasing. These correspond to partitions of at most $n-s$ parts and a sequence of s nonnegative integers, which add up to m. The partitions with at most $n-s$ parts are in bijection with the partitions with largest part $n-s$ (by transposing their Ferrers diagrams). The generating function for the latter is given by a well-known formula of Euler and is equal to

$$\frac{1}{(1-q)(1-q^2)\cdots(1-q^{n-s})}.$$

The generating function for the number of sequences of s nonnegative integers with a given sum is trivially $1/(1-q)^s$ and so we have that

$$G_P(q) = \frac{1}{(1-q)(1-q^2)\cdots(1-q^{n-s})} \frac{1}{(1-q)^s}.$$

After substitution in (4) we obtain the statement of the lemma. $\square$

Proof of Theorem 2.:

The descent set of a tableau T is the set of all i, such that $i+1$ is in a lower row than i in T, denote it by $D(T)$. By the properties of RSK (see e.g. [4], lemma 7.23.1) we have that the descent set of a permutation, $D(w) = \{i : w_i > w_{i+1}\}$ is the same as the descent set of its recording tableau, or by the symmetry of RSK, $D(w^{-1})$ is the same as the descent set of the insertion tableau P. Write $\mathrm{maj}(T) = \sum_{i \in D(T)} i$. Hence we have that

$$\sum_{w \in \Pi_{n,k}} q^{\mathrm{maj}(w^{-1})} = \sum_{w \in \Pi_{n,k}, \mathrm{rsk}(w)=(P,Q)} q^{\mathrm{maj}(P)} = \sum_{(P,Q) \in \Pi^{\mathrm{rsk}}_{n,k}} q^{\mathrm{maj}(P)} \tag{5}$$

From the proof of Theorem 1 we have the equality (3) on sets of pairs (P, Q),

$$\Pi^{\mathrm{rsk}}_{n,k} = C^{\mathrm{rsk}}_{n,k} \setminus (D_{n,k,1} \setminus (D_{n,k,2} \setminus \cdots \setminus (D_{n,k,k-1} \setminus D_{n,k,k}))),$$

or alternatively, $\Pi^{\mathrm{rsk}}_{n,k} = C^{\mathrm{rsk}}_{n,k} \setminus E_{n,k,1}$ and $E_{n,k,l} = D_{n,k,l} \setminus E_{n,k,l+1}$. Hence the statistic $q^{\mathrm{maj}(P)}$ on these sets will also respect the equalities between them; *i.e.* we have

$$\begin{aligned}\sum_{(P,Q)\in\Pi^{\mathrm{rsk}}_{n,k}} q^{\mathrm{maj}(P)} &= \sum_{(P,Q)\in C^{\mathrm{rsk}}_{n,k}\setminus E_{n,k,1}} q^{\mathrm{maj}(P)} \\ &= \sum_{(P,Q)\in C^{\mathrm{rsk}}_{n,k}} q^{\mathrm{maj}(P)} - \sum_{(P,Q)\in E_{n,k,1}} q^{\mathrm{maj}(P)} \\ &= \sum_{(P,Q)\in C^{\mathrm{rsk}}_{n,k}} q^{\mathrm{maj}(P)} - \sum_{(P,Q)\in D_{n,k,1}} q^{\mathrm{maj}(P)} + \sum_{(P,Q)\in E_{n,k,2}} q^{\mathrm{maj}(P)} = \cdots \\ &= \sum_{(P,Q)\in C^{\mathrm{rsk}}_{n,k}} q^{\mathrm{maj}(P)} - \sum_{(P,Q)\in D_{n,k,1}} q^{\mathrm{maj}(P)} + \cdots + (-1)^k \sum_{(P,Q)\in D_{n,k,k}} q^{\mathrm{maj}(P)}.\end{aligned} \tag{6}$$

Again, by the RSK correspondence, $\mathrm{maj}(P) = \mathrm{maj}(w^{-1})$ and Lemma 1 we have that

$$\sum_{(P,Q)\in C^{\mathrm{rsk}}_{n,k}} q^{\mathrm{maj}(P)} = \sum_{w\in C_{n,k}} q^{\mathrm{maj}(w^{-1})} = [n]_q \cdots [n-k+1]_q. \tag{7}$$

In order to evaluate $\sum_{(P,Q)\in D_{n,k,s}} q^{\mathrm{maj}(P)}$ we note that pairs $(P, Q) \in D_{n,k,s}$ are in correspondence with triples $(P, Q', \mathbf{a} = \{a_1, \ldots, a_s\})$, where P remains the same and $(P, Q') \in C^{\mathrm{rsk}}_{n,k-s}$. Hence

$$\begin{aligned}\sum_{(P,Q)\in D_{n,k,s}} q^{\mathrm{maj}(P)} &= \sum_{(P,Q',\mathbf{a})} q^{\mathrm{maj}(P)} \\ &= \sum_{\mathbf{a}\in\binom{[n-k+1,\ldots,n]}{s}} \sum_{(P,Q')\in C^{\mathrm{rsk}}_{n,k-s}} q^{\mathrm{maj}(P)} \\ &= \binom{k}{s}[n]_q \cdots [n-k+s+1]_q.\end{aligned} \tag{8}$$

Substituting the equations for (7) and (8) into (6) and comparing with (5) we obtain the statement of the theorem. □

We can apply the same argument for the preservation of the insertion tableaux and their descent sets to the bijection $T_{n,k,s} \leftrightarrow \Pi^{\mathrm{rsk}}_{n,k-s} \times \binom{[n-k+1,\ldots,n]}{s}$ in Proposition 1. We see that the insertion tableaux P in this bijection, $(P, Q) \leftrightarrow (P, Q', b_1, \ldots, b_s)$ remains the same and so do the corresponding descent sets and major indices

$$\sum_{(P,Q)\in T_{n,k,s}} q^{\mathrm{maj}(P)} = \binom{k}{s} \sum_{(P,Q')\in\Pi^{\mathrm{rsk}}_{n,k-s}} q^{\mathrm{maj}(P)} = \binom{k}{s} \sum_{w\in\Pi_{n,k-s}} q^{\mathrm{maj}(w^{-1})}.$$

Hence we have the following corollary to the bijection in Proposition 1 and Lemma 1.

Proposition 2 *We have that*

$$\sum_{s=0}^{k} \binom{k}{s} \sum_{w \in \Pi_{n,k-s}} q^{\mathrm{maj}(w^{-1})} = [n]_q \cdots [n-k+1]_q. \tag{9}$$

5 Permutations only.

Since the original question was posed only in terms of permutations, we will now give proofs of the main theorems without passing on to the pairs of tableaux. The constructions we will introduce is inspired from application of the inverse RSK to the pairs of tableaux considered in our proofs so far. However, since the pairs (P, Q) of tableaux in the sets $D_{n,k,s}$ are not pairs of Standard Young Tableaux we cannot apply directly the inverse RSK to the bijection in the proof of Theorem 1. This requires us to find new constructions and sets of permutations.

We will say that an increasing subsequence of length m of a permutation π satisfies the LLI-m (Least Lexicographic Indices) property if it is the first appearance of an increasing subsequence of length m (*i.e.* if a is the index of its last element, then $\bar{\pi} = \pi_1, \ldots, \pi_{a-1}$ has $\mathrm{is}(\bar{\pi}) < m$) and the indices of its elements are smallest lexicographically among all such increasing subsequences. For example, in $\pi = 2513467$, 234 is LLI-3. Let $n \geq 2s$ and let $C_{n,s,a}$ with $a \in [n-s+1, \ldots, n]$ be the set of permutations in $C_{n,s}$, for which there is an increasing subsequence of length $n-s+1$ and whose LLI-$(n-s+1)$ sequence has its last element at position a.

We define a map $\Phi : C_{n,s} \setminus \Pi^{\mathrm{rsk}}_{n,s} \to C_{n,s-1} \times [n-s+1, \ldots, n]$ for $n \geq 2s$ as follows. A permutation $\pi \in C_{n,s} \setminus \Pi^{\mathrm{rsk}}_{n,s}$ has a LLI-$(n-s+1)$ subsequence σ which would necessarily start with π_1 since $n \geq 2s$ and $\pi \in C_{n,s}$. Let $\sigma = \pi_1, \ldots, \pi_l, \pi_{i_{l+1}}, \ldots, \pi_{i_{n-s+1}}$ for some $l \geq 0$; if $a = i_{n-s+1}$ then $\pi \in C_{n,s,a}$. Let w be obtained from π by setting $w_{i_j} = \pi_{i_{j+1}}$ for $l+1 \leq j \leq n-s$, and then inserting $\pi_{i_{l+1}}$ right after π_l, all other elements preserve their (relative) positions. For example, if $\pi = 12684357 \in C_{8,4} \setminus \Pi^{\mathrm{rsk}}_{8,4}$, then 12457 is LLI-5, $a = 8$ and $w = 12468537$. Set $\Phi(\pi) = (w, a)$.

Lemma 2 *The map Φ is well-defined and injective. We have that*

$$C_{n,s-1} \times a \setminus \Phi(C_{n,s,a}) = \bigcup_{n-s+2 \leq b \leq a} C_{n,s-1,b}.$$

Proof: Let again $\pi \in C_{n,s,a}$ and $\Phi(\pi) = (w, a)$. It is clear by the LLI condition that we must have $\pi_{i_{l+1}} < \pi_{l+1}$ as otherwise

$$\pi_1, \ldots, \pi_{l+1}, \pi_{i_{l+1}}, \ldots, \pi_{i_{n-s}}$$

would be increasing of length $n-s+1$ and will have lexicographically smaller indices. Then the first $n-s+1$ elements of w will be increasing and $w \in C_{n,s-1}$.

To show injectivity and describe the coimage we will describe the inverse map $\Psi : \Phi(C_{n,s,a}) \to C_{n,s,a}$. Let $(w, a) \in \Phi(C_{n,s,a})$ with $(w, a) = \Phi(\pi)$ for some $\pi \in C_{n,s,a}$ and let $\bar{w} = w_1 \cdots w_a$.

Notice that $\bar{w}$ cannot have an increasing subsequence $\{y_i\}$ of length $n-s+2$. To show this, let $\{x_i\}$ be the subsequence of w which was the LLI-$(n-s+1)$ sequence of π. If there were a sequence $\{y_i\}$, this could have happened only involving the forward shifts of x_i and some of the x's and y's should coincide (in the beginning at least). By the pigeonhole principle there must be two pairs of indices $p_1 < q_1$

and $p_2 < q_2$ (q_1 and q_2 might be auxiliary, *i.e.* off the end of $\bar{w}$), such that in $\bar{w}$ we have $x_{p_1} = y_{p_2}$ and $x_{q_1} = y_{q_2}$ and between them there are strictly more elements of y and no more coincidences, *i.e.* $q_2 - p_2 > q_1 - p_1$. Then $x_{p_1-1} < x_{p_1} < y_{p_2+1}$ and in π (after shifting $\{x_i\}$ back) we will have the subsequence $x_1, \ldots, x_{p_1-1}, y_{p_2+1}, \ldots, y_{q_2-1}, x_{q_1} = y_{q_2}, \ldots, x_{n-s+1}$, which will be increasing and of length $p_1 - 1 + q_2 - p_2 + n - s + 1 - q_1 \geq n - s + 1$. By the LLI property we must have that y_{p_2+1} appears after x_{p_1}, but then $x_1, \ldots, x_{p_1-1}, x_{p_1}, y_{p_2+1}, \ldots, y_{q_2-1}, x_{q_1} = y_{q_2}, \ldots, x_{n-s}$ will be increasing of length at least $n-s+1$ appearing before $\{x_i\}$ in π. This violates the other LLI condition of no $n-s+1$ increasing subsequences before x_{n-s+1}.

Now let $\sigma = w_1, \ldots, w_r, w_{i_{r+1}}, \ldots, w_{i_{n-s+1}}$ with $i_{r+1} > n - s + 1$ be the $(n - s + 1)$-increasing subsequence of $\bar{w}$ with largest lexicographic index sequence. Let w' be obtained from w by assigning $w'_{i_j} = w_{i_{j-1}}$ for $r + 1 \leq j \leq n - s + 1$, where $i_r = r$, $w'_{a_1} = w_{i_{n-s+1}}$ and then deleting the entry w_r at position r.

We claim that the LLI-$(n - s + 1)$ sequence of w' is exactly σ. Suppose the contrary and let $\{y_i\}$ be the LLI-$(n - s + 1)$ subsequence of w'. Since $n \geq 2s$ we have $y_1 = w_1 = \sigma_1, \ldots, y_{i_j} = w_{i_j} = \sigma_j$ for all $1 \leq j \leq l$ for some $l \geq r$. If there are no more coincidences between y and σ afterwards, then the sequence $w_1 = y_1, \ldots, w_{i_l} = y_l, y_{l+1}, \ldots, y_{n-s+1}$ is increasing of length $n - s + 1$ in $\bar{w}$ (in the same order) and of lexicographically larger index than σ, since the index of y_{l+1} is after the index of y_l in w' equal to the index of σ_{l+1} in w.

Hence there must be at least one more coincidence, let $y_p = \sigma_q$ be the last such coincidence. Again, in $\bar{w}$ the sequence $\sigma_1, \ldots, \sigma_q, y_{p+1}, \ldots, y_{n-s+1}$ appears in this order and is increasing with the index of y_{p+1} larger than the one of σ_{q+1} in w, so its length must be at most $n-s$, *i.e.* $q+(n-s+1)-p \leq n-s$, so $q \leq p-1$. We see then that there are more y's between y_l and $y_p = \sigma_q$ than there are σ's there, so we can apply an argument similar to the one in the previous paragraph. Namely, there are indices p_1, q_1, p_2, q_2, such that $y_{p_1} = \sigma_{q_1}$, $y_{p_2} = \sigma_{q_2}$ with no other coincidences between them and $q_2 - q_1 < p_2 - p_1$. Then the sequence $\sigma_1, \ldots, \sigma_{q_1}, y_{p_1+1}, \ldots, y_{p_2-1}, \sigma_{q_2+1}, \ldots$ is increasing in this order in w, has length $q_1+p_2-1-p_1+n-s+1-q_2 = (n-s+1)+(p_2-p_1)-(q_2-q_1)-1 \geq n-s+1$ and the index of y_{p_1+1} in w is larger than the index of σ_{q_1+1} (which is the index of y_{p_1} in w'). We thus reach a contradiction, showing that we have found the inverse map of Φ is given by $\Psi(w, a) = w'$ and, in particular, that Φ is injective.

We have also shown that the image of Φ consists exactly of these permutations, which do not have an increasing subsequence of length $n - s + 2$ within their first a elements. Therefore the coimage of Φ is the set of permutations in $C_{n,s-1}$ with $n - s + 2$ increasing subsequence within its first a elements, so the ones in $C_{n,s-1,b}$ for $n - s + 2 \leq b \leq a$. □

We can now proceed to the proof of Theorem 1. We have that $C_{n,k} \setminus \Pi^{\mathrm{rsk}}_{n,k}$ is exactly the set of permutations in $C_{n,k}$ with some increasing subsequence of length $n - k + 1$, hence

$$C_{n,k} \setminus \Pi^{\mathrm{rsk}}_{n,k} = \bigcup_{n-k+1 \leq a_1 \leq n} C_{n,k,a_1}.$$

On the other hand, applying the lemma we have that

$$\begin{aligned}
\bigcup_{n-k+1\le a_1\le n} C_{n,k,a_1} \simeq \bigcup_{n-k+1\le a_1\le n} \Phi(C_{n,k,a_1}) \qquad (10) \\
= \bigcup_{n-k+1\le a_1\le n} \left(C_{n,k-1}\times a_1 \setminus \bigcup_{n-k+2\le a_2\le a_1} C_{n,k-1,a_2} \right) \\
= C_{n,k-1}\times \binom{[k]}{1} \setminus \left(\bigcup_{n-k+2\le a_2\le a_1\le n} C_{n,k-1,a_2}\times a_1 \right) \\
\simeq C_{n,k-1}\times \binom{[k]}{1} \setminus \left(C_{n,k-1}\times \binom{[k]}{2} \setminus \left(\bigcup_{n-k+3\le a_3\le a_2\le a_1\le n} C_{n,k-2,a_3}\times (a_2,a_1) \right)\right) \\
= \cdots = C_{n,k-1}\times \binom{[k]}{1} \setminus \left(C_{n,k-2}\binom{[k]}{2} \setminus \cdots \setminus \left(C_{n,k-r}\times \binom{[k]}{r} \setminus \ldots \right) \ldots \right),
\end{aligned}$$

where $\simeq$ denotes the equivalence under Φ and $\binom{[k]}{r}$ represent the $r-$tuples $(a_r,\ldots,a_1)$ where $n-k+r\le a_r\le a_{r-1}\le\cdots\le a_1\le n$. Since $\#C_{n,k-r}\times\binom{[k]}{r}=\binom{n}{k-r}(k-r)!\binom{k}{r}$, Theorem 1 follows.

As for the $q-$analogue, Theorem 2, it follows immediately from the set equalities (10) and Lemma 1 once we realize that the map Φ does not change the major index of the inverse permutation, as shown in the following small lemma.

Lemma 3 *Let* $D(w)=\{i+1$ *before* i *in* $w\}$. *Then* $D(w)=D(\Phi(w))$ *and thus* $\mathrm{maj}(w^{-1})=\sum_{i\in D(w)} i$ *remains the same after applying* Φ.

Proof: To see this, notice that i and $i+1$ could hypothetically change their relative order after applying Φ only if exactly one of them is in the LLI-$(n-s+1)$ sequence of w, denote this sequence by $\sigma = w_1,\ldots,w_{i_{n-s+1}}$.

Let $w_p=i$ and $w_q=i+1$. We need to check only the cases when $p=i_r$ and $i_{r-1}<q<i_r$ or $q=i_r$ and $i_{r-1}<p<i_r$, since otherwise i and $i+1$ preserve their relative order after shifting σ one step forward by applying Φ. In either case, we see that the sequence $w_1,\ldots,w_{i_{r-1}},w_p$ (or w_q), $w_{i_{r+1}},\ldots,w_{i_{n-s+1}}$ is increasing of length $n-s+1$ in w and has lexicographically smaller indices than σ, violating the LLI property. Thus these cases are not possible and the relative order of i and $i+1$ is preserved, so $D(w)=D(\Phi(w))$. $\square$

We now have that the equalities and equivalences in (10) are equalities on the sets $D(w)$ and so preserve the $\mathrm{maj}(w^{-1})$ statistic, leading directly to Theorem 2.

References

[1] Adriano Garsia, *A new recursion in the theory of Macdonald polynomials*, `http://math.ucsd.edu/~garsia/lectures/MIT-09-newRec.pdf`.

[2] A. M. Garsia and A. Goupil, *Character polynomials, their q-analogs and the Kronecker product*, Electron. J. Combin. **16** (2009), no. 2, Special volume in honor of Anders Bjorner, Research Paper 19, 40. MR2576382

[3] Richard P. Stanley, *Enumerative combinatorics. Vol. 1*, Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 1997. MR1442260 (98a:05001)

[4] ———, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999. MR1676282 (2000k:05026)

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 851–861

Cyclic sieving for longest reduced words in the hyperoctahedral group

T. K. Petersen[1] and L. Serrano[2†]

[1]*DePaul University, Chicago, IL*
[2]*University of Michigan, Ann Arbor, MI*

Abstract. We show that the set $R(w_0)$ of reduced expressions for the longest element in the hyperoctahedral group exhibits the cyclic sieving phenomenon. More specifically, $R(w_0)$ possesses a natural cyclic action given by moving the first letter of a word to the end, and we show that the orbit structure of this action is encoded by the generating function for the major index on $R(w_0)$.

Résumé. Nous montrons que l'ensemble $R(w_0)$ des expressions réduites pour l'élément le plus long du groupe hyperoctaédral présente le phénomène cyclique de tamisage. Plus précisément, $R(w_0)$ possède une action naturelle cyclique donnée par le déplacement de la première lettre d'un mot vers la fin, et nous montrons que la structure d'orbite de cette action est codée par la fonction génératrice pour l'indice majeur sur $R(w_0)$.

Resumen. En este artículo demostramos que el conjunto $R(w_0)$ de expresiones reducidas del elemento mas largo del grupo hiperoctaedro presenta el fenómeno de tamizado cíclico. Para ser mas precisos, $R(w_0)$ posee una acción cíclica natural dada por el movimiento de la primera letra de una palabra al final, y nosotros mostramos que la estructura de las orbitas de esta acción está codificada por la función generatriz del indice mayor en $R(w_0)$.

Keywords: Cyclic sieving, hyperoctahedral group, standard Young tableau, shifted staircase, reduced word

1 Introduction and main result

Suppose we are given a finite set X, a finite cyclic group $C = \langle\omega\rangle$ acting on X, and a polynomial $X(q) \in \mathbb{Z}[q]$ with integer coefficients. Following Reiner, Stanton, and White [RSW], we say that the triple $(X, C, X(q))$ exhibits the *cyclic sieving phenomenon* (CSP) if for every integer $d \geq 0$, we have that $|X^{\omega^d}| = X(\zeta^d)$ where $\zeta \in \mathbb{C}$ is a root of unity of multiplicitive order $|C|$ and X^{ω^d} is the fixed point set of the action of the power ω^d. In particular, since the identity element fixes everything in any group action, we have that $|X| = X(1)$ whenever $(X, C, X(q))$ exhibits the CSP.

If the triple $(X, C, X(q))$ exhibits the CSP and ζ is a primitive $|C|^{th}$ root of unity, we can determine the cardinalities of the fixed point sets $X^1 = X$, X^ω, $X^{\omega^2}, \ldots, X^{\omega^{|C|-1}}$ via the polynomial evaluations $X(1), X(\zeta), X(\zeta^2), \ldots, X(\zeta^{|C|-1})$. These fixed point set sizes determine the cycle structure of the canonical image of ω in the group of permutations of X, S_X. Therefore, to find the cycle structure of the

[†]Partially supported by NSF grant DMS-0555880

image of any bijection $\omega : X \to X$, it is enough to determine the order of the action of ω on X and find a polynomial $X(q)$ such that $(X, \langle c\rangle, X(q))$ exhibits the CSP.

The cyclic sieving phenomenon has been demonstrated in a variety of contexts. The paper of Reiner, Stanton, and White [RSW] itself includes examples involving noncrossing partitions, triangulations of polygons, and cosets of parabolic subgroups of Coxeter groups. An example of the CSP with standard Young tableaux is due to Rhoades [Rh] and will discussed further in Section 4. Now we turn to the CSP of interest to this note.

Let $w_0 = w_0^{(B_n)}$ denote the longest element in the type B_n Coxeter group. Given generating set $S = \{s_1, \ldots, s_n\}$ for B_n, (s_1 being the "special" reflection), we will write a reduced expression for w_0 as a word in the subscripts. For example, $w_0^{(B_3)}$ can be written as

$$s_1 s_2 s_1 s_3 s_2 s_3 s_1 s_2 s_3;$$

we will abbreviate this product by 121323123. It turns out that if we cyclically permute these letters, we always get another reduced expression for w_0. Said another way, $s_i w_0 s_i = w_0$ for $i = 1, \ldots, n$. The same is not true for longest elements of other classical types. In type A, we have $s_i w_0^{(A_n)} s_{n+1-i} = w_0^{(A_n)}$, and for type D,

$$w_0^{(D_n)} = \begin{cases} s_i w_0^{(D_n)} s_i & \text{if } n \text{ even or } i > 2, \\ s_i w_0^{(D_n)} s_{3-i} & \text{if } n \text{ odd and } i = 1, 2. \end{cases}$$

Let $R(w_0)$ denote the set of reduced expressions for w_0 in type B_n and let $c : R(w_0) \to R(w_0)$ denote the action of placing the first letter of a word at the end. Then the orbit in $R(w_0^{(B_3)})$ of the word above is:

$$\{121323123 \to 213231231 \to 132312312 \to 323123121 \to 231231213$$

$$\to 312312132 \to 123121323 \to 231213231 \to 312132312\}.$$

As the length of w_0 is n^2, we clearly have $c^{n^2} = 1$, and the size of any orbit divides n^2. For an example of a smaller orbit, notice that the word 213213213 has cyclic order 3.

For any word $w = w_1 \ldots w_l$, (e.g., a reduced expression for w_0), a *descent* of w is defined to be a position i in which $w_i > w_{i+1}$.The *major index* of w, $\mathrm{maj}(w)$, is defined as the sum of the descent positions. For example, the word $w = 121323123$ has descents in positions $2, 4$, and 6, so its major index is $\mathrm{maj}(w) = 2 + 4 + 6 = 12$. Let $f_n(q)$ denote the generating function for this statistic on words in $R(w_0)$:

$$f_n(q) = \sum_{w \in R(w_0)} q^{\mathrm{maj}(w)}.$$

The following is our main result.

Theorem 1 *The triple $(R(w_0), \langle c\rangle, X(q))$ exhibits the cyclic sieving phenomenon, where*

$$X(q) = q^{-n\binom{n}{2}} f_n(q).$$

For example, let us consider the case $n = 3$. We have

$$\begin{aligned} X(q) &= q^{-9} \sum_{w \in R(w_0^{(B_3)})} q^{\mathrm{maj}(w)} \\ &= 1 + q^2 + 2q^3 + 2q^4 + 2q^5 + 4q^6 + 3q^7 + 4q^8 + 4q^9 \\ &\quad +4q^{10} + 3q^{11} + 4q^{12} + 2q^{13} + 2q^{14} + 2q^{15} + q^{16} + q^{18}. \end{aligned}$$

Let $\zeta = e^{\frac{2\pi i}{9}}$. Then we compute:

$$\begin{array}{lll} X(1) = 42 & X(\zeta^3) = 6 & X(\zeta^6) = 6 \\ X(\zeta) = 0 & X(\zeta^4) = 0 & X(\zeta^7) = 0 \\ X(\zeta^2) = 0 & X(\zeta^5) = 0 & X(\zeta^8) = 0 \end{array}$$

Thus, the 42 reduced expressions for $w_0^{(B_3)}$ split into two orbits of size three (the orbits of 123123123 and 132132132) and four orbits of size nine.

We prove Theorem 1 by relating it to another instance of the CSP, namely Rhoades' recent (and deep) result [Rh, Thm 3.9] for the set $SYT(n^m)$ of rectangular standard Young tableaux with respect to the action of *promotion* (defined in Section 2). To make the connection, we rely on a pair of remarkable bijections due to Haiman [H1, H2]. The composition of Haiman's bijections maps to $R(w_0)$ from the set of square tableaux, $SYT(n^n)$. In this note our main goal is to show that Haiman's bijections carry the orbit structure of promotion on $SYT(n^n)$ to the orbit structure of c on $R(w_0)$.

We conclude this section by remarking that this approach was first outlined by Rhoades [Rh, Thm 8.1]. One purpose of this article is to fill some nontrivial gaps in his argument. A second is to justify the new observation that the polynomial $X(q)$ can be expressed as the generating function for the major index on $R(w_0)$.

2 Promotion on standard Young tableaux

For λ a partition, let $SYT(\lambda)$ denote the set of standard Young tableaux of shape λ. If λ is a *strict partition*, i.e., with no equal parts, then let $SYT'(\lambda)$ denote the set of standard Young tableaux of shifted shape λ. We now describe the action of *jeu de taquin promotion*, first defined by Schützenberger [Sch].

We will consider promotion as a permutation of tableaux of a fixed shape (resp. shifted shape), $p : SYT(\lambda) \to SYT(\lambda)$ (resp. $p : SYT'(\lambda) \to SYT'(\lambda)$). Given a tableau T with $\lambda \vdash n$, we form $p(T)$ with the following algorithm. (We denote the entry in row a, column b of a tableau T, by $T_{a,b}$.)

1. Remove the entry 1 in the upper left corner and decrease every other entry by 1. The empty box is initialized in position $(a, b) = (1, 1)$.

2. Perform jeu de taquin:

   (a) If there is no box to the right of the empty box and no box below the empty box, then go to 3).

   (b) If there is a box to the right or below the empty box, then swap the empty box with the box containing the smaller entry, i.e., $p(T)_{a,b} := \min\{T_{a,b+1}, T_{a+1,b}\}$. Set $(a, b) := (a', b')$, where (a', b') are the coordinates of box swapped, and go to 2a).

3. Fill the empty box with n.

Here is an example:

$$T = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 4 & 8 \\ \hline 3 & 6 & 7 \\ \cline{1-3} 5 \\ \cline{1-1}\end{array} \quad \mapsto \quad \begin{array}{|c|c|c|c|}\hline 1 & 3 & 6 & 7 \\ \hline 2 & 5 & 8 \\ \cline{1-3} 4 \\ \cline{1-1}\end{array} = p(T).$$

As a permutation, promotion naturally splits $SYT(\lambda)$ into disjoint orbits. For a general shape λ there seems to be no obvious pattern to the sizes of the orbits. However, for certain shapes, notably Haiman's "generalized staircases" more can be said [H2] (see also Edelman and Greene [EG, Cor. 7.23]). In particular, rectangles fall into this category, with the following result.

Theorem 2 ([H2], Theorem 4.4) *If $\lambda \vdash N = bn$ is a rectangle, then $p^N(T) = T$ for all $T \in SYT(\lambda)$.*

Thus for $n \times n$ square shapes λ, $p^{n^2} = 1$ and the size of every orbit divides n^2. With $n = 3$, here is an orbit of size 3:

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 6 & 8 \\ \hline 4 & 7 & 9 \\ \hline\end{array} \to \begin{array}{|c|c|c|}\hline 1 & 4 & 7 \\ \hline 2 & 5 & 8 \\ \hline 3 & 6 & 9 \\ \hline\end{array} \to \begin{array}{|c|c|c|}\hline 1 & 3 & 6 \\ \hline 2 & 4 & 7 \\ \hline 5 & 8 & 9 \\ \hline\end{array} \to \cdots. \tag{1}$$

There are 42 standard Young tableaux of shape $(3, 3, 3)$, and there are 42 reduced expressions in the set $R(w_0^{(B_3)})$. Stanley first conjectured that $R(w_0^{(B_3)})$ and $SYT(n^n)$ are equinumerous, and Proctor suggested that rather than $SYT(n^n)$, a more direct correspondence might be given with $SYT'(2n - 1, 2n - 3, \ldots, 1)$, that is, with shifted standard tableaux of "doubled staircase" shape. (That the squares and doubled staircases are equinumerous follows easily from hook length formulas.)

Haiman answers Proctor's conjecture in such a way that the structure of promotion on doubled staircases corresponds precisely to cyclic permutation of words in $R(w_0)$ [H2, Theorem 5.12]. Moreover, in [H1, Proposition 8.11], he gives a bijection between standard Young tableaux of square shape and those of doubled staircase shape that (as we will show) commutes with promotion.

As an example, his bijection carries the orbit in 1 to this shifted orbit:

$$\begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 4 & 5 & 8 \\ \hline \multicolumn{1}{c|}{} & 3 & 6 & 9 \\ \cline{2-4} \multicolumn{1}{c}{} & \multicolumn{1}{c|}{} & 7 \\ \cline{3-3}\end{array} \to \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 3 & 4 & 7 \\ \hline \multicolumn{1}{c|}{} & 5 & 6 & 8 \\ \cline{2-4} \multicolumn{1}{c}{} & \multicolumn{1}{c|}{} & 9 \\ \cline{3-3}\end{array} \to \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 3 & 6 & 9 \\ \hline \multicolumn{1}{c|}{} & 4 & 5 & 7 \\ \cline{2-4} \multicolumn{1}{c}{} & \multicolumn{1}{c|}{} & 8 \\ \cline{3-3}\end{array} \to \cdots.$$

Both of these orbits of tableaux correspond to the orbit of the reduced word 132132132.

3 Haiman's bijections

We first describe the bijection between reduced expressions and shifted standard tableaux of doubled staircase shape. This bijection is described in Section 5 of [H2].

Let T in $SYT'(2n - 1, 2n - 3, \ldots, 1)$. Notice the largest entry in T, (i.e., n^2), occupies one of the outer corners. Let $r(T)$ denote the row containing this largest entry, numbering the rows from the bottom up. The *promotion sequence* of T is defined to be $\Phi(T) = r_1 \cdots r_{n^2}$, where $r_i = r(p^i(T))$. Using the example above of

$$T = \begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 4 & 5 & 8 \\ \hline \multicolumn{1}{c|}{} & 3 & 6 & 9 \\ \cline{2-4} \multicolumn{1}{c}{} & \multicolumn{1}{c|}{} & 7 \\ \cline{3-3}\end{array},$$

we see $r(T) = 2, r(p(T)) = 1, r(p^2(T)) = 3$, and since $p^3(T) = T$, we have

$$\Phi(T) = 132132132.$$

Haiman's result is the following.

Theorem 3 ([H2], Theorem 5.12) *The map $T \mapsto \Phi(T)$ is a bijection $SYT'(2n-1, 2n-3, \ldots, 1) \to R(w_0)$.*

By construction, one can see that applying promotion to T will cyclically shift the letters in $\Phi(T)$. Therefore, we have

$$\Phi(p(T)) = c(\Phi(T)),$$

i.e., Φ is an orbit-preserving bijection

$$(SYT'(2n-1, 2n-3, \ldots, 1), p) \longleftrightarrow (R(w_0), c).$$

Next, we will describe the bijection

$$H : SYT(n^n) \to SYT'(2n-1, 2n-3, \ldots, 1)$$

between squares and doubled staircases. Though not obvious from the definition below, we will demonstrate that H commutes with promotion.

We assume the reader is familiar with the *Robinson-Schensted-Knuth insertion* algorithm (RSK). (See [Sta, Section 7.11], for example.) This is a map between words w and pairs of tableaux $(P, Q) = (P(w), Q(w))$. We say P is the *insertion tableau* and Q is the *recording tableau*.

There is a similar correspondence between words w and pairs of shifted tableaux $(P', Q') = (P'(w), Q'(w))$ called *shifted mixed insertion* due to Haiman [H1]. (See also Sagan [Sa] and Worley [W].) Serrano defined a semistandard generalization of shifted mixed insertion in [Ser]. Throughout this paper we refer to semistandard shifted mixed insertion simply as *mixed insertion*. Details can be found in [Ser, Section 1.1].

Theorem 4 ([Ser] Theorem 2.26) *Let w be a word. If we view $Q(w)$ as a skew shifted standard Young tableau and apply jeu de taquin to obtain a standard shifted Young tableau, the result is $Q'(w)$ (independent of any choices in applying jeu de taquin).*

For example, if $w = 332132121$, then

$$(P, Q) = \left(\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 2 & 2 & 2 \\ \hline 3 & 3 & 3 \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 3 & 6 & 8 \\ \hline 4 & 7 & 9 \\ \hline \end{array} \right),$$

$$(P', Q') = \left(\begin{array}{ccccc} 1 & 1 & 1 & 2' & 3' \\ & 2 & 2 & 3' & \\ & & 3 & & \end{array}, \begin{array}{ccccc} 1 & 2 & 4 & 5 & 8 \\ & 3 & 6 & 9 & \\ & & 7 & & \end{array} \right).$$

Performing jeu de taquin we see:

$$\begin{array}{ccccc} \square & \square & 1 & 2 & 5 \\ & \square & 3 & 6 & 8 \\ & & 4 & 7 & 9 \end{array} \to \begin{array}{ccccc} \square & \square & 1 & 2 & 5 \\ & 3 & 4 & 6 & 8 \\ & & 7 & 9 & \end{array} \to \begin{array}{ccccc} \square & 1 & 2 & 5 & 8 \\ & 3 & 4 & 6 & \\ & & 7 & 9 & \end{array} \to \begin{array}{ccccc} 1 & 2 & 4 & 5 & 8 \\ & 3 & 6 & 9 & \\ & & 7 & & \end{array}.$$

Haiman's bijection is precisely $H(Q) = Q'$. That is, given a standard square tableau Q, we embed it in a shifted shape and apply jeu de taquin to create a standard shifted tableau. That this is indeed a bijection follows from Theorem 4, but is originally found in [H1, Proposition 8.11].

Remark 5 *Haiman's bijection H applies more generally between rectangles and "shifted trapezoids", i.e., for $m \leq n$, we have $H : SYT(n^m) \to SYT'(n+m-1, n+m-3, \ldots, n-m+1)$. All the results presented here extend to this generality, with similar proofs. We restict to squares and doubled staircases for clarity of exposition.*

We will now fix the tableaux P and P' to ensure that the insertion word w has particularly nice properties. We will use the following lemma.

Lemma 6 ([Ser], Proposition 1.8) *Fix a word w. Let $P = P(w)$ be the RSK insertion tableau and let $P' = P'(w)$ be the mixed insertion tableau. Then the set of words that mixed insert into P' is contained in the set of words that RSK insert into P.*

Now we apply Lemma 6 to the word

$$w = \underbrace{n \cdots n}_{n} \cdots \underbrace{2 \cdots 2}_{n} \underbrace{1 \cdots 1}_{n}.$$

If we use RSK insertion, we find P is an $n \times n$ square tableau with all 1s in row first row, all 2s in the second row, and so on. With such a choice of P it is not difficult to show that any other word u inserting to P has the property that in any initial subword $u_1 \cdots u_i$, there are at least as many letters $(j+1)$ as letters j. Such words are sometimes called (reverse) *lattice words* or (reverse) *Yamanouchi words*. Notice also that any such u has n copies of each letter i, $i = 1, \ldots, n$. We call the words inserting to this choice of P *square words*.

On the other hand, if we use mixed insertion on w, we find P' as follows (with $n = 4$):

$$\begin{array}{|c|c|c|c|c|c|c|}
\hline
1 & 1 & 1 & 1 & 2' & 3' & 4' \\
\hline
 & 2 & 2 & 2 & 3' & 4' & \\
\hline
 & & 3 & 3 & 4' & & \\
\hline
 & & & 4 & & & \\
\hline
\end{array}\ .$$

In general, on the "shifted half" of the tableau we see all 1s in the first row, all 2s in the second row, and so on. In the "straight half" we see only prime numbers, with $2'$ on the first diagonal, $3'$ on the second diagonal, and so on. Lemma 6 tells us that every u that mixed inserts to P' is a square word. But since the sets of recording tableaux for P and for P' are equinumerous, we see that the set of words mixed inserting to P' is precisely the set of all square words.

Remark 7 *Yamanouchi words give a bijection with square standard Young tableaux that circumvents insertion completely. In reading the word from left to right, if $w_i = j$, we put letter i in the leftmost unoccupied position of row $n+1-j$. (See [Sta, Proposition 7.10.3(d)].)*

We will soon characterize promotion in terms of operators on insertion words. First, some lemmas.

For a tableau T (shifted or not) let ΔT denote the result of all but step (3) of promotion. That is, we delete the smallest entry and perform jeu de taquin, but we do not fill in the empty box. The following lemma says that, in both the shifted and unshifted cases, this can be expressed very simply in terms of our insertion word. The first part of the lemma is a direct application of the theory of jeu de taquin (see, e.g., [Sta, A1.2]); the second part is [Ser, Lemma 3.9].

Lemma 8 *For a word $w = w_1 w_2 \cdots w_l$, let $\widehat{w} = w_2 \cdots w_l$. Then we have*

$$Q(\widehat{w}) = \Delta Q(w),$$

and

$$Q'(\widehat{w}) = \Delta Q'(w).$$

The operator e_j acting on words $w = w_1 \cdots w_l$ is defined in the following way. Consider the subword of w formed only by the letters j and $j+1$. Consider every $j+1$ as an opening bracket and every j as a closing bracket, and pair them up accordingly. The remaining word is of the form $j^r(j+1)^s$. The operator e_j leaves all of w invariant, except for this subword, which it changes to $j^{r-1}(j+1)^{s+1}$ (assuming $r > 0$, otherwise e_j is not defined on w). This operator is widely used in the theory of *crystal graphs*.

As an example, we calculate $e_2(w)$ for the word $w = 3121221332$. The subword formed from the letters 3 and 2 is

$$3 \cdot 2 \cdot 22 \cdot 332,$$

which corresponds to the bracket sequence $()))(()$. Removing paired brackets, one obtains $))($, corresponding to the subword

$$\cdots\cdot 22 \cdot 3 \cdot\cdot.$$

We change the last 2 to a 3 and keep the rest of the word unchanged, obtaining $e_2(w) = 31212\mathbf{3}1332$.

The following lemma shows that this operator leaves the recording tableau unchanged. The unshifted case is found in work of Lascoux, Leclerc, and Thibon [LLT, Theorem 5.5.1]; the shifted case follows from the unshifted case, and the fact that the mixed recording tableau of a word is uniquely determined by its RSK recording tableau (Theorem 4).

Lemma 9 *Recording tableaux are invariant under the operators e_i. That is,*

$$Q(e_i(w)) = Q(w),$$

and

$$Q'(e_i(w)) = Q'(w).$$

Let $\overline{e} = e_1 \cdots e_{n-1}$ denote the composite operator given by applying first e_{n-1}, then e_{n-2} and so on. It is clear that if $w = w_1 \cdots w_{n^2}$ is a square word, then $\overline{e}(\widehat{w})1$ is again a square word.

Theorem 10 *Let $w = w_1 \cdots w_{n^2}$ be a square word. Then,*

$$p(Q(w)) = Q(\overline{e}(\widehat{w})1),$$

and

$$p(Q'(w)) = Q'(\overline{e}(\widehat{w})1).$$

In other words, Haiman's bijection commutes with promotion:

$$p(H(Q)) = H(p(Q)).$$

Proof: By Lemma 8, we see that $Q(\widehat{w})$ is only one box away from $p(Q(w))$. Further, repeated application of Lemma 9 shows that

$$Q(\widehat{w}) = Q(e_{n-1}(\widehat{w})) = Q(e_{n-2}(e_{n-1}(\widehat{w}))) = \cdots = Q(\overline{e}(\widehat{w})).$$

The same lemmas apply show $Q'(\overline{e}(\widehat{w}))$ is one box away from $p(Q'(w))$.

All that remains is to check that the box added by inserting 1 into $P(\overline{e}(\widehat{w}))$ (resp. $P'(\overline{e}(\widehat{w}))$) is in the correct position. But this follows from the observation that $\overline{e}(\widehat{w})1$ is a square word, and square words insert (resp. mixed insert) to squares (resp. doubled staircases). $\square$

4 Rhoades' result

Rhoades [Rh] proved an instance of the CSP related to the action of promotion on rectangular tableaux. His result is quite deep, employing Kahzdan-Lusztig cellular representation theory in its proof.

Recall that for any partition $\lambda \vdash n$, we have that the standard tableaux of shape λ are enumerated by the Frame-Robinson-Thrall *hook length formula*:

$$f^\lambda = |SYT(\lambda)| = \frac{n!}{\prod_{(i,j)\in\lambda} h_{ij}},$$

where the product is over the boxes (i,j) in λ and h_{ij} is the *hook length* at the box (i,j), i.e., the number of boxes directly east or south of the box (i,j) in λ, counting itself exactly once. To obtain the polynomial used for cyclic sieving, we replace the hook length formula with a natural q-analogue. First, recall that for any $n \in \mathbb{N}$, $[n]_q := 1 + q + \cdots + q^{n-1}$ and $[n]_q! := [n]_q[n-1]_q \cdots [1]_q$.

Theorem 11 ([Rh], Theorem 3.9) *Let $\lambda \vdash N$ be a rectangular shape and let $X = SYT(\lambda)$. Let $C := \mathbb{Z}/N\mathbb{Z}$ act on X via promotion. Then, the triple $(X, C, X(q))$ exhibits the cyclic sieving phenomenon, where*

$$X(q) = \frac{[N]_q!}{\Pi_{(i,j)\in\lambda}[h_{ij}]_q}$$

is the q-analogue of the hook length formula.

Now thanks to Theorem 10 we know that H preserves orbits of promotion, and as a consequence we see the CSP for doubled staircases.

Corollary 12 *Let $X = SYT'(2n-1, 2n-3, \ldots, 1)$, and let $C := \mathbb{Z}/n^2\mathbb{Z}$ act on X via promotion. Then the triple $(X, C, X(q))$ exhibits the cyclic sieving phenomenon, where*

$$X(q) = \frac{[n^2]_q!}{[n]_q^n \prod_{i=1}^{n-1}([i]_q \cdot [2n-i]_q)^i}$$

is the q-analogue of the hook length formula for an $n \times n$ square Young diagram.

Because of Theorem 3 the set $R(w_0)$ also exhibits the CSP.

Corollary 13 ([Rh], Theorem 8.1) *Let $X = R(w_0)$ and let $X(q)$ as in Corollary 12. Let $C := \mathbb{Z}/n^2\mathbb{Z}$ act on X by cyclic rotation of words. Then the triple $(X, C, X(q))$ exhibits the cyclic sieving phenomenon.*

Corollary 13 is the CSP for $R(w_0)$ as stated by Rhoades. This is nearly our main result (Theorem 1), but for the definition of $X(q)$.

In spirit, if $(X, C, X(q))$ exhibits the CSP, the polynomial $X(q)$ should be some q-enumerator for the set X. That is, it should be expressible as

$$X(q) = \sum_{x \in X} q^{s(x)},$$

where s is an intrinsically defined statistic for the elements of X. Indeed, nearly all known instances of the cyclic sieving phenomenon have this property. For example, it is known ([Sta, Cor 7.21.5]) that the q-analogue of the hook-length formula can be expressed as follows:

$$f^{\lambda}(q) = q^{-\kappa(\lambda)} \sum_{T \in SYT(\lambda)} q^{\mathrm{maj}(T)}, \tag{2}$$

where $\kappa(\lambda_1, \ldots, \lambda_l) = \sum_{1 \le i \le l} (i-1)\lambda_i$ and for a tableau T, $\mathrm{maj}(T)$ is the sum of all i such that i appears in a row above $i+1$. Thus $X(q)$ in Theorem 11 can be described in terms of a statistic on Young tableaux.

With this point of view, Corollaries 12 and 13 are aesthetically unsatisfying. Section 5 is given to showing that $X(q)$ can be defined as the generating function for the major index on words in $R(w_0)$. It would be interesting to find a combinatorial description for $X(q)$ in terms of a statistic on $SYT'(2n-1, 2n-3, \ldots, 1)$ as well, though we have no such description at present.

5 Combinatorial description of $X(q)$

As stated in the introduction, we will show that

$$X(q) = q^{-n\binom{n}{2}} \sum_{w \in R(w_0)} q^{\mathrm{maj}(w)}.$$

If we specialize equation (2) to square shapes, we see that $\kappa(n^n) = n\binom{n}{2}$ and

$$X(q) = q^{-n\binom{n}{2}} \sum_{T \in SYT(n^n)} q^{\mathrm{maj}(T)}.$$

Thus it suffices to exhibit a bijection between square tableaux and words in $R(w_0)$ that preserves major index. In fact, the composition $\Psi := \Phi H$ has a stronger feature.

Define the *cyclic descent set* of a word $w = w_1 \cdots w_l$ to be the set

$$D(w) = \{i : w_i > w_{i+1}\} \pmod{l}$$

That is, we have descents in the usual way, but also a descent in position 0 if $w_l > w_1$. Then $\mathrm{maj}(w) = \sum_{i \in D(w)} i$. For example with $w = 132132132$, $D(w) = \{0, 2, 3, 5, 6, 8\}$ and $\mathrm{maj}(w) = 0+2+3+5+6+8 = 24$.

Similarly, we follow [Rh] in defining the cyclic descent set of a square (in general, rectangular) Young tableau. For T in $SYT(n^n)$, define $D(T)$ to be the set of all i such that i appears in a row above $i+1$, along with 0 if n^2-1 is above n^2 in $p(T)$. Major index is $\mathrm{maj}(T)=\sum_{i\in D(T)} i$. We will see that Ψ preserves cyclic descent sets, and hence, major index. Using our earlier example of $w=132132132$, one can check that

$$T=\Psi^{-1}(w)=\begin{array}{|c|c|c|}\hline 1&2&5\\\hline 3&6&8\\\hline 4&7&9\\\hline\end{array}$$

has $D(T)=D(w)$, and so $\mathrm{maj}(T)=\mathrm{maj}(w)$.

Lemma 14 *Let $T\in SYT(n^n)$, and let $w=\Psi(T)$ in $R(w_0)$. Then $D(T)=D(w)$.*

Proof: First, we observe that both types of descent sets shift cyclically under their respective actions:

$$D(p(T))=\{i-1 \pmod{n^2} : i\in D(T)\},$$

and

$$D(c(w))=\{i-1 \pmod{n^2} : i\in D(w)\}.$$

For words under cyclic rotation, this is obvious. For tableaux under promotion, this is a lemma of Rhoades [Rh, Lemma 3.3].

Because of this cyclic shifting, we see that $i\in D(T)$ if and only if $0\in D(p^i(T))$. Thus, it suffices to show that $0\in D(T)$ if and only if $0\in D(w)$. (Actually, it is easier to determine if n^2-1 is a descent.)

Let $S=\Phi^{-1}(w)$ be the shifted doubled staircase tableau corresponding to w. We have $n^2-1\in D(w)$ if and only if n^2 is in a higher row in $p^{-1}(S)$ than in S. But since n^2 occupies the same place in $p^{-1}(S)$ as n^2-1 occupies in S, this is to say n^2-1 is above n^2 in S. On the other hand, $n^2-1\in D(T)$ if and only if n^2-1 is above n^2 in T. It is straightforward to check that since S is obtained from T by jeu de taquin into the upper corner, the relative heights of n^2 and n^2-1 (i.e., whether n^2 is below or not) are the same in S as in T. This completes the proof. □

This lemma yields the desired result for $X(q)$.

Theorem 15 *The q-analogue of the hook length formula for an $n\times n$ square Young diagram is, up to a shift, the major index generating function for reduced expressions of the longest element in the hyperoctahedral group:*

$$\sum_{w\in R(w_0)} q^{\mathrm{maj}(w)}=q^{n\binom{n}{2}}\cdot\frac{[n^2]_q!}{[n]_q^n\prod_{i=1}^{n-1}([i]_q\cdot[2n-i]_q)^i}.$$

Theorem 15, along with Corollary 13, completes the proof of our main result, Theorem 1. Because this result can be stated purely in terms of the set $R(w_0)$ and a natural statistic on this set, it would be interesting to obtain a self-contained proof, i.e., one that does not appeal to Haiman's or Rhoades' work. Why must a result about cyclic rotation of words rely on promotion of Young tableaux?

Acknowledgements

We thank Brendon Rhoades for encouraging us to write this note. Thanks also to Kevin Dilks, John Stembridge, and Alex Yong for fruitful discussions on this and related topics, and to Sergey Fomin for comments on the manuscript.

References

[EG] P. EDELMAN AND C. GREENE: Balanced tableaux, ADV. IN MATH. **63** (1987), 42–99.

[H1] M. HAIMAN: On mixed insertion, symmetry, and shifted Young tableaux, *J. Combin. Theory Ser. A* **50** (1989), no. 2, 196–225.

[H2] M. HAIMAN: Dual equivalence with applications, including a conjecture of Proctor, *Discrete Math. 99* (1992), 79–113.

[LLT] A. LASCOUX, B. LECLERC, AND J. -Y. THIBON: The plactic monoid, *in* "M. Lothaire, Algebraic combinatorics on words", Cambridge University Press, Cambridge, 2002 (Chapter 6).

[RSW] V. REINER, D. STANTON, AND D. WHITE: The cyclic sieving phenomenon, *J. Combin. Theory Ser. A* **108** (2004), no. 1, 17–50.

[Rh] B. RHOADES: Cyclic sieving, promotion, and representation theory, *J. Combin. Theory Ser. A* **117** (2010), no. 1, 38–76.

[Sa] B. SAGAN: Shifted tableaux, Schur Q-functions, and a conjecture of R. P. Stanley, *J. Combin. Theory Ser. A* **45** (1987), 62–103.

[Ser] L. SERRANO: The shifted plactic monoid, arXiv: 0811.2057.

[Sch] M. P. SCHÜTZENBERGER: Promotion des morphismes d'ensembles ordonnés, *Discrete Mathematics* **2**, (1972), 73–94.

[Sta] R. STANLEY: *Enumerative Combinatorics Vol 2*, Cambridge University Press, Cambridge, UK, 1999.

[W] D. R. WORLEY: A theory of shifted Young tableaux, Ph.D. thesis, MIT, 1984; available at `http://hdl.handle.net/1721.1/15599`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 862–873

The cluster and dual canonical bases of $\mathbb{Z}[x_{11}, \ldots, x_{33}]$ are equal

Brendon Rhoades[1]

[1]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA, 02139*

Abstract. The polynomial ring $\mathbb{Z}[x_{11}, \ldots, x_{33}]$ has a basis called the dual canonical basis whose quantization facilitates the study of representations of the quantum group $U_q(\mathfrak{sl}_3(\mathbb{C}))$. On the other hand, $\mathbb{Z}[x_{11}, \ldots, x_{33}]$ inherits a basis from the cluster monomial basis of a geometric model of the type D_4 cluster algebra. We prove that these two bases are equal. This extends work of Skandera and proves a conjecture of Fomin and Zelevinsky. This also provides an explicit factorization of the dual canonical basis elements of $\mathbb{Z}[x_{11}, \ldots, x_{33}]$ into irreducible polynomials.

Résumé. L'anneau de polynômes $\mathbb{Z}[x_{11}, \ldots, x_{33}]$ a une base appelée base duale canonique, et dont une quantification facilite l'étude des représentations du groupe quantique $U_q(\mathfrak{sl}_3(\mathbb{C}))$. D'autre part, $\mathbb{Z}[x_{11}, \ldots, x_{33}]$ admet une base issue de la base des monômes d'amas de l'algèbre amassée géométrique de type D_4. Nous montrons que ces deux bases sont égales. Ceci prolonge les travaux de Skandera et démontre une conjecture de Fomin et Zelevinsky. Ceci fournit également une factorisation explicite en polynômes irréductibles des éléments de la base duale canonique de $\mathbb{Z}[x_{11}, \ldots, x_{33}]$.

Keywords: cluster algebra, dual canonical basis

---

1 Introduction

For $n \geq 0$, let $\mathcal{A}_n$ denote the polynomial ring $\mathbb{Z}[x_{11}, \ldots, x_{nn}]$ in the n^2 commuting variables $(x_{ij})_{1 \leq i,j \leq n}$. The algebra $\mathcal{A}_n$ has an obvious $\mathbb{Z}$-basis of monomials in the variables x_{ij}, which we call the *natural* basis. In addition to the natural basis, the ring $\mathcal{A}_n$ has many other interesting bases such as a bitableau basis defined by Mead and popularized by Désarménien, Kung, and Rota [2] having applications in invariant theory and the dual canonical basis of Lusztig [8] and Kashiwara [5] whose quantization facilitates the study of representations of the quantum group $U_q(\mathfrak{sl}_n(\mathbb{C}))$. Given two bases of $\mathcal{A}_n$, it is natural to compare them by examining the corresponding transition matrix. For example, in [9] it is shown that these latter two bases are related via a transition matrix which may be taken to be unitriangular (i.e., upper triangular with 1's on the main diagonal) with respect to an appropriate ordering of basis elements.

Cluster algebras are a certain class of commutative rings introduced by Fomin and Zelevinsky [3] to study total positivity and dual canonical bases. Any cluster algebra comes equipped with a distinguished set of generators called *cluster variables* which are grouped into finite overlapping subsets called *clusters*, all of which have the same cardinality. The cluster algebras with a finite number of clusters have a classification similar to the Cartan-Killing classification of finite-dimensional simple complex Lie algebras [4]. In this classification, it turns out that the cluster algebra of type D_4 is a localization of the ring $\mathcal{A}_3$ (see for example [10]) and the ring $\mathcal{A}_3$ inherits a $\mathbb{Z}$-basis consisting of cluster monomials. We call this basis the *cluster basis*. Fomin and Zelevinsky conjectured that the cluster basis and the dual canonical basis of $\mathcal{A}_3$ are equal, and Skandera showed that any two of the natural, cluster, and dual canonical bases of $\mathcal{A}_3$ are related via a unitriangular transition matrix when basis elements are ordered appropriately [10]. In this paper we strengthen Skandera's result and prove Fomin and Zelevinsky's conjecture with the following result (definitions will be postponed until Section 2).

Theorem 1.1 *The dual canonical and cluster basis of $\mathcal{A}_3$ are equal.*

Since each of the cluster and frozen variables of $\mathcal{A}_3$ are irreducible polynomials, this result can be viewed as giving a complete factorization of the dual canonical basis elements of $\mathcal{A}_3$ into irreducibles. Because the natural $GL_3(\mathbb{C})$ action on $\mathbb{C} \otimes_{\mathbb{Z}} \mathcal{A}_3$ is multiplicative, this could aid in constructing representing matrices for this action with respect to the dual canonical basis.

Theorem 1.1 will turn out to be the classical $q = 1$ specialization of a result (Theorem 3.6) comparing two bases of a noncommutative quantization $\mathcal{A}_3^{(q)}$ of the polynomial ring $\mathbb{A}_3$. In Section 2 we define the cluster basis of the classical ring $\mathcal{A}_3$. In Section 3 we introduce the quantum polynomial ring $\mathcal{A}_n^{(q)}$ together with its dual canonical basis and a quantum analogue of the cluster basis of $\mathcal{A}_3$. In Section 4 we comment on possible extensions of the results in this extended abstract.

2 The Cluster Basis of $\mathbb{Z}[x_{11}, \ldots, x_{33}]$

We shall not find it necessary to use a great deal of the general theory of cluster algebras to define and study the cluster basis of $\mathcal{A}_3$. Rather, we simply will define a collection of 16 polynomials in $\mathcal{A}_3$ to be cluster variables and associate to each of them a certain decorated octogon, define an additional 5 polynomials to be frozen variables, define (extended) clusters in terms of noncrossing conditions on decorated octogons, and define cluster monomials to be products of elements of an extended cluster.

For any two subsets $I, J \subseteq 3$ of equal size, define the $(I, J)-$minor $\Delta_{I,J}(x)$ of $x = (x_{i,j})_{1\leq i,j\leq 3}$ to be the determinant of the submatrix of x with row set I and column set J. Define additionally two more polynomials, the 132- and 213-Kazhdan-Lusztig immanants of x, by

$$\mathrm{Imm}_{132}(x) = x_{11}x_{23}x_{32} - x_{12}x_{23}x_{31} - x_{13}x_{21}x_{32} + x_{13}x_{22}x_{31}$$

and

$$\mathrm{Imm}_{213}(x) = x_{12}x_{21}x_{33} - x_{12}x_{23}x_{31} - x_{13}x_{21}x_{32} + x_{13}x_{22}x_{31}.$$

The *cluster variables* are the 16 elements of $\mathcal{A}_3$ shown in Figure 2.1 [10, p. 3], with the associated decorated octogons. Every octogon is decorated with either a pair of parallel nonintersecting nondiameters or a diameter colored one of two colors, red or blue.

A *centrally symmetric modified triangulation* of the octogon is a maximal collection of the above octogon decorations without crossings, where we adopt the convention that identical diameters of different colors do not cross and distinct diameters of the same color do not cross. Every centrally symmetric modified triangulation of the octogon consists of four decorations, and a *cluster* is the associated four element set of polynomials corresponding to the decorations in such a triangulation. There are 50 centrally symmetric modified triangulations of the octogon, and hence 50 clusters. Four examples of centrally symmetric modified triangulations are shown in Figure 2.2. The corresponding clusters are, from left to

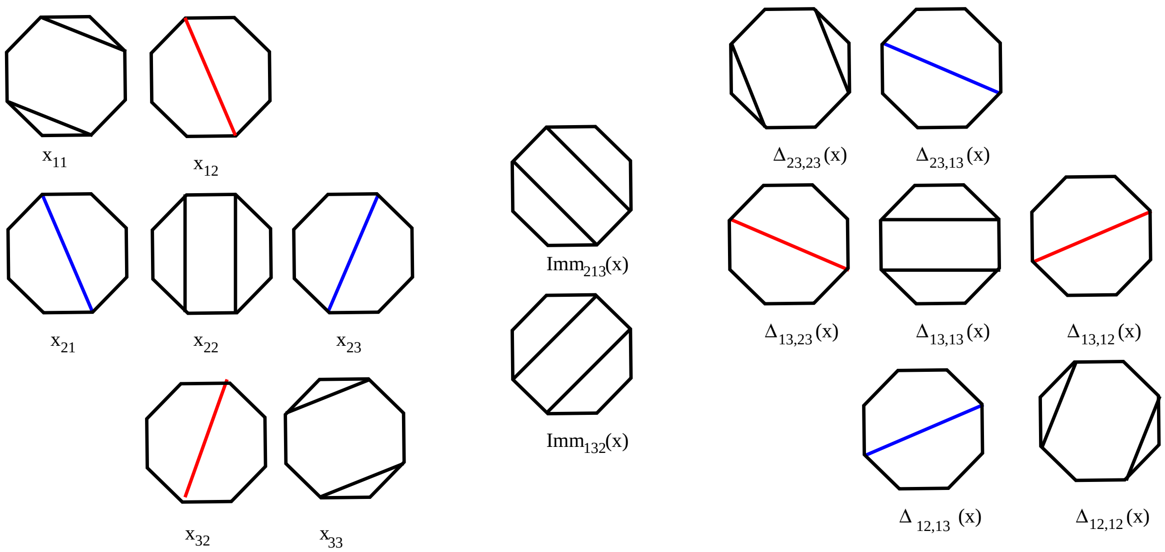


Fig. 2.1: Cluster variables in $\mathbb{Z}[x_{11}, \ldots, x_{33}]$

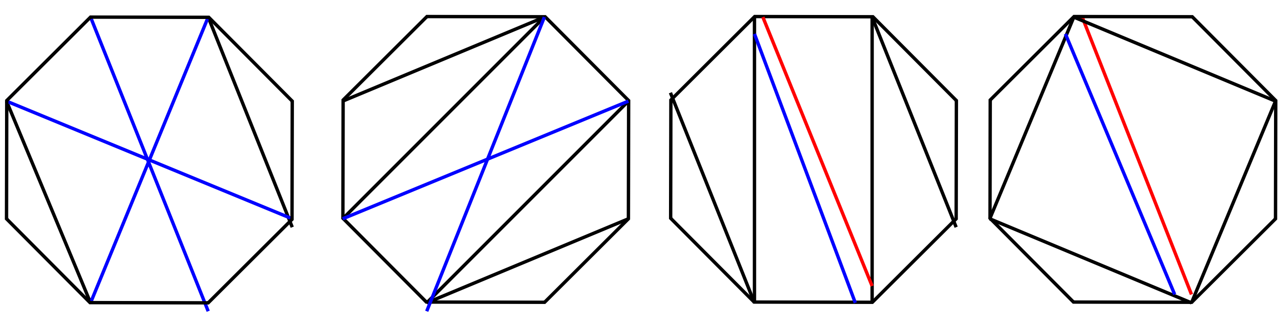


Fig. 2.2: Four centrally symmetric modified triangulations corresponding to clusters

right, $\{x_{21}, x_{23}, \Delta_{23,13}(x), \Delta_{23,23}(x)\}$, $\{x_{23}, x_{33}, \Delta_{12,13}(x), \mathrm{Imm}_{132}(x)\}$, $\{x_{12}, x_{21}, x_{22}, \Delta_{23,23}(x)\}$, and $\{x_{11}, x_{12}, x_{21}, \Delta_{12,12}(x)\}$.

We define additionally a set $\mathcal{F}$ consisting of the five polynomials

$$\mathcal{F} := \{x_{13}, \Delta_{12,23}(x), \Delta_{123,123}(x) = \det(x), \Delta_{23,12}(x), x_{31}\}.$$

Elements in $\mathcal{F}$ are called *frozen variables* and the union of $\mathcal{F}$ with any cluster is an *extended cluster*. A *cluster monomial* is a product of the form $z_1^{a_1} \cdots z_9^{a_9}$, where $\{z_1, \ldots, z_9\}$ is an extended cluster and the a_i are nonnegative integers. Observe that the same cluster monomial can arise from different extended clusters. The *cluster basis* of $\mathcal{A}_3$ is the set of all possible cluster monomials.

Skandera [10] develops a bijection ϕ between the cluster basis and the set $\mathrm{Mat}_3(\mathbb{N})$ as follows. For any cluster or frozen variable z, let $\phi(z)$ be the lexicographically greatest matrix $A = (a_{ij})$ for which the monomial $\prod x_{ij}^{a_{ij}}$ appears with nonzero coefficient in the expansion of z in the natural basis. Given an arbitrary cluster monomial $z_1^{a_1} \cdots z_9^{a_9}$, extend the definition of ϕ via

$$\phi(z_1^{a_1} \cdots z_9^{a_9}) := a_1\phi(z_1) + \cdots + a_9\phi(z_9).$$

The fact that ϕ is a bijection [10] implies that the set of cluster monomials is related to the natural basis of $\mathcal{A}_3$ via a unitriangular, integer transition matrix, and thus is actually a $\mathbb{Z}$-basis for $\mathcal{A}_3$ (the fact that the cluster monomials form a basis for $\mathcal{A}_3$ is also a consequence of more theory of finite type cluster algebras).

Example 2.1 *Consider the cluster corresponding to the leftmost centrally symmetric modified triangulation in Figure 2.2, i.e.* $\{x_{21}, x_{23}, \Delta_{23,13}(x), \Delta_{23,23}(x)\}$. *An example of a cluster monomial drawn from the corresponding extended cluster is*

$$z := x_{21}^7 x_{23}^0 \Delta_{23,13}(x)^2 \Delta_{23,23}(x)^1 x_{13}^0 \Delta_{12,23}(x)^2 \Delta_{123,123}(x)^0 \Delta_{23,12}(x)^0 x_{31}^7.$$

We have that

$$\phi(z) = 7\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 0\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + 2\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \cdots = \begin{pmatrix} 0 & 2 & 0 \\ 9 & 1 & 2 \\ 7 & 0 & 3 \end{pmatrix}.$$

3 The Quantum Polynomial Ring

For $n \geq 0$, define the *quantum polynomial ring* $\mathcal{A}_n^{(q)}$ to be the unital associative (noncommutative) $\mathbb{Z}[q^{\pm 1/2}]$-algebra generated by the n^2 variables $x = (x_{ij})_{1 \leq i,j \leq n}$ and subject to the relations

$$x_{ik}x_{il} = qx_{il}x_{ik} \tag{1}$$
$$x_{ik}x_{jk} = qx_{jk}x_{ik} \tag{2}$$
$$x_{il}x_{jk} = x_{jk}x_{il} \tag{3}$$
$$x_{ik}x_{jl} = x_{jl}x_{ik} + (q - q^{-1})x_{il}x_{jk}, \tag{4}$$

where $i < k$ and $k < l$. It follows from these relations that the specialization of $\mathcal{A}_n^{(q)}$ to $q = 1$ recovers the classical polynomial ring $\mathcal{A}_n$. The center of $\mathcal{A}_n^{(q)}$ is generated by the *quantum determinant* $\det_q(x) :=$

$\sum_{w\in S_n}(-q)^{\ell(w)}x_{1,w(1)}\cdots x_{n,w(n)}$. Factoring the extension $\mathbb{C}\otimes_{\mathbb{Z}}\mathcal{A}_n^{(q)}$ by the ideal $(\det_q(x)-1)$ yields the quantum coordinate ring $\mathcal{O}_q(SL_n(\mathbb{C}))$ of the special linear group. Given two ring elements $f,g\in\mathcal{A}_n^{(q)}$, we say that f is a *q-shift* of g if there is a number a so that $f=q^a g$.

The natural basis of $\mathcal{A}_n$ lifts to a $\mathbb{Z}[q^{\pm 1/2}]$-basis of the quantum polynomial ring $\mathcal{A}_n^{(q)}$ given by $\{X^A := x_{11}^{a_{11}}\cdots x_{nn}^{a_{nn}} \mid A=(a_{ij})\in \mathrm{Mat}_n(\mathbb{N})\}$, where the terms in the product are in lexicographical order (see, for example, [12]). We call this basis the *quantum natural basis* (QNB). We will find it convenient to work with a $\mathbb{Z}[q^{\pm 1/2}]$-basis for $\mathcal{A}_n^{(q)}$ whose elements are q-shifts of QNB elements. Following [12], for any matrix $A=(a_{ij})\in\mathrm{Mat}_n(\mathbb{N})$, define the number $e(A) := -\frac{1}{2}\sum_i\sum_{j<k}(a_{ij}a_{ik}+a_{ji}a_{ki})$ and the quantum polynomial $X(A) := q^{e(A)}X^A\in\mathcal{A}_n^{(q)}$. The set $\{X(A)\,|\,A\in\mathrm{Mat}_n(\mathbb{N})\}$ is also a $\mathbb{Z}[q^{\pm 1/2}]$-basis of $\mathcal{A}_n^{(q)}$, called the *modified quantum natrual basis* (MQNB).

As with the classical polynomial ring $\mathcal{A}_n$, the quantum ring $\mathcal{A}_n^{(q)}$ admits a natural $\mathbb{N}$-grading by degree. Finer than this grading is an $\mathbb{N}^n\times\mathbb{N}^n$-grading, where the $(r_1,\ldots,r_n)\times(c_1,\ldots,c_n)$-graded piece is the $\mathbb{Z}[q^{\pm 1/2}]$-linear span of all MQNB elements $X(A)$ for matrices $A\in\mathrm{Mat}_n(\mathbb{N})$ with row vector $row(A)=(r_1,\ldots,r_n)$ and column vector $col(A)=(c_1,\ldots,c_n)$. It is routine to check from (1)-(4) that this grading is well-defined.

The ring $\mathcal{A}_n^{(q)}$ is equipped with an involutive bar antiautomorphism defined by the $\mathbb{Z}$-linear extension of $\overline{q^{1/2}}=q^{-1/2}$ and $\overline{x_{ij}}=x_{ij}$. It follows readily from relations (1)-(4) that $\bar{\cdot}$ is well-defined. Observe that the bar involution specializes to the identity map at $q=1$. The *dual canonical basis* (DCB) of $\mathcal{A}_n^{(q)}$ arises naturally when attempting to find bases of $\mathcal{A}_n^{(q)}$ consisting of bar invariant polynomials.

Define a partial order $\leq_{Br}$ on $\mathrm{Mat}_n(\mathbb{N})$ called *Bruhat order* by letting $\leq_{Br}$ be the transitive closure of $A\prec_{Br}B$ if B can be obtained from $A=(a_{ij})_{1\leq i,j\leq n}$ by a 2×2 submatrix transformation of the form

$$\begin{pmatrix} a_{ik} & a_{il} \\ a_{jk} & a_{jl}\end{pmatrix}\mapsto\begin{pmatrix} a_{ik}-1 & a_{il}+1 \\ a_{jk}+1 & a_{jl}-1\end{pmatrix},$$

for $i<j$ and $k<l$ with $a_{ik},a_{ji}>0$. Observe that the restriction of $\leq_{Br}$ to the set of permutation matrices is isomorphic to the ordinary (strong) Bruhat order on the symmetric group S_n. Observe also that matrix transposition and antitransposition are automorphisms of the poset $(\mathrm{Mat}_n(\mathbb{N}),\leq_{Br})$.

Theorem 3.1 *There exists a unique $\mathbb{Z}[q^{\pm 1/2}]$-basis*

$$\{b(A)\,|\,A\in\mathrm{Mat}_n(\mathbb{N})\}$$

of $\mathcal{A}_n^{(q)}$ where $b(A)$ is homogeneous with respect to the $\mathbb{N}^n\times\mathbb{N}^n$-grading of $\mathcal{A}_n^{(q)}$ with degree $row(A)\times col(A)$ and the b(A) satisfy
(1) (Bar invariance) $\overline{b(A)}=b(A)$ for all $A\in\mathrm{Mat}_n(\mathbb{N})$, and
(2) (Triangularity) For all $A\in\mathrm{Mat}_n(\mathbb{N})$, the basis element $b(A)$ expands in the MQNB as

$$b(A)=X(A)+\sum_{B>_{Br}A}\beta_{A,B}(q^{1/2})X(B),$$

where the $\beta_{A,B}$ are polynomials in $q^{1/2}\mathbb{Z}[q^{1/2}]$.

This basis is called the dual canonical basis.

Proof: If one replaces Bruhat order with the dual of lexicographical order on $\mathrm{Mat}_n(\mathbb{N})$, this result is [12, Theorem 3.2]. However, as noted in the first paragraph of the proof of [12, Corollary 3.4], one has that the coefficient $\beta_{A,B}(q^{1/2})$ in the expansion

$$b(A) = X(A) + \sum_{B <_{lex} A} \beta_{A,B}(q^{1/2}) X(B),$$

is nonzero only if $A >_{Br} B$. □

While the DCB is important in the study of the representation theory of the quantum group $U_q(\mathfrak{sl}_n(\mathbb{C}))$[5] [8] , the lack of an elementary formula for the expansion of the $b(A)$ in the MQNB can make computations involving the DCB difficult. Setting $q = 1$ in the DCB elements yields a basis of the classical polynomial ring $\mathcal{A}_n$, also called the dual canonical basis.

Due to the triangularity condition (2) of Theorem 3.1, we will frequently need to analyze expansions of quantum ring elements in the (M)QNB. To find these expansions, we use relations (1)-(4) to put arbitrary ring elements in lexicographical order. While the somewhat exotic relation (4) can make for complicated expansions, if we are only interested in the Bruhat leading term the situation improves. Given any product $m \in \mathcal{A}_n^{(q)}$ of the generators x_{ij} and a ground ring element $\beta(q) \in \mathbb{Z}[q^{\pm 1/2}]$, define the *content* $C(\beta(q)m)$ of $\beta(q)m$ to be the $n \times n$ matrix whose (i, j)-entry is equal to the number of occurances of x_{ij} in m.

Lemma 3.2 *(Leading Lemma) Let $f = m + m_1 + \cdots + m_r \in \mathcal{A}_n^{(q)}$ be an element of $\mathcal{A}_n^{(q)}$ such that the m_k are monomials and $C(m) <_{Br} C(m_k)$ for all k. Write $m = \beta(q)x_{i_1 j_1} \cdots x_{i_r j_r}$, where $\beta(q) \in \mathbb{Z}[q^{\pm 1/2}]$.*

The expansion of f in the QNB has unique term $X^{C(m)}$ with $C(m)$ Bruhat minimal among the contents of the terms in the QNB expansion of f and the coefficient of $X^{C(m)}$ in this expansion is

$$q^{-y}\beta(q),$$

where y is given by

$$y = |\{(k < \ell) \,|\, i_k = i_\ell, j_k > j_\ell\}| + |\{(k < \ell) \,|\, j_k = j_\ell, i_k > i_\ell\}|.$$

Proof: Observe that the application of relations (1)-(3) to the ring element f do not change the contents of the monomial constituents of f. Moreover, the application of relation (4) to any monomial m_0 in $\mathcal{A}_n^{(q)}$ yields a sum $m_0' + m_0''$, where m_0' has the same content and coefficient as m_0 and m_0'' has content which is greater in Bruhat order than the content of m_0. The value of y follows from the exponents of q which appear in the quasicommutativity relations (1)-(3). □

In the classical setting $q = 1$, Skandera [11] discovered an explicit formula for dual canonical basis elements of $\mathcal{A}_n$ which involves certain polynomials called *immanants*. Given a permutation $w \in S_m$ and an $m \times m$ matrix $y = (y_{ij})_{1 \le i,j \le m}$ with entries drawn from the set $\{x_{ij} \,|\, 1 \le i, j \le n\}$, define the *w-KL immanant* of y to be

$$\mathrm{Imm}_w(y) := \sum_{v \in S_m} Q_{v,w}(1) y_{1,v(1)} \cdots y_{m,v(m)}.$$

Here $Q_{v,w}(q)$ is the *inverse Kazhdan-Lusztig polynomial* corresponding to the permutations v and w (see [6] or [1]). It can be shown that the KL immanant $\mathrm{Imm}_1(y)$ corresponding to the identity permutation $1 \in S_m$ is equal to the determinant $\det(y)$.

Any (weak) composition $\alpha \models m$ with n parts induces a function $[m] \to [n]$, also denoted α, which maps the interval $(\alpha_1 + \cdots + \alpha_{i-1}, \alpha_1 + \cdots \alpha_i]$ onto i for all i. We also have the associated parabolic subgroup $S_\alpha \cong S_{\alpha_1} \times \cdots \times S_{\alpha_n}$ of S_m which stabilizes all of the above intervals. Given a pair $\alpha, \beta \models m$ of compositions of m both having n parts, we define the *generalized submatrix* $x_{\alpha,\beta}$ of x to be the $m \times m$ matrix satisfying $(x_{\alpha,\beta})_{ij} := x_{\alpha(i),\beta(j)}$ for all $1 \leq i, j \leq m$. Let $\Lambda_m(\alpha, \beta)$ denote the set of Bruhat maximal permutations in the set of double cosets $S_\alpha \backslash S_m / S_\beta$. Skandera's work [11, Section 2] implies that the dual canonical basis of $\mathcal{A}_n$ is equal to the set

$$\bigcup_{m \geq 0} \bigcup_{\alpha,\beta} \{\mathrm{Imm}_w(x_{\alpha,\beta}) \,|\, w \in \Lambda_m(\alpha, \beta)\}. \tag{5}$$

The lack of an elementary description of the inverse KL polynomials is the most difficult part in using Skandera's formula to write down DCB elements.

Returning to the quantum setting, for two subsets $I, J \subseteq [n]$ with $|I| = |J|$, the *quantum minor* $\Delta^{(q)}_{I,J}(x) \in \mathcal{A}^{(q)}_n$ is the quantum determinant of the submatrix of x with row set I and column set J. Restricting to the case $n = 3$, we define the *quantum 132- and 213-KL immanants*, denoted $\mathrm{Imm}^{(q)}_{132}(x)$ and $\mathrm{Imm}^{(q)}_{213}(x)$, to be the elements of $\mathcal{A}^{(q)}_3$ given by

$$\mathrm{Imm}^{(q)}_{132}(x) = x_{11}x_{23}x_{32} - qx_{12}x_{23}x_{31} - qx_{13}x_{21}x_{32} + q^2 x_{13}x_{22}x_{31}$$

and

$$\mathrm{Imm}^{(q)}_{213}(x) = x_{12}x_{21}x_{33} - qx_{12}x_{23}x_{31} - qx_{13}x_{21}x_{32} + q^2 x_{13}x_{22}x_{31}.$$

Define *quantum cluster* and *quantum frozen variables* to be the polynomials obtained by replacing every minor in the classical quantum or frozen variable definition by its corresponding quantum minor and the classical polynomials $\mathrm{Imm}_{132}(x)$ and $\mathrm{Imm}_{213}(x)$ by their quantum counterparts. Define a quantum (extended) cluster to be the set of quantum (frozen and) cluster variables corresponding to polynomials in a classical (extended) cluster.

To define the quantum cluster monomials, fix a total order $\{z'_1 < z'_2 < \cdots < z'_{21}\}$ on the union of the quantum cluster and frozen variables. A *quantum cluster monomial* is any product of the form $z_1^{a_1} \cdots z_9^{a_9} \in \mathcal{A}^{(q)}_3$, where $\{z_1 < \cdots < z_9\}$ is an ordered quantum extended cluster and the a_i are nonnegative integers. Skandera's map ϕ yields a bijection (also denoted ϕ) between the set of quantum cluster monomials and $\mathrm{Mat}_3(\mathbb{N})$. This bijection, combined with the Leading Lemma, implies that the transition matrix between the set of quantum cluster monomials and the QNB is triangular with units on the diagonal with respect to any order of basis elements which is obtained from a linear extension of Bruhat order. Therefore, the set of all quantum cluster monomials is a $\mathbb{Z}[q^{\pm 1/2}]$-basis for $\mathcal{A}^{(q)}_3$, called the *quantum cluster basis* (QCB). Observe also that every QCB element is homogeneous with respect to the $\mathbb{N}^3 \times \mathbb{N}^3$-grading of the ring $\mathcal{A}^{(q)}_3$.

It is natural to ask how much the QCB depends on the initial choice of total order $\{z'_1 < z'_2 < \cdots < z'_{21}\}$ on the union of the quantum cluster and frozen variables. While it is the case that different choices of the total order $\{z'_1 < z'_2 < \cdots < z'_{21}\}$ can lead to different quantum cluster monomials, we will show in Observation 3.5 that these ring elements differ only up to a q-shift (which may depend not only on the order chosen, but also on the quantum cluster monomial in question). Thus, the QCB is independent of this choice of order 'up to q-shift'.

Our computational work with the ring $\mathcal{A}_n^{(q)}$ will be economized by means of a collection of algebra maps. Define maps τ and α on the generators of $\mathcal{A}_n^{(q)}$ by the formulas $\tau(x_{ij}) = x_{ji}$ and $\alpha(x_{ij}) = x_{(n-j+1)(n-i+1)}$. It is routine to check from the relations (1)-(4) that τ extends to an involutive $\mathbb{Z}[q^{\pm 1/2}]$-algebra automorphism $\tau : \mathcal{A}_n^{(q)} \to \mathcal{A}_n^{(q)}$ and that α extends to an involutive $\mathbb{Z}[q^{\pm 1/2}]$-algebra antiautomorphism $\alpha : \mathcal{A}_n^{(q)} \to \mathcal{A}_n^{(q)}$. The maps τ and α will be called the *transposition* and *antitransposition* maps, respectively, because they act on the matrix $x = (x_{ij})$ of generators by transposition and antitransposition. In addition, for any two subsets $I, J \subseteq [n]$, we can form the subalgebra $\mathcal{A}_n^{(q)}(I,J)$ of $\mathcal{A}_n^{(q)}$ generated by $\{x_{ij} \mid i \in I, j \in J\}$. Writing $I = \{i_1 < \cdots < i_r\}$ and $J = \{j_1 < \cdots < j_s\}$, we have a $\mathbb{Z}[q^{\pm 1/2}]$-algebra isomorphism $c_{I,J} : \mathcal{A}_n^{(q)}(I,J) \to \mathcal{A}_n^{(q)}([r],[s])$ given by $c_{I,J} : x_{i_a,j_b} \mapsto x_{a,b}$. The map $c_{I,J}$ will be called the *compression map* corresponding to I and J because it acts on the matrix x of generators by compression into the northwest corner. The set of quantum cluster and frozen variables is closed under taking images under τ, α, and the compression maps.

Observation 3.3 *Let z be a quantum cluster or frozen variable. Then, $\alpha(z)$ and $\tau(z)$ are quantum cluster or frozen variables with $\phi(\alpha(z)) = \phi(z)^{T'}$ and $\phi(\tau(z)) = \phi(z)^T$, where $\cdot^T$ denotes matrix transposition and $\cdot^{T'}$ denotes matrix antitransposition. Moreover, if the row support of $\phi(z)$ is contained in $I \subseteq [3]$ and the column support of $\phi(z)$ is contained in $J \subseteq [3]$, then the image of z under the compression map $c_{I,J}$ corresponding to I and J is a quantum cluster or frozen variable whose image under ϕ is obtained by compressing the nonzero rows and columns of $\phi(z)$ to the northwest.*

The proof of this observation is a direct computation. For example, the quantum cluster variable $z = \Delta^{(q)}_{23,13}(x)$ satisfies $\phi(z) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, which has row support $\{2,3\}$ and column support $\{1,3\}$. The image of z under the compression map $c_{23,13}$ is $y = \Delta^{(q)}_{12,12}(x)$, which satisfies $\phi(y) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. The 7 equivalence classes of quantum cluster variables under the maps τ and α are shown in Figure 3.1. The actions of τ, α, and the compression maps remain well-defined on the level of quantum clusters. The transposition, antitransposition, and compression maps act on MQNB elements in a simple way.

Observation 3.4 *Let $A \in \mathrm{Mat}_n(\mathbb{N})$. We have the following formulae involving the MQNB, retaining notation from Observation 3.3:*
(1) $\tau(X(A)) = X(A^T)$
(2) $\alpha(X(A)) = X(A^{T'})$.
Moreover, if the row support of A is contained in I and the column support of A is contained in J for subsets $I, J \subseteq [n]$, then
(3) $c_{I,J}(X(A)) = X(C_{I,J}(A))$,
where $C_{I,J}(A)$ is obtained by compressing the entries in A with row indicies in I and column indicies in J to the northwest.

Proof: (3) is trivial. To verify (1) and (2), one applies the maps τ and α to $X(A)$ and uses the defining relations (3.1)-(3.3) to get the desired result. □

It is natural to ask to what extent the choice of total order $\{z'_1 < \cdots < z'_{21}\}$ on the union of the quantum cluster and frozen variables affects the quantum cluster monomials. The effects of this choice turn out to

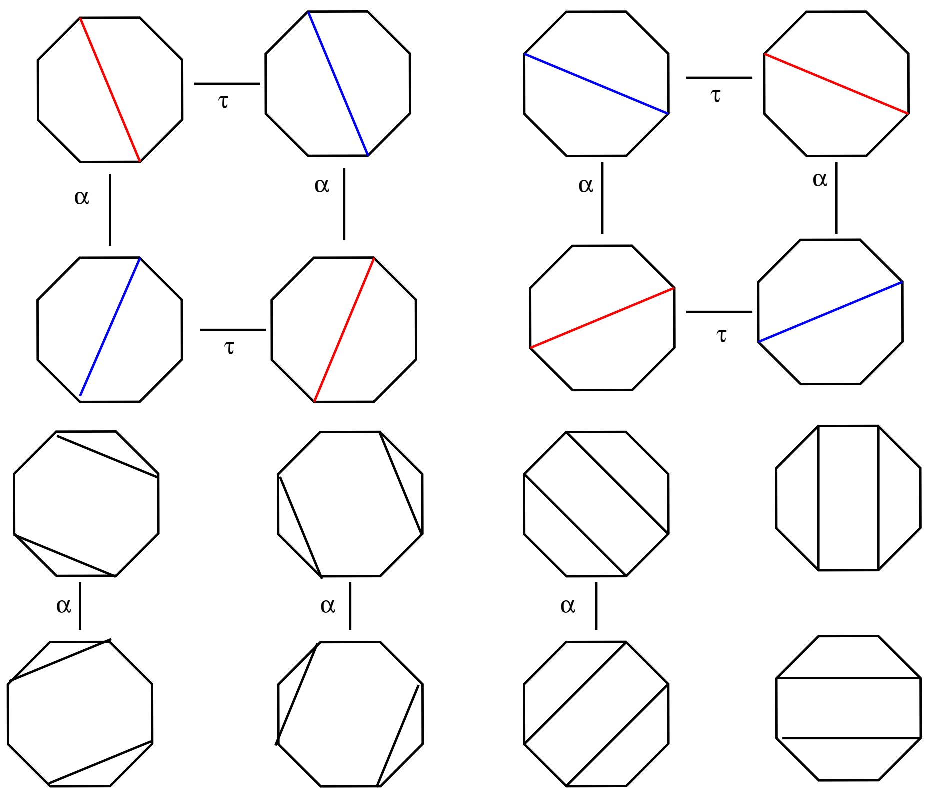


Fig. 3.1: Equivalence classes of cluster variables under α and τ

be quite benign. More precisely, we observe that different choices of total orders only affect the QCB by q-shifts. Two ring elements $f, g \in \mathcal{A}_n^{(q)}$ are said to *quasicommute* if $fg = q^a gf$ for some a.

Observation 3.5 *Let z, z' be a pair of polynomials which appear in the same quantum extended cluster. Then, z and z' quasicommute and moreover $zz' = q^a z' z$ for some $a \in \mathbb{Z}$.*

The proof of the above observation is a straightforward, albeit tedious calculation using the relations (1)-(4). It follows from Lemma 5.1 of [12] that any quantum frozen variable quasicommutes with any quantum frozen or cluster variable, so it is enough to show that cluster variables appearing in the same (nonextended) cluster pairwise quasicommute. By use of the transposition map τ, the antitransposition map α, and the compression maps, we need only check this observation on 7 pairs of quantum cluster variables. The resulting identities are as follows.

$$\begin{array}{cc}
\Delta^{(q)}_{12,12}(x)\Delta^{(q)}_{13,23}(x) = q\Delta^{(q)}_{13,23}(x)\Delta^{(q)}_{12,12}(x) & \mathrm{Imm}^{(q)}_{132}(x)\Delta^{(q)}_{13,12}(x) = \Delta^{(q)}_{13,12}(x)\mathrm{Imm}^{(q)}_{132}(x) \\
\Delta^{(q)}_{12,12}(x)\mathrm{Imm}^{(q)}_{132}(x) = q^2\mathrm{Imm}^{(q)}_{132}(x)\Delta^{(q)}_{12,12}(x) & \Delta^{(q)}_{13,12}(x)\Delta^{(q)}_{12,13}(x) = \Delta^{(q)}_{12,13}(x)\Delta^{(q)}_{13,12}(x) \\
\Delta^{(q)}_{13,12}(x)\Delta^{(q)}_{13,23}(x) = q\Delta^{(q)}_{13,23}(x)\Delta^{(q)}_{13,12}(x) & x_{11}x_{12} = qx_{12}x_{11} \\
x_{12}x_{21} = x_{21}x_{12} &
\end{array}$$

Theorem 1.1 follows from specializing the following result at $q = 1$.

Theorem 3.6 *Every DCB element of $\mathcal{A}_3^{(q)}$ is a q-shift of a unique QCB element of $\mathcal{A}_3^{(q)}$.*

Proof: (Sketch) Fixed a quantum cluster monomial $z_1^{a_1} \cdots z_9^{a_9}$ arising from an ordered quantum extended cluster $\{z_1 < \cdots < z_9\}$. We show that $z_1^{a_1} \cdots z_9^{a_9}$ has a q-shift which satisfies the bar invariance condition (1) of Theorem 3.1 and that this same q-shift also satisfies the triangularity condition (2) of Theorem 3.1.

To check that a q-shift of $z_1^{a_1} \cdots z_9^{a_9}$ is invariant under the bar involution, one first shows by direct computation that every quantum frozen and cluster variable is invariant under the bar involution. Using the transposition map τ, the antitransposition map α, and the compression maps $c_{I,J}$ reduces our computations here to checking that the four elements $x_{11}, \Delta^{(q)}_{12,12}(x), \mathrm{Imm}^{(q)}_{132}(x), \Delta^{(q)}_{123,123}(x) \in \mathcal{A}_3^{(q)}$ are bar invariant. Therefore, $\overline{z_i} = z_i$ for $1 \le i \le 9$. By Observation 3.5, we have that for all $1 \le i < j \le 9$, there exists $b_{ij} \in \mathbb{Z}$ so that $z_i z_j = q^{b_{ij}} z_j z_i$. Therefore, we have

$$\begin{aligned}
\overline{z_1^{a_1} \cdots z_9^{a_9}} &= z_9^{a_9} \cdots z_1^{a_1} \\
&= q^c z_1^{a_1} \cdots z_9^{a_9},
\end{aligned}$$

where the exponent c is given by

$$c = -\sum_{1 \le i < j \le 9} a_i a_j b_{ij}.$$

It follows that the ring element $q^{-c/2} z_1^{a_1} \cdots z_9^{a_9}$ is invariant under the bar involution. We now must check that $q^{-c/2} z_1^{a_1} \cdots z_9^{a_9}$ also satisfies the triangularity condition of Theorem 3.1. This verification is omitted from this extended abstract. $\square$

Example 3.1 *Consider the quantum analogue*

$$z := x_{21}^7 \Delta_{23,13}^{(q)}(x)^2 \Delta_{23,23}^{(q)}(x)^1 \Delta_{12,23}^{(q)}(x)^2 x_{31}^7$$

of the cluster monomial of Example 2.1. By Theorem 3.6, the ring element z is a q-shift of a DCB element. Computing the $q = 1$ specialization of this DCB element using Skandera's characterization of the DCB of $\mathcal{A}_3$ would involve computing inverse Kazhdan-Lusztig polynomials corresponding to pairs of elements in the symmetric group on 24 letters.

4 Future Directions

In this paper, by means of a series of computations, we have proven that the dual canonical basis and the cluster monomial basis of the classical polynomial ring $\mathcal{A}_3$ are equal by showing that they have quantizations which differ by a q-shift. In doing so, we discovered how DCB elements for $\mathcal{A}_3$ and $\mathcal{A}_3^{(q)}$ decompose into irreducibles and found an easy way to write down any DCB element of these rings - up to a q-shift, just choose a decorated octogon and write down some monomial in the elements of the related extended (quantum) cluster. It is natural to ask how much of this can be extended to rings $\mathcal{A}_n$ and $\mathcal{A}_n^{(q)}$ for $n > 3$. It turns out that there are obstructions to finding such results from both the theory of cluster monomial bases and dual canonical bases.

For $n > 3$ there is a known cluster algebra structure on a subalgebra of $\mathcal{A}_n$ which gives rise to a linearly independent set of cluster monomials. Unfortunately, these cluster monomials do not span $\mathcal{A}_n$ for $n > 3$. Moreover, for $n > 3$ this cluster algebra is of infinite type, i.e., it has infinitely many clusters. Since these clusters are not given at the outset but rather are determined by a 'mutation' procedure starting with some initial cluster and 'mutation matrix' (see [3]), this would seem to make the cluster monomials in these algebras difficult to work with.

Leaving aside the present lack of a cluster algebra structure on $\mathcal{A}_n$, one can still ask how dual canonical basis elements of $\mathcal{A}_n$ and its quantization $\mathcal{A}_n^{(q)}$ factor. By Theorem 3.6 and the fact that quantum cluster monomials are arbitrary products of the ring elements in some quantum extended cluster, we have the following result in $\mathcal{A}_3^{(q)}$.

Corollary 4.1 *Let b be any element in the dual canonical basis of $\mathcal{A}_3^{(q)}$. Then, a q-shift of b^k is in the dual canonical basis of $\mathcal{A}_3^{(q)}$ for any $k \geq 0$.*

For n large, Leclerc [7] has shown that there exist elements b of the DCB of $\mathcal{A}_n^{(q)}$ such that b^2 is not a q-shift of a DCB element of $\mathcal{A}_n^{(q)}$ (so-called *imaginary* vectors). In light of the construction of cluster monomials, Leclerc's result is troubling if one wants to find a cluster-style interpretation of the factorization of all of the DCB elements of $\mathcal{A}_n^{(q)}$ for $n > 3$.

5 Acknowledgements

The author is grateful to Sergey Fomin and Mark Skandera for helpful conversations.

References

[1] A. BJÖRNER AND F. BRENTI. *Combinatorics of Coxeter groups*. Springer (1995).

[2] J. DÉSARMÉNIEN, J. P. S. KUNG, AND G.-C. ROTA. Invariant theory, Young bitableaux and combinatorics. *Advances in Math*, **27** (1978).

[3] S. FOMIN AND A. ZELEVINSKY. Cluster algebras I: Foundations. *J. Amer. Math. Soc.*, **15** (2002) pp. 497–529.

[4] S. FOMIN AND A. ZELEVINSKY. Cluster algebras II: finite type classification. *Invent. Math.*, **154** (2003) pp. 63–121.

[5] M. KASHIWARA. On crystal bases of the Q-analog of universal enveloping algebras. *Duke Math J.*, **63** (1991) pp. 465–516.

[6] D. KAZHDAN AND G. LUSZTIG. Representations of Coxeter groups and Hecke algebras. *Inv. Math.*, **53** (1979) pp. 165–184.

[7] B. LECLERC. Imaginary vectors in the dual canonical basis of $u_q(n)$. *Transform. Groups*, **8** (2003) pp. 95–104.

[8] G. LUSZTIG. Canonical bases arising from quantized enveloping algebras. *J. Amer. Math. Soc.*, **3** (1990) pp. 447–498.

[9] B. RHOADES AND M. SKANDERA. Bitableaux and zero sets of dual canonical basis elements (2009). Submitted. Available at http://www.math.berkeley.edu/ brhoades/research .

[10] M. SKANDERA. The Cluster Basis of $\mathbb{Z}[x_{1,1}, ..., x_{3,3}]$. *Electr. J. Combin.*, **14**, 1 (2007).

[11] M. SKANDERA. On the dual canonical and Kazhdan-Lusztig bases and 3412, 4231-avoiding permutations. *J. Pure Appl. Algebra*, **212** (2008) pp. 1086–1104.

[12] H. ZHANG. On dual canonical bases. *J. Phys. A: Math. Gen.*, **37** (2004) pp. 7879–7893.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 874–883

Crossings and nestings in set partitions of classical types

Martin Rubey[1] and Christian Stump[2]

[1]*Institut für Algebra, Zahlentheorie und Diskrete Mathematik, Leibniz Universität Hannover, Welfengarten 1, D-30167 Hannover, Germany*
[2]*LaCIM, Université du Québec à Montréal, 201 Président-Kennedy, Montréal (Québec) H2X 3Y7 , Canada*

Abstract. In this extended abstract, we investigate bijections on various classes of set partitions of classical types that preserve openers and closers. On the one hand we present bijections for types B and C that interchange crossings and nestings, which generalize a construction by Kasraoui and Zeng for type A. On the other hand we generalize a bijection to type B and C that interchanges the cardinality of a maximal crossing with the cardinality of a maximal nesting, as given by Chen, Deng, Du, Stanley and Yan for type A.

For type D, we were only able to construct a bijection between non-crossing and non-nesting set partitions. For all classical types we show that the set of openers and the set of closers determine a non-crossing or non-nesting set partition essentially uniquely.

Résumé. Dans ce résumé, nous étudions des bijections entre diverses classes de partitions d'ensemble de types classiques qui préservent les "openers" et les "closers". D'un part, nous présentons des bijections pour les types B et C qui échangent croisées et emboîtées, qui généralisent une construction de Kasraoui et Zeng pour le type A. D'autre part, nous généralisons une bijection pour le type B et C qui échange la cardinalité d'un croisement maximal avec la cardinalité d'un emboîtement maximal comme il a été fait par Chen, Deng, Du, Stanley et Yan pour le type A.

Pour le type D, nous avons seulement construit une bijection entre les partitions non croisées et non emboîtées. Pour tout les types classiques, nous montrons que l'ensemble des "openers" et l'ensemble des "closers" déterminent une partition non croisées ou non emboîtées essentiellement de façon unique.

Keywords: non-crossing partitions, non-nesting partitions, k-crossing partitions, k-nesting partitions, bijective combinatorics

Introduction

The lattice of non-crossing set partitions was first considered by Germain Kreweras in [18]. It was later reinterpreted by Paul Edelman, Rodica Simion and Daniel Ullman as a well-behaved sub-lattice of the intersection lattice for the Coxeter arrangement of type A, see e.g. [6, 7, 22]. Natural combinatorial interpretations of non-crossing partitions for the classical reflection groups were then given by Christos A. Athanasiadis and Vic Reiner in [3, 20].

On the other hand, non-nesting partitions were simultaneously introduced for all crystallographic reflection groups by Alex Postnikov as anti-chains in the associated root poset, see [20, Remark 2]; for further

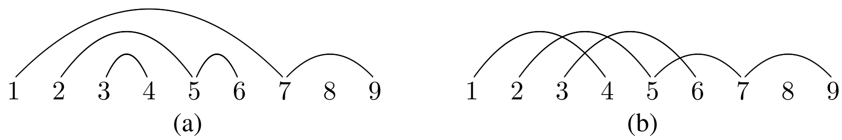


Fig. 1: A non-crossing (a) and a non-nesting (b) set partition of $[9]$.

information on reflection groups as well as for a definition of Coxeter arrangements and root posets see e.g. [14]

Within the last years, several bijections between non-crossing and non-nesting partitions have been constructed. In particular, type (block-size) preserving bijections were given by Christos A. Athanasiadis [2] for type A and by Alex Fink and Benjamin I. Giraldo [9] for the other classical reflection groups. Recently, Ricardo Mamede and Alessandro Conflitti [5, 19] constructed bijections for types A, B and D which turn out to be subsumed by the bijections we present here.

In the case of set partitions of type A, also the *number* of crossings and nestings was considered: Anisse Kasraoui and Jiang Zeng constructed a bijection which interchanges crossings and nestings in [16]. Finally, in a rather different direction, William Y.C. Chen, Eva Y.P. Deng, Rosena R.X. Du, Richard P. Stanley [4] have shown that the number of set partitions where a *maximal crossing* has cardinality k and a *maximal nesting* has cardinality ℓ is the same as the number of set partitions where a maximal crossing has cardinality ℓ and a maximal nesting has cardinality k.

In this extended abstract, we present bijections on various classes of set partitions of classical types that preserve openers and closers. In particular, the bijection by Anisse Kasraoui and Jiang Zeng as well as the bijection by William Y.C. Chen, Eva Y.P. Deng, Rosena R.X. Du, Richard P. Stanley enjoy this property. We give generalizations of these bijections for the other classical reflection groups, whenever possible. Furthermore we show that the bijection is in fact (essentially) unique for the class of non-crossing and non-nesting set partitions.

1 Set partitions for classical types

A *set partition* of $[n] := \{1, 2, 3, \dots, n\}$ is a collection $\mathcal{B}$ of pairwise disjoint, non-empty subsets of $[n]$, called *blocks*, whose union is $[n]$. We visualize $\mathcal{B}$ by placing the numbers $1, 2, \dots, n$ in this order on a line and then joining *consecutive* elements of each block by an arc, see Figure 1 for examples.

The *openers* $\mathrm{op}(\mathcal{B})$ are the non-maximal elements of the blocks in $\mathcal{B}$, whereas the *closers* $\mathrm{cl}(\mathcal{B})$ are its non-minimal elements. For example, the set partitions in Figure 1 both have $\mathrm{op}(\mathcal{B}) = \{1, 2, 3, 5, 7\}$ and $\mathrm{cl}(\mathcal{B}) = \{4, 5, 6, 7, 9\}$.

A pair $(\mathcal{O}, \mathcal{C}) \subseteq [n] \times [n]$ is an *opener-closer configuration*, if $|\mathcal{O}| = |\mathcal{C}|$ and

$$|\mathcal{O} \cap [k]| \geq |\mathcal{C} \cap [k+1]| \quad \text{for} \quad k \in \{0, 1, \dots, n-1\},$$

or, equivalently, $(\mathcal{O}, \mathcal{C}) = \big(\mathrm{op}(\mathcal{B}), \mathrm{cl}(\mathcal{B})\big)$ for some set partition $\mathcal{B}$ of n.

A set partition of type B_n or C_n is a set partition $\mathcal{B}$ of $[\pm n] := \{1, 2, \dots, n, -1, -2, \dots, -n\}$, such that

$$B \in \mathcal{B} \Leftrightarrow -B \in \mathcal{B} \tag{1}$$

and such that there exists at most one block $B_0 \in \mathcal{B}$ (called the *zero block*) for which $B_0 = -B_0$.

A set partition $\mathcal{B}$ of type D_n is a set partition of type B_n (or C_n) where the zero block, if present, must not consist of a single pair $\{i, -i\}$.

2 Crossings and nestings in set partitions of type A

In this extended abstract, we refine the following well known correspondences between non-crossing and non-nesting set partitions. For ordinary set partitions, a *crossing* consists of a pair of arcs (i, j) and (i', j') such that $i < i' < j < j'$,

$$1 \quad \ldots \quad i \quad < \quad i' \quad < \quad j \quad < \quad j' \quad \ldots \quad n.$$

On the other hand, a *nesting* consists of a pair of arcs (i, j) and (i', j') such that $i < i' < j' < j$,

$$1 \quad \ldots \quad i \quad < \quad i' \quad < \quad j' \quad < \quad j \quad \ldots \quad n.$$

A set partition of $[n]$ is called *non-crossing* (resp. *non-nesting*) if the number of crossings (resp. the number of nestings) equals 0.

It has been known for a long time that the numbers of non-crossing and non-nesting set-partitions of $[n]$ coincide. More recently, Anisse Kasraoui and Jiang Zeng have shown in [16] that much more is true:

Theorem 2.1 *There is an explicit bijection on set partitions of* $[n]$*, preserving the set of openers and the set of closers, and interchanging the number of crossings and the number of nestings.*

The construction in [16] also proves the following corollary:

Corollary 2.2 *For any opener-closer configuration* $(\mathcal{O}, \mathcal{C}) \subseteq [n] \times [n]$*, there exists a unique non-crossing set partition* $\mathcal{B}$ *of* $[n]$ *and a unique non-nesting set partition* $\mathcal{B}'$ *of* $[n]$ *such that*

$$\mathrm{op}(\mathcal{B}) = \mathrm{op}(\mathcal{B}') = \mathcal{O} \quad \textit{and} \quad \mathrm{cl}(\mathcal{B}) = \mathrm{cl}(\mathcal{B}') = \mathcal{C}.$$

Apart from the number of crossings or nestings, another natural statistic to consider is the cardinality of a 'maximal crossing' and of a 'maximal nesting': a *maximal crossing* of a set partition is a set of largest cardinality of mutually crossing arcs, whereas a *maximal nesting* is a set of largest cardinality of mutually nesting arcs. For example, in Figure 1(a), the arcs $\{(1, 7), (2, 5), (3, 4)\}$ form a maximal nesting of cardinality 3. In Figure 1(b) the arcs $\{(1, 4), (2, 5), (3, 6)\}$ form a maximal crossing.

The following symmetry property was shown by William Y.C. Chen, Eva Y.P. Deng, Rosena R.X. Du, Richard P. Stanley and Catherine H. Yan [4]:

Theorem 2.3 *There is an explicit bijection on set partitions, preserving the set of openers and the set of closers, and interchanging the cardinalities of a maximal crossing and a maximal nesting.*

Since a 'maximal crossing' of a non-crossing partition and a 'maximal nesting' of a non-nesting partition both have cardinality 1, Corollary 2.2 implies that this bijection coincides with the bijection by Anisse Kasraoui and Jiang Zeng for non-crossing and non-nesting partitions. In particular, we obtain the curious fact that in this case, the bijection maps non-crossing partitions with k nestings and maximal nesting having cardinality ℓ to non-nesting partitions with k crossings and maximal crossing having cardinality ℓ.

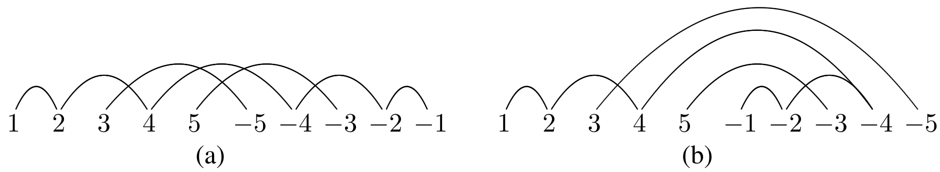


Fig. 2: The nesting (a) and the crossing (b) diagram of a set partition of type C_5.

We have to stress however, that in general it is not possible to interchange the number of crossings and the cardinality of a maximal crossing with the number of nestings and the cardinality of a maximal nesting simultaneously.

Example 2.4 *For $n = 8$, there is a set partition with one crossing, six nestings and the cardinalities of a maximal crossing and a maximal nesting equal both one, namely $\{\{1,7\},\{2,8\},\{3,4,5,6\}\}$. However, there is no set partition with six crossings, one nesting and cardinalities of a maximal crossing and a maximal nesting equal to one. To check, the four set partitions with six crossings and one nesting are*

$$\{\{1,4,6\},\{2,5,8\},\{3,7\}\}, \quad \{\{1,4,7\},\{3,5,8\},\{2,6\}\},$$
$$\{\{1,4,8\},\{2,5,7\},\{3,6\}\}, \quad \{\{1,5,8\},\{2,4,7\},\{3,6\}\}.$$

3 Crossings and nestings in set partitions of type C

Type independent definitions for *non-crossing* and *non-nesting* set partitions have been available for a while now, see for example [1, 2, 3, 20]. However, it turns out that the notions of crossing and nesting is more tricky, and we do not have a type independent definition. In this section we generalize the results of the previous section to type C.

We want to associate two pictures to each set partition, namely the ‘crossing’ and the ‘nesting diagram’. To this end, we define two orderings on the set $[\pm n]$: the *nesting order* for type C_n is

$$1 < 2 < \cdots < n < -n < \cdots < -2 < -1,$$

whereas the *crossing order* is

$$1 < 2 < \cdots < n < -1 < -2 < \cdots < -n.$$

The *nesting diagram* of a set partition $\mathcal{B}$ of type C_n is obtained by placing the numbers in $[\pm n]$ in *nesting order* on a line and then joining consecutive elements of each block of $\mathcal{B}$ by an arc, see Figure 2(a) for an example.

The *crossing diagram* of a set partition $\mathcal{B}$ of type C_n is obtained from the *nesting diagram* by reversing the order of the negative numbers. More precisely, we place the numbers in $[\pm n]$ in *crossing order* on a line and then join *consecutive elements in the nesting order* of each block of $\mathcal{B}$ by an arc, see Figure 2(b) for an example. We stress that the same elements are joined by arcs in both diagrams. Observe furthermore that the symmetry property (1) implies that if (i, j) is an arc, then its negative $(-j, -i)$ is also an arc.

A *crossing* is a pair of arcs that crosses in the crossing diagram, and a *nesting* is a pair of arcs that nests in the nesting diagram.

The *openers* $\mathrm{op}(\mathcal{B})$ are the *positive* non-maximal elements of the blocks in $\mathcal{B}$, the *closers* $\mathrm{cl}(\mathcal{B})$ the *positive* non-minimal elements. Thus, openers and closers are the start and end points of the arcs in the positive part of the nesting (or crossing) diagram. For example, the set partition displayed in Figure 2 has openers $\{1,2,3,4,5\}$ and closers $\{2,4\}$. For convenience, we call the negatives of the elements in $\mathrm{op}(\mathcal{B})$ *negative closers* and the negatives of the elements in $\mathrm{cl}(\mathcal{B})$ *negative openers*.

In type C_n, $(\mathcal{O},\mathcal{C}) \subseteq [n]\times[n]$ is an *opener-closer configuration*, if

$$|\mathcal{O}\cap[k]| \geq |\mathcal{C}\cap[k+1]| \quad \text{for} \quad k\in\{0,1,\dots,n-1\}.$$

Note that we do not require that $|\mathcal{O}| = |\mathcal{C}|$.

Theorem 3.1 *There is an explicit bijection on set partitions of type C_n, preserving the set of openers and the set of closers, and interchanging the number of crossings and the number of nestings.*

Furthermore, we will also see the following analog of Corollary 2.2:

Corollary 3.2 *For any opener-closer configuration $(\mathcal{O},\mathcal{C}) \subseteq [n]\times[n]$, there exists a unique non-crossing set partition $\mathcal{B}$ and a unique non-nesting set partition $\mathcal{B}'$, both of type C_n, such that*

$$\mathrm{op}(\mathcal{B}) = \mathrm{op}(\mathcal{B}') = \mathcal{O} \quad \textit{and} \quad \mathrm{cl}(\mathcal{B}) = \mathrm{cl}(\mathcal{B}') = \mathcal{C}.$$

Theorem 3.3 *There is an explicit bijection on set partitions of type C_n, preserving the set of openers and the set of closers, and interchanging the cardinalities of a maximal crossing and a maximal nesting.*

The bijection in Theorem 3.1 is constructed in an analogous way as the bijection in Theorem 2.1 whereas the rough idea of our bijection in the Theorem 3.3 is as follows: we render a type C_n set partition in the language of 0-1-fillings of a certain polyomino. We will do this in such a way that maximal nestings correspond to north-east chains of ones of maximal length.

Interpreting this filling as a growth diagram in the sense of Sergey Fomin and Tom Roby [10, 11, 12, 21] enables us to define a transformation on the filling that maps – technicalities aside – the length of the longest north-east chain to the length of the longest south-east chain. This filling can then again be interpreted as a C_n set partition, where south-east chains of maximal length correspond to maximal crossings. Many variants of the transformation involved are described in Christian Krattenthaler's article [17], we employ yet another (slight) variation.

We remark that the bijection in Theorem 3.3 is not an involution and that it does not, as discussed above, exchange the number of crossings and the number of nestings.

Remark 1 *Using the same methods as in the proof the above theorem, one can also deduce a conjecture due to Daniel Soll and Volkmar Welker [23, Conjecture 13]. Namely, we consider generalized triangulations of the $2n$-gon that are invariant under rotation of $180°$, and such that at most k diagonals are allowed to cross mutually. Daniel Soll and Volkmar Welker then conjectured that the number of such triangulations that are maximal, i.e., where one cannot add any diagonal without introducing a $k+1$ crossing, coincides with the number of* fans *of k Dyck paths that are symmetric with respect to a vertical axis.*

The corresponding theorem for type A was discovered and proved by Jakob Jonsson [15]. A (nearly) bijective proof very similar to ours was given by Christian Krattenthaler in [17]. A simple bijection for the case of 2-triangulations was recently given by Sergi Elizalde in [8].

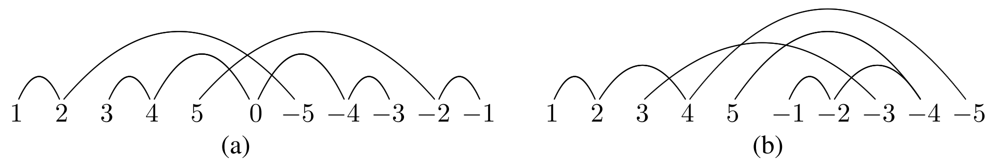


Fig. 3: The nesting (a) and the crossing (b) diagram of a set partition of type B_5.

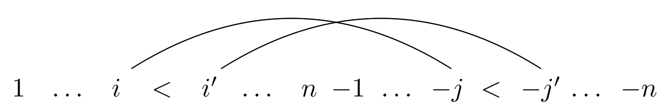
4 Crossings and nestings in set partitions of type B

The definition of non-crossing set partitions of type B_n coincides with the definition in type C_n, and the *crossing diagram* is also the same. However, the combinatorial model for non-nesting set partitions changes slightly: we define the *nesting order* for type B_n as

$$1 < 2 < \cdots < n < 0 < -n < \cdots < -2 < -1.$$

The *nesting diagram* of a set partition $\mathcal{B}$ is then obtained by placing the numbers in $[\pm n] \cup \{0\}$ in *nesting order* on a line and joining consecutive elements of each block of $\mathcal{B}$ by an arc, where the zero block is augmented by the number 0, if present. See Figure 3(a) for an example. The definition of *openers* $\mathrm{op}(\mathcal{B})$ and *closers* $\mathrm{cl}(\mathcal{B})$ is the same as in type C, the number 0 is neither an opener nor a closer.

These changes are actually dictated by the general, type independent definitions for non-crossing and non-nesting set partitions. Moreover, it turns out that we need to ignore certain crossings and nestings that appear in the diagrams: a *crossing* is a pair of arcs that crosses in the crossing diagram, except if both arcs connect a positive and a negative element and at least one of them connects a positive element with an element smaller in absolute value. Pictorially,



is not a crossing, if $j < i$ or $j' < i'$.

Similarly, a *nesting* is a pair of arcs that nests in the nesting diagram, except if both arcs connect a positive element or 0 with a negative element or 0, and at least one of them connects a positive element with an element smaller in absolute value.

Example 4.1 *The set partition in Figure 3(b) has three crossings:* $(3, -3)$ *crosses* $(2, 4)$*,* $(4, -5)$*, and* $(-4, -2)$*. It does not cross* $(5, -4)$ *by definition.*

The set partition in Figure 3(a) has three nestings: $(2, -5)$ *nests* $(3, 4)$ *and* $(4, 0)$*, and* $(5, -2)$ *nests* $(-4, -3)$*. However,* $(5, -2)$ *does not nest* $(0, -4)$ *by definition.*

With this definition, we have a theorem that is only slightly weaker than in type C:

Theorem 4.2 *There is an explicit bijection on set partitions of type* B_n*, preserving the set of openers and the set of closers, and mapping the number of nestings to the number of crossings.*

Again, we obtain an analog of Corollary 2.2:

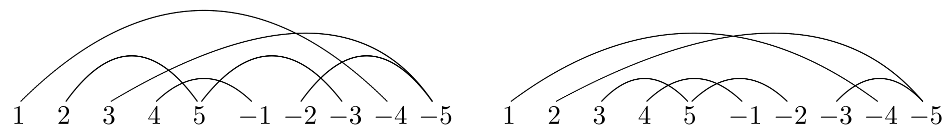


Fig. 4: Two non-crossing set partition of type D_5. Both are obtained from each other by interchanging 5 and -5.

Corollary 4.3 *For any opener-closer configuration* $(\mathcal{O},\mathcal{C}) \subseteq [n]\times[n]$*, there exists a unique non-crossing set partition* $\mathcal{B}$ *and a unique non-nesting set partition* $\mathcal{B}'$*, both of type* B_n*, such that*

$$\mathrm{op}(\mathcal{B}) = \mathrm{op}(\mathcal{B}') = \mathcal{O} \quad \textit{and} \quad \mathrm{cl}(\mathcal{B}) = \mathrm{cl}(\mathcal{B}') = \mathcal{C}.$$

Using Theorem 3.3 and Theorem 4.2, one can deduce the following theorem:

Theorem 4.4 *There is an explicit bijection on set partitions of type* B_n*, preserving the set of openers and the set of closers, and interchanging the cardinalities of a maximal crossing and a maximal nesting.*

5 Non-crossing and non-nesting set partitions in type D

In type D we do not have any good notion of crossing or nesting, we can only speak properly about *non-crossing* and *non-nesting* set partitions.

A combinatorial model for non-crossing set partition of type D_n was given by Christos A. Athanasiadis and Vic Reiner in [3]. For our purposes it is better to use a different description of the same model: let $\mathcal{B}$ be a set partition of type D_n and let $\{(i_1,-j_1),\ldots,(i_k,-j_k)\}$ for positive $i_\ell, j_\ell < n$ be the ordered set of arcs in $\mathcal{B}$ starting in $\{1,\ldots,n-1\}$ and ending in its negative. $\mathcal{B}$ is called *non-crossing* if

(i) $(i,-i)$ is an arc in $\mathcal{B}$ implies $i = n$,

and if it is non-crossing in the sense of type C_n with the following exceptions:

(ii) arcs in $\mathcal{B}$ containing n *must* cross all arcs $(i_\ell,-j_\ell)$ for $\ell > k/2$,

(iii) arcs in $\mathcal{B}$ containing $-n$ *must* cross all arcs $(i_\ell,-j_\ell)$ for $\ell \leq k/2$,

(iv) two arcs in $\mathcal{B}$ containing n and $-n$ *may* cross.

Here, (i) is equivalent to say that if $\mathcal{B}$ contains a zero block B_0 then $n \in B_0$ and observe that (i) together with the non-crossing property of $\{(i_1,-j_1),\ldots,(i_k,-j_k)\}$ imply that $k/2 \in \mathbb{N}$, see Figure 4 for an example.

Note that all conditions hold for a set partition $\mathcal{B}$ if and only if they hold for the set partition obtained from $\mathcal{B}$ by interchanging n and $-n$.

A set partition of type D_n is called *non-nesting* if it is non-nesting in the sense of [2]. This translates to our notation as follows: let $\mathcal{B}$ be a set partition of type D_n. Then $\mathcal{B}$ is called non-nesting if

(i) $(i,-i)$ is an arc in $\mathcal{B}$ implies $i = n$,

and if it is non-nesting in the sense of type C_n with the following exceptions:

(ii) arcs $(i,-n)$ and (j,n) for positive $i < j < n$ in $\mathcal{B}$ are allowed to nest, as do

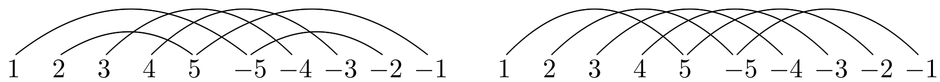


Fig. 5: Two non-nesting set partition of type D_5. Both are obtained from each other by interchanging 5 and -5.

(iii) arcs $(i, -j)$ and $(n, -n)$ for positive $k < i, j < n$ in $\mathcal{B}$ where (k, n) is another arc in $\mathcal{B}$ (which exists by the definition of set partitions in type D_n).

Again, (i) means that if $B_0 \in \mathcal{B}$ is a zero block then $n \in B_0$. (ii) and (iii) come from the fact that the positive roots $e_i + e_n$ and $e_j - e_n$ for $i \leq j$ are comparable in the root poset of type C_n but are not comparable in the root poset of type D_n. As for non-crossing set partitions in type D_n, all conditions hold if and only if they hold for the set partition obtained by interchanging n and $-n$. See Figure 5 for an example. The definition of *openers* $\mathrm{op}(\mathcal{B})$, *closers* $\mathrm{cl}(\mathcal{B})$ and *opener-closer configuration* is as in type C.

Proposition 5.1 *Let $(\mathcal{O}, \mathcal{C}) \subseteq [n]$ be an opener-closer configuration. Then there exists a non-crossing set partition $\mathcal{B}$ of type D_n with $\mathrm{op}(\mathcal{B}) = \mathcal{O}$ and $\mathrm{cl}(\mathcal{B}) = \mathcal{C}$ if and only if*

$$|\mathcal{O}| - |\mathcal{C}| \text{ is even or } n \in \mathcal{O}, \mathcal{C}. \tag{2}$$

Moreover, there exist exactly two non-crossing set partitions of type D_n having this opener-closer configuration if both conditions hold, otherwise, it is unique.

As in types A, B and C, the analogue proposition holds also for non-nesting set partitions of type D_n:

Proposition 5.2 *Let $(\mathcal{O}, \mathcal{C}) \subseteq [n]$ be an opener-closer configuration. Then there exists a non-nesting set partition $\mathcal{B}$ of type D_n with $\mathrm{op}(\mathcal{B}) = \mathcal{O}$ and $\mathrm{cl}(\mathcal{B}) = \mathcal{C}$ if and only if*

$$|\mathcal{O}| - |\mathcal{C}| \text{ is even or } n \in \mathcal{O}, \mathcal{C}. \tag{3}$$

Furthermore, there exist exactly two non-nesting set partitions of type D_n having this opener-closer configuration if both conditions hold, otherwise, it is unique.

References

[1] Drew Armstrong, *Generalized Noncrossing Partitions and Combinatorics of Coxeter Groups*, Ph.D. thesis, Cornell University, 2007, math.CO/0611106v1, to appear in Mem. Amer. Math. Soc.

[2] Christos A. Athanasiadis, *On noncrossing and nonnesting partitions for classical reflection groups*, Electronic Journal of Combinatorics **5** (1998), Research Paper 42, 16 pp. (electronic). MR MR1644234 (99i:05204)

[3] Christos A. Athanasiadis and Victor Reiner, *Noncrossing partitions for the group D_n*, SIAM Journal on Discrete Mathematics **18** (2004), no. 2, 397–417 (electronic). MR MR2112514 (2006b:06004)

[4] William Y. C. Chen, Eva Y. P. Deng, Rosena R. X. Du, Richard P. Stanley, and Catherine H. Yan, *Crossings and nestings of matchings and partitions*, Transactions of the American Mathematical Society **359** (2007), no. 4, 1555–1575 (electronic), math.CO/0501230. MR MR2272140 (2007i:05015)

[5] Alessandro Conflitti and Ricardo Mamede, *On noncrossing and nonnesting partitions of type d*, Preprint (2009), math.CO/08905.4371.

[6] Paul H. Edelman, *Chain enumeration and noncrossing partitions*, Discrete Mathematics **31** (1980), no. 2, 171–180. MR MR583216 (81i:05018)

[7] Paul H. Edelman and Rodica Simion, *Chains in the lattice of noncrossing partitions*, Discrete Mathematics **126** (1994), no. 1-3, 107–119. MR MR1264480 (95f:05012)

[8] Sergi Elizalde, *A bijection between 2-triangulations and pairs of non-crossing Dyck paths*, Preprint (2006), math.CO/0610235.

[9] Alex Fink and Benjamin Iriarte Giraldo, *Bijections between noncrossing and nonnesting partitions for classical reflection groups*, Proceedings of the 21st International Conference on Formal Power Series and Algebraic Combinatorics, Discrete Mathematics and Theoretical Computer Science, DMTCS, 2009, math.CO/0810.2613v3, pp. 399–412.

[10] Sergey Fomin, *The generalized Robinson-Schensted-Knuth correspondence*, Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta imeni V. A. Steklova Akademii Nauk SSSR (LOMI) **155** (1986), no. Differentsialnaya Geometriya, Gruppy Li i Mekh. VIII, 156–175, 195. MR MR869582 (88b:06003)

[11] Sergey V. Fomin, *Duality of graded graphs*, Journal of Algebraic Combinatorics **3** (1994), no. 4, 357–404. MR MR1293822 (95i:05088)

[12] ______, *Schensted algorithms for dual graded graphs*, Journal of Algebraic Combinatorics **4** (1995), no. 1, 5–45. MR MR1314558 (95m:05246)

[13] Curtis Greene, *An extension of Schensted's theorem*, Advances in Mathematics **14** (1974), 254–265. MR MR0354395 (50 #6874)

[14] James E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics 29 (1990)

[15] Jakob Jonsson, *Generalized triangulations and diagonal-free subsets of stack polyominoes*, Journal of Combinatorial Theory, Series A **112** (2005), no. 1, 117–142.

[16] Anisse Kasraoui and Jiang Zeng, *Distribution of crossings, nestings and alignments of two edges in matchings and partitions*, Electronic Journal of Combinatorics **13** (2006), no. 1, Research Paper 33, 12 pp. (electronic), math.CO/0601081. MR MR2212506 (2006k:05021)

[17] Christian Krattenthaler, *Growth diagrams, and increasing and decreasing chains in fillings of Ferrers shapes*, Advances in Applied Mathematics **37** (2006), no. 3, 404–431, math.CO/0510676. MR MR2261181 (2007h:05011)

[18] Germain Kreweras, *Sur les partitions non croisées d'un cycle*, Discrete Mathematics **1** (1972), no. 4, 333–350. MR MR0309747 (46 #8852)

[19] Ricardo Mamede, *A bijection between noncrossing and nonnesting partitions of types A and B*, Proceedings of the 21st International Conference on Formal Power Series and Algebraic Combinatorics, Discrete Mathematics and Theoretical Computer Science, DMTCS, 2009, math.CO/0810.1422, pp. 599–612.

[20] Victor Reiner, *Non-crossing partitions for classical reflection groups*, Discrete Mathematics **177** (1997), no. 1-3, 195–222. MR MR1483446 (99f:06005)

[21] Tom Roby, *Applications and extensions of Fomin's generalization of the Robinson-Schensted correspondence to differential posets*, Ph.D. thesis, M.I.T., Cambridge, Massachusetts, 1991.

[22] Rodica Simion and Daniel Ullman, *On the structure of the lattice of noncrossing partitions*, Discrete Mathematics **98** (1991), no. 3, 193–206. MR MR1144402 (92j:06003)

[23] Daniel Soll and Volkmar Welker, *Type-B generalized triangulations and determinantal ideals*, Discrete Mathematics (2006), math.CO/0607159.

[24] Richard P. Stanley, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999, With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. MR MR1676282 (2000k:05026)

[25] Christian Stump, *Non-crossing partitions, non-nesting partitions and Coxeter sortable elements in types A and B*, Preprint (2008), math.CO/0808.2822.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 884–891

The Hodge Structure of the Coloring Complex of a Hypergraph (Extended Abstract)

Sarah C Rundell[1] and Jane H Long[2]

[1] *Department of Math and Computer Science, Denison University, Granville, OH, 43023, USA*
[2] *Department of Mathematics, Stephen F. Austin State University, Nacodogches, TX, 75962*

Abstract. Let G be a simple graph with n vertices. The coloring complex $\Delta(G)$ was defined by Steingrímsson, and the homology of $\Delta(G)$ was shown to be nonzero only in dimension $n-3$ by Jonsson. Hanlon recently showed that the Eulerian idempotents provide a decomposition of the homology group $H_{n-3}(\Delta(G))$ where the dimension of the j^{th} component in the decomposition, $H^{(j)}_{n-3}(\Delta(G))$, equals the absolute value of the coefficient of λ^j in the chromatic polynomial of G, $\chi_G(\lambda)$. Let H be a hypergraph with n vertices. In this paper, we define the coloring complex of a hypergraph, $\Delta(H)$, and show that the coefficient of λ^j in $\chi_H(\lambda)$ gives the Euler Characteristic of the j^{th} Hodge subcomplex of the Hodge decomposition of $\Delta(H)$. We also examine conditions on a hypergraph, H, for which its Hodge subcomplexes are Cohen-Macaulay, and thus where the absolute value of the coefficient of λ^j in $\chi_H(\lambda)$ equals the dimension of the j^{th} Hodge piece of the Hodge decomposition of $\Delta(H)$.

Résumé. Soit G un graphe simple à n sommets. Le complexe de coloriage $\Delta(G)$ a été défini par Steingrímsson et Jonsson a prouvé que l'homologie de $\Delta(G)$ est non nulle seulement en dimension $n-3$. Hanlon a récemment prouvé que les idempotents eulériens fournissent une décomposition du groupe d'homologie $H_{n-3}(\Delta(G))$ où la dimension de la j^e composante dans la décomposition de $H_{n-3}(j)(\Delta(G))$ est égale à la valeur absolue du coefficient de λ^j dans le polynôme chromatique de G, $\chi_G(\lambda)$. Soit H un hypergraphe à n sommets. Dans ce texte, nous définissons le complexe de coloration d'un hypergraphe $\Delta(H)$ et nous prouvons que le coefficient de λ^j dans $\chi_H(\lambda)$ donne la caractéristique d'Euler du j^e sous-complexe de Hodge dans la décomposition de Hodge de $\Delta(H)$. Nous examinons également des conditions sur un hypergraphe H pour lesquelles les sous-complexes de Hodge sont Cohen-Macaulay. Ainsi la valeur absolue du coefficient de λ^j de $\chi_H(\lambda)$ est égale à la dimension du j^e sous-complexe de Hodge dans la décomposition de Hodge de $\Delta(H)$.

Keywords: coloring complex, hypergraph, chromatic polynomial

1 Preliminaries

Definition 1.1 *A* hypergraph*, H, is an ordered pair, (V, E), where V is a set of vertices and E is a set of subsets of the vertices of V. A hypergraph is said to be* uniform of rank r *if all of its hyperedges have size r.*

Throughout this paper, H will denote a hypergraph whose vertex set V is $\{1, ..., n\}$.

Definition 1.2 *The* chromatic polynomial *of H, denoted $\chi_H(\lambda)$, is the number of ways to color the vertices of the hyperedges of H with at most λ colors, so that the vertices of each edge are colored with at least two colors.*

Example 1.3 *Let H be the hypergraph on 5 vertices with hyperedges $\{1,2,3\}$ and $\{3,4,5\}$. It follows that $\chi_H(\lambda) = \lambda^5 - (\lambda^3 + \lambda^3) + \lambda = \lambda^5 - 2\lambda^3 + \lambda$.*

The following theorem is a generalization of a well-known result for the chromatic polynomial of a graph and will be used in the proof of Theorem 4.1. The same counting argument that is used to prove the deletion-contraction formula for the chromatic polynomial of a graph can be used to prove the following theorem.

Theorem 1.4 ***Deletion-Contraction Property*** *[11] Let e be a hyperedge of H. Let $H - e$ denote the hypergraph obtained from H by deleting the edge e, and let H/e denote the hypergraph obtained from H by identifying the vertices in e. Then*

$$\chi_H(\lambda) = \chi_{H-e}(\lambda) - \chi_{H/e}(\lambda).$$

Example 1.5 *Let H be as in Example 1.3, and let $e = \{1,2,3\}$. Then $H - e$ is the hypergraph on 5 vertices with hyperedge $\{3,4,5\}$, and H/e is the hypergraph on 3 vertices with edge $\{123,4,5\}$. Then $\chi_{H-e}(\lambda) = \lambda^2(\lambda^3 - \lambda)$ and $\chi_{H/e} = \lambda^3 - \lambda$. Thus, by Theorem 1.4*

$$\chi_H(\lambda) = \lambda^5 - 2\lambda^3 + \lambda.$$

2 The Coloring Complex

We begin by defining Steingrímsson's [10] coloring complex following the presentation in Jonsson [8].

Let $(B_1, ..., B_{r+2})$ be an ordered partition of $\{1, ..., n\}$ where at least one of the B_i contains a hyperedge of H, and let Δ_r be the set of ordered partitions $(B_1, ..., B_{r+2})$.

Definition 2.1 *The* coloring complex *of H is the sequence:*

$$... \to C_r \xrightarrow{\delta_r} C_{r-1} \xrightarrow{\delta_{r-1}} ... \xrightarrow{\delta_1} C_0 \xrightarrow{\delta_0} C_{-1} \xrightarrow{\delta_{-1}} 0$$

where C_r is the vector space over a field of characteristic zero with basis Δ_r and

$$\partial_r((B_1, ..., B_{r+2})) := \sum_{i=1}^{r+1} (-1)^i (B_1, ..., B_i \bigcup B_{i+1}, ..., B_{r+2}).$$

Notice that $\partial_{r-1} \circ \partial_r = 0$. Then:

Definition 2.2 *The r^{th} homology group of $\Delta(H)$, $H_r(\Delta(H)) := ker(\partial_r)/im(\partial_{r+1})$.*

It is worth noting that Hultman [7] defined a complex that includes both Steingrímsson's coloring complex and the coloring complex of a hypergraph as a special case.

For the proof of our main result, Theorem 4.1, we will need the following Lemma. Lemma 2.3 is the hypergraph version of Lemma 1.3 in Jonsson [8]. We follow the notation in Hanlon [3].

Lemma 2.3 *Let E be the hypergraph with the single hyperedge e and the same vertices as H, and let $C_r(\Delta(H), \Delta(E))$ be the vector space spanned by ordered partitions in $\Delta_r(H)$ with no block containing e. Then*

$$C_r(\Delta(H), \Delta(E)) \cong C_r(\Delta(H - e))/(C_r(\Delta(H - e)) \cap C_r(\Delta(E)))$$

and

$$C_r(\Delta(H - e)) \cap C_r(\Delta(E)) \cong C_r(\Delta(H/e)).$$

Note that $C_*(\Delta(H))/C_*(\Delta(E))$ is the same as $C_*(\Delta(H), \Delta(E))$ in the relative homology sense, and we will use this fact later in some of our proofs.

3 Eulerian Idempotents

Recently, Hanlon [3] showed that there is a Hodge decomposition of $H_{n-3}(\Delta(G))$ for a graph G; we will discuss this result and its generalization to the hypergraph case. In order to describe this decomposition, we must first define and describe the Eulerian idempotents. The Eulerian idempotents have many interesting properties and have proved useful in many different algebraic and combinatorial problems. For more information on Eulerian idempotents see [2], [9], [4], and [5].

Define a *descent* of a permutation $\pi \in S_n$ to be a couple of consecutive numbers $(i, i+1)$ such that $\pi(i) > \pi(i+1)$. It follows from Loday's [9] definition that the Eulerian idempotents $e_r^{(j)}$ can be defined by the identity:

Definition 3.1 *The Eulerian idempotents are defined by*

$$\sum_{j=1}^{n} t^j e_n^{(j)} = \sum_{\pi \in S_n} \binom{n+t-des(\pi)-1}{n} sgn(\pi)\pi,$$

where $des(\pi)$ *is the number of descents of* π.

There are several important properties of the Eulerian idempotents which are due to Gerstenhaber and Schack [2]. In their paper, they show that the Eulerian idempotents are mutually orthogonal idempotents and that their sum is the unit element in $\mathbb{C}[S_n]$. So then for any S_n-module, M, we have that

$$M = \bigoplus_j e_n^{(j)} M.$$

Notice that we can define an action of S_{r+2} on Δ_r. Namely, if $\sigma \in S_{r+2}$, then $\sigma \cdot (B_1, ..., B_{r+2}) = (B_{\sigma^{-1}(1)}, ..., B_{\sigma^{-1}(r+2)})$. This action then makes C_r into an S_{r+2}-module.

Hanlon [3] notes (or this result can be derived from the work of Gerstenhaber and Schack [2]) in Lemma 2.1 of his paper that for any graph G and for each r, j,

$$\partial_r \circ e_{r+2}^{(j)} = e_{r+1}^{(j)} \circ \partial_r.$$

This implies then that, for each j, $C_r^{(j)} = e_{r+2}^{(j)} C_r$ is a subcomplex of $(C_*(\Delta(G)), \partial_*)$. We may then consider the homology of the subcomplex, and it will be denoted by $H_*^{(j)}(\Delta(G))$. So then we have

$$H_r(\Delta(G)) = \bigoplus_j H_r^{(j)}(\Delta(G)).$$

The above decomposition is called the *Hodge decomposition* of $H_*(\Delta(G))$.

Hanlon [3] showed that there is a Hodge decomposition of the top homology group of $\Delta(G)$, i.e. $H_{n-3}(\Delta(G)) = \bigoplus_{j=1}^{n-1} H_{n-3}^{(j)}(\Delta(G))$. Further, he showed that the dimension of the j^{th} Hodge piece is equal to the absolute value of the coefficient of λ^j in the chromatic polynomial of G.

In the case where H is a hypergraph, the Hodge decomposition of $\Delta(H)$ can be defined by following the same process as above. However, in general, the homology of $\Delta(H)$ is not concentrated in one dimension, and thus Hanlon's result does not in general hold for an arbitrary hypergraph. In this paper, we will provide a generalization of Hanlon's result and study instances where the absolute value of the coefficient of λ^j in the chromatic polynomial of H is equal to the dimension of $H_*^{(j)}(\Delta(H))$.

4 The Relationship Between the Chromatic Polynomial of H and $\Delta(H)$

In this section, we will provide a generalization of Hanlon's result to hypergraphs. In our study, we will need the following definition:

Let $X^{(j)}$ denote the Euler Characteristic of the j^{th} Hodge piece of $\Delta(H)$. In particular,

$$\begin{aligned} X^{(j)} &= \sum_{i=-1}^{n-r-1} (-1)^i \dim(C_i^{(j)}(\Delta(H)) \\ &= \sum_{i=-1}^{n-r-1} (-1)^i \dim(H_i^{(j)}(\Delta(H)) \end{aligned}$$

Theorem 4.1 *Let H be a hypergraph on n vertices. Then*

$$X^{(j)}(\Delta(H)) = -[\lambda^j](\chi_H(-\lambda) - (-\lambda)^n).$$

where $[\lambda^j]\chi_H(-\lambda)$ denotes the coefficient of λ^j in $\chi_H(-\lambda)$.

Hultman [7] notes that a hypergraph H on n vertices, without inclusions among edges, may be associated with a subspace arrangement embeddable in the braid arrangement A_n. Following his construction and Theorem 5.7 of his paper, he notes that this then gives an interpretation of the chromatic polynomial of such hypergraphs in terms of Hilbert polynomials. Thus, for such hypergraphs, Theorem 4.1 above gives a relationship between the Euler Characteristics of the Hodge pieces of $\Delta(H)$ and coefficients in Hilbert polynomials.

5 Star Hypergraphs

The previous theorem leads to the following

Question 1 *For which hypergraphs H is it the case that, for each j, there exists at most one r for which $\dim(H_r^{(j)}(\Delta(H))$ is nonzero and therefore*

$$\dim(H_r^{(j)}(\Delta(H)) = (-1)^{r+1}[\lambda^j](\chi_H(-\lambda) - (-\lambda)^n)?$$

Recall the following definition:

Definition 5.1 *A simplicial complex Δ is* Cohen-Macaulay *over a ring R if $H_i(link_\Delta(\sigma); R) = 0$ for all $\sigma \in \Delta$ and $i < \dim(link_\Delta(\sigma))$.*

Notice that if $\Delta(H)$ is Cohen-Macaulay, then the above statement is true. In this section, we will show that if H is a star hypergraph, then $\Delta(H)$ is Cohen-Macaulay and thus satisfies the condition in Question 1.

Definition 5.2 *Let H be a uniform hypergraph of rank k with no singleton vertices. H is a* star hypergraph *if all of its hyperedges intersect in a common set of size $k-1$.*

Before we show that $\Delta(H)$ is Cohen-Macaulay, we need the following definitions.

Definition 5.3 *A simplicial complex is* pure *if all of its maximal faces have the same dimension.*

The definition of a constructible complex is due to Hochster [6].

Definition 5.4 *1. Simplices, including the empty set, are constructible.*

2. If Δ_1 and Δ_2 are n-dimensional constructible complexes and $\Delta_1 \cap \Delta_2$ is an $(n-1)$-dimensional constructible complex, then $\Delta_1 \cup \Delta_2$ is constructible.

As noted in Björner [1], constructible complexes are Cohen-Macaulay. We will show that for H, a star hypergraph, $\Delta(H)$ is constructible and thus Cohen-Macaulay.

Theorem 5.5 *If H is a star hypergraph, then $\Delta(H)$ is constructible, and hence,*

$$\begin{aligned}\dim(H^{(j)}_{n-k-1}(\Delta(H))) &= (-1)^{n-k}[\lambda^j](\chi_H(-\lambda)-(-\lambda)^n).\\ &= (-1)^{n-k}[\lambda^j](-\lambda(-\lambda-1)^{n-(k-1)}-(-\lambda)^{n-(k-1)})\end{aligned}$$

Corollary 5.6 *If H is a star hypergraph, then the dimension of $H_{n-k-1}(\Delta(H))$ equals the absolute value of the sum of the coefficients of $\chi_H(\lambda)$ minus one.*

6 Cohen-Macaulay Hodge Subcomplexes of $\Delta(H)$

Under certain conditions, while $\Delta(H)$ may not be Cohen-Macaulay, the Hodge subcomplexes of $\Delta(H)$ are Cohen-Macaulay and thus provide an alternative answer to Question 1. In this section, we will study these conditions. We begin by examining a class of hypergraphs which are Cohen-Macaulay when $n \leq 5$.

Theorem 6.1 *Let H be a uniform hypergraph of rank $k \neq 2$ having n vertices, $n \leq 5$. If each hyperedge of H intersects at least one other hyperedge of H in a set of size $k-1$, then $\Delta(H)$ is Cohen-Macaulay and hence,*

$$\dim(H^{(j)}_{n-k-1}(\Delta(H))) = (-1)^{n-k}[\lambda^j](\chi_H(-\lambda)-(-\lambda)^n).$$

We now present some results concerning the homology of hypergraphs built from H by including an additional edge, beginning with two lemmas concerning their homology. For any hypergraph H, and any hyperedge e of H, it can be seen directly from the definitions that there exists a short exact sequence

$$0 \to C_r(\Delta(E)) \to C_r(\Delta(H)) \to C_r(\Delta(H))/C_r(\Delta(E)) \to 0$$

where $\Delta(E)$ denotes the edge complex of e and $C_*(\Delta(H),\Delta(E))$ in degree r is understood to represent the vector space spanned by ordered partitions in $C_r(\Delta(H))$ with no block containing e; the boundary maps on $C_*(\Delta(H),\Delta(E))$ are induced by the boundary maps on $C_*(\Delta(H))$. We can show

Lemma 6.2 *Let H be a uniform hypergraph of rank k and e be a hyperedge of H. Then, for $r < n-k-1$, $H_r(\Delta(H)) \cong H_r(\Delta(H), \Delta(E))$ and $H_{n-k-1}(\Delta(H)) \cong H_{n-k-1}(\Delta(E)) \oplus H_{n-k-1}(\Delta(H), \Delta(E))$.*

Lemma 6.3 *Consider a uniform, rank k hypergraph H on n vertices such that $H - e$ is constructible for some hyperedge e in the edge set of H. Then, for $r < n - k - 1$, $H_r(\Delta(H)) \cong H_{r-1}(\Delta(H/e))$.*

From this lemma we can deduce:

Corollary 6.4 *Let H be a uniform, rank k hypergraph on n vertices such that, for some hyperedge e, $H - e$ is a star hypergraph and the largest size of any intersection of e with another edge of H is 1. Then the homology of $\Delta(H)$ is nonzero in dimensions $n - k - 1$ and $n - 2k + 1$. The dimension of $H_{n-k-1}(\Delta(H))$ equals the sum of the absolute values of the coefficients of $\chi_{H-e}(\lambda)$, and the dimension of $H_{n-2k+1}(\Delta(H))$ equals the sum of the absolute values of the coefficients of $\chi_{H/e}(\lambda)$ minus one.*

It is more difficult to analyze the case where H is a uniform, rank k hypergraph on n vertices such that, for some hyperedge e, $H - e$ is a star hypergraph and the largest size of any intersection of e with another edge of H is greater than one. It would be interesting, however, to know whether there is a nice formula for the dimensions of $H_r(\Delta(H))$ in this case.

We now have the following theorem from which we can deduce an answer to Question 1.

Theorem 6.5 *Let H be a uniform hypergraph of rank $k \neq 2$ having n vertices, $n \leq 5$, then all Hodge subcomplexes of $\Delta(H)$ are Cohen-Macaulay.*

Notice that this result is no longer true when $n = 6$:

Example 6.6 *Let H be the hypergraph with 6 vertices and edges $\{1234\}$, $\{1256\}$, and $\{3456\}$. It is straightforward to see that the dimensions of the Hodge pieces of the homology of $\Delta(H)$ are:*

$$\begin{array}{lll} \dim(H_1^{(1)}(\Delta(H))) = 0 & \dim(H_0^{(1)}(\Delta(H))) = 2 & \dim(H_{-1}^{(1)}(\Delta(H))) = 0 \\ \dim(H_1^{(2)}(\Delta(H))) = 3 & \dim(H_0^{(2)}(\Delta(H))) = 3 & \\ \dim(H_1^{(3)}(\Delta(H))) = 3 & & \end{array}$$

These results lead to the following open questions:

Question 2 *Is there a general condition for uniform hypergraphs with $n \geq 6$ vertices so that the Hodge subcomplexes of $\Delta(H)$ are Cohen-Macaulay?*

Question 3 *Let H be a uniform hypergraph of rank k. Is there a general formula for the dimensions of the homology groups of $\Delta(H)$?*

7 The Chromatic Polynomial of $\Delta(H)$ and the Hodge Decomposition of $H_*(\Delta(H))$

When each of the Hodge subcomplexes of $\Delta(H)$ have exactly one nonzero Hodge piece, there are some interesting relationships between the signs of the coefficients of the chromatic polynomial of $\Delta(H)$ and the Hodge decompositions of the homology groups of $\Delta(H)$. We present these results below.

Theorem 7.1 *Suppose the j^{th} Hodge subcomplex of $\Delta(H)$ has at most one nonzero Hodge piece for all j. If both*

1. *the coefficient of λ^j has the same sign as the coefficient of λ^{j-1} and*

2. *the homology of the j^{th} Hodge subcomplex is nonzero in the decomposition of $H_k(\Delta(H))$*

then the $(j-1)^{st}$ Hodge subcomplex is nonzero in the decomposition of $H_l(\Delta(H))$, where l is the largest integer, with the opposite parity as k, that is less than k and greater than -1, and for which $H_l(\Delta(H))$ is nonzero.

Consider the following example:

Example 7.2 *Let H be the hypergraph of 5 vertices with edges $\{1,2,3\}$, $\{2,4,5\}$, and $\{3,4,5\}$. It can be verified that $\chi_H(\lambda) = \lambda^5 - 3\lambda^3 + \lambda^2 + \lambda$, and the dimensions of the Hodge pieces of the homology of $\Delta(H)$ are:*

$$\begin{array}{lll} \dim(H_1^{(1)}(\Delta(H))) = 0 & \dim(H_0^{(1)}(\Delta(H))) = 1 & \dim(H_{-1}^{(1)}(\Delta(H))) = 0 \\ \dim(H_1^{(2)}(\Delta(H))) = 1 & \dim(H_0^{(2)}(\Delta(H))) = 0 & \\ \dim(H_1^{(3)}(\Delta(H))) = 3 & & \end{array}$$

In this case, the coefficient of λ^2 and λ have the same parity, and $\dim(H_1^{(2)}(\Delta(H))) = 1$. Notice that 0 is the only integer less than 1 and greater than -1 which is even. Thus, according to the theorem, $\dim(H_0^{(1)}(\Delta(H)) = 1$.

Similarly, one can show that:

Theorem 7.3 *Suppose the j^{th} Hodge subcomplex of $\Delta(H)$ has at most one nonzero Hodge piece for all j. If both*

1. *the coefficient of λ^j has the opposite sign as the coefficient of λ^l where l is the largest integer less than j such that l has the same parity as j (and all coefficients of λ^i are zero for $l < i < j$) and*

2. *if the homology of the j^{th} Hodge subcomplex is nonzero in the decomposition of $H_k(\Delta(H))$*

then the l^{th} Hodge subcomplex is nonzero in the decomposition of $H_m(\Delta(H))$, where m is the largest integer, with the opposite parity as l, that is less than k and greater than -1, and for which $H_m(\Delta(H))$ is nonzero.

Example 7.4 *Let H be the hypergraph of 5 vertices with edges $\{1,2,3\}$ and $\{3,4,5\}$. As determined earlier in the paper, $\chi_H(\lambda) = \lambda^5 - 2\lambda^3 + \lambda$, and the dimensions of the Hodge pieces of the homology of $\Delta(H)$ are:*

$$\begin{array}{lll} \dim(H_1^{(1)}(\Delta(H))) = 0 & \dim(H_0^{(1)}(\Delta(H))) = 1 & \dim(H_{-1}^{(1)}(\Delta(H))) = 0 \\ \dim(H_1^{(2)}(\Delta(H))) = 0 & \dim(H_0^{(2)}(\Delta(H))) = 0 & \\ \dim(H_1^{(3)}(\Delta(H))) = 2 & & \end{array}$$

8 Acknowledgement

The authors thank Philip Hanlon for his suggestion of the problem.

References

[1] A. Björner, "Topological methods", in *Handbook of Combinatorics*, R. Graham, M. Grötschel, and L. Lovász (Eds.), North-Holland/Elsevier, Amsterdam, 1995, pp. 1819-1872.

[2] M. Gerstenhaber and S. D. Schack, "The shuffle bialgebra and the cohomology of commutative algebras", *J. Pure Appl. Algebra*, **70**, No. 3 (1991), 263-272.

[3] P. Hanlon, "A Hodge decomposition interpretation for the coefficients of the chromatic polynomial", *Proc. Amer. Math. Soc.*, **136** (2008), 3741-3749.

[4] P. Hanlon, "Hodge structures on posets," *Proc. Amer. Math. Soc.*, **134** (2006), 1857-1867.

[5] P. Hanlon, "The action of S_n on the components of the Hodge decomposition of Hochschild homology", *Michigan Math. J.*, **37** (1990), 105-124.

[6] M. Hochster, "Rings of invariants of tori, Cohen-Macaulay rings generated by monomials, and polytopes," *Ann. Math.*, **96**(1972), 318-337.

[7] A. Hultman, "Link complexes of subspace arrangements," *European J. Combinatorics*, **28** (2007), 781 - 790.

[8] J. Jonsson, "The topology of the coloring complex", *J. Alg. Combinatorics*, **21**(2005), 311-329.

[9] J.-L. Loday, "Partition eulerienne et operations en homologie cyclique", *C. R. Acad. Sci. Paris Ser. I Math.*, **307** (1988), 283-286.

[10] E. Steingrímsson, "The coloring ideal and coloring complex of a graph," *J. Alg. Comb.* **14** (2001), 73-84.

[11] G. Whittle, "A geometric theory of hypergraph coloring", *Aequationes Math.*, **43** (1992), 45 - 58.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 892–903

Schubert complexes and degeneracy loci

Steven V Sam †

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract. The classical Thom–Porteous formula expresses the homology class of the degeneracy locus of a generic map between two vector bundles as an alternating sum of Schur polynomials. A proof of this formula was given by Pragacz by expressing this alternating sum as the Euler characteristic of a Schur complex, which gives an explanation for the signs. Fulton later generalized this formula to the situation of flags of vector bundles by using alternating sums of Schubert polynomials. Building on the Schubert functors of Kraśkiewicz and Pragacz, we introduce Schubert complexes and show that Fulton's alternating sum can be realized as the Euler characteristic of this complex, thereby providing a conceptual proof for why an alternating sum appears.

Résumé. La formule classique de Thom–Porteous exprime la classe d'homologie du locus de la dégénérescence d'une fonction générique entre deux fibrés vectoriels comme une somme alternée des polynômes de Schur. Un preuve de cette formule a été donnée par Pragacz en exprimant ce alternant somme comme la caractéristique d'Euler d'un complexe de Schur, ce qui donne une explication pour les signes. Fulton puis généralisée cette formule à la situation des drapeaux de fibrés vectoriels à l'aide alternant des sommes de polynômes de Schubert. S'appuyant sur le Schubert foncteurs de Kraśkiewicz et Pragacz, nous introduisons les complexes de Schubert et montrent que la somme alternée de Fulton peuvent être réalisées en tant que Euler caractéristique de ce complexe, fournissant ainsi une preuve conceptuelle pour lesquelles une somme alternée apparaît.

Keywords: Schubert polynomials, Schubert complexes, degeneracy loci, balanced labelings, Thom–Porteous formula

1 Introduction

Let X be a smooth variety, and let $\varphi\colon E \to F$ be a map of vector bundles over X, with ranks e and f respectively. Given a number $k \leq \min(e, f)$, let $D_k(\varphi)$ be the degeneracy locus of points x where the rank of φ restricted to the fiber of x is at most k. Then $\operatorname{codim} D_k(\varphi) \leq (e-k)(f-k)$, and in the case of equality, the Thom–Porteous formula gives an expression for the homology class of $D_k(\varphi)$ in the Chow groups of X in terms of the Chern classes of E and F using super Schur polynomials. Also in the case of equality, the Schur complex associated with the rectangular partition $(f-k) \times (e-k)$ [ABW] of φ is a linear locally free resolution for a coherent sheaf whose support is $D_k(\varphi)$. Interpreted appropriately, the Euler characteristic of this complex recovers the Thom–Porteous formula. Hence the complex provides a "linear approximation" of the syzygies of $D_k(\varphi)$.

The situation was generalized by Fulton as follows. We provide the additional data of a flag of subbundles $E_\bullet$ for E and a flag of quotient bundles $F_\bullet$ for F, and we can define degeneracy loci for an array of

† Steven Sam was supported by an NSF graduate fellowship and an NDSEG fellowship during the preparation of this work.

numbers which specifies the ranks of maps $E_p \to F_q$. In particular, the rank functions which give rise to irreducible degeneracy loci are indexed by permutations in a natural way. Under the right codimension assumptions, one can express its homology class as a substitution of a double Schubert polynomial with the Chern classes of the quotients E_i/E_{i-1} and the kernels $\ker(F_j \to F_{j-1})$. The motivation for this work was to complete the analogy of this situation with the previous one by constructing "Schubert complexes" which would be acyclic whenever the degeneracy loci has the right codimension.

Building on the constructions for Schubert functors by Kraśkiewicz and Pragacz of [KP], we construct these complexes over an arbitrary (commutative) ring R from the data of two free R-modules M_0, M_1, with given flags of submodules, respectively, quotient modules, and a map $\partial\colon M_0 \to M_1$. We show that they are generically acyclic (in the sense of [BE]) and that in general they are acyclic when a certain ideal defined in terms of minors of ∂ has the right depth, i.e., they are "depth-sensitive." This allows us to extend the construction to an arbitrary variety (or more generally, an arbitrary scheme). We will stick to the language of varieties, however the results can be generalized as necessary. Again, the complexes are linear and provide a "linear approximation" to the syzygies of Fulton's degeneracy loci. We remark here that as a special case of Fulton's degeneracy loci, one gets Schubert varieties inside of arbitrary partial flag varieties.

Our main result is that in the situation of Fulton's theorem, the complex is acyclic and the Euler characteristic provides the formula in the same sense as above. A majority of the hard work goes into proving that our constructed complexes are acyclic under the appropriate depth assumption. Our proof uses techniques from commutative algebra, algebraic geometry, and combinatorics, and will appear in the full version of this paper. In the present article, we offer a short sketch of the proof.

Using the work of Fomin, Greene, Reiner, and Shimozono [FGRS], we can also construct explicit bases for the terms of the Schubert complex in the case that M_0 and M_1 are free. This basis naturally extends their notion of "balanced labelings" and their generating function gives an alternative expression for double Schubert polynomials. Furthermore, the complex naturally affords a representation of the Lie superalgebra of upper triangular matrices (with respect to the given flags) in $\mathrm{Hom}(M_0, M_1)$, and its supercharacter is the double Schubert polynomial.

The article is structured as follows. In Section 2 we recall some facts about double Schubert polynomials and balanced labelings. We introduce balanced super labelings and explain their relationship with the double Schubert polynomials. In Section 3 we extend the construction for Schubert functors to the $\mathbf{Z}/2$-graded setting and describe a basis for them naturally indexed by the balanced super labelings. In Section 4 we construct the Schubert complex from this $\mathbf{Z}/2$-graded Schubert functor. We mention the relevant facts and sketch the idea of a proof that these complexes are generically acyclic, and that in general the acyclicity of the complex is controlled by depth of a Schubert determinantal ideal. We also give some examples of Schubert complexes. Finally, in Section 5, we relate the acyclicity of the Schubert complexes to a degeneracy locus formula of Fulton.

2 Double Schubert polynomials.

2.1 Preliminaries.

Let Σ_n be the permutation group on the set $\{1, \dots, n\}$. Let s_i denote the transposition which switches i and $i+1$. Then Σ_n is generated by $\{s_1, \dots, s_{n-1}\}$, and for $w \in \Sigma_n$, we define the **length** of w to be the least number of $\ell(w)$ such that $w = s_{i_1} \cdots s_{i_{\ell(w)}}$. Such a minimal expression is a **reduced decomposition** for w. We can also write $\ell(w) = \#\{i < j \mid w(i) > w(j)\}$. There is a unique word w_0 with maximal length, which is the permutation defined by $w_0(i) = n + 1 - i$.

We will use two partial orders on Σ_n. The **weak Bruhat order**, denoted by $u \le_W w$, holds if some reduced decomposition of u is the suffix of some reduced decomposition of w. We denote the **strong Bruhat order** by $u \le w$, which holds if some reduced decomposition of w contains a subword that is a reduced decomposition of u. It follows from the definition that $u \le w$ if and only if $u^{-1} \le w^{-1}$. For a permutation w, let $r_w(p,q) = \{i \le p \mid w(i) \le q\}$ be its **rank function**. Then $u \le w$ if and only if $r_u(p,q) \ge r_w(p,q)$ for all p and q (the inequality on rank functions is reversed).

Given a polynomial (with arbitrary coefficient ring) in the variables $\{x_i\}_{i\ge 1}$, let ∂_i be the **divided difference operator**

$$(\partial_i P)(x_1, x_2, \dots) = \frac{P(\dots, x_{i-1}, x_i, x_{i+1}, \dots) - P(\dots, x_{i-1}, x_{i+1}, x_i, \dots)}{x_i - x_{i+1}}. \tag{2.1}$$

The operators ∂_i satisfy the braid relations: $\partial_i\partial_j = \partial_j\partial_i$ when $|i-j| > 1$ and $\partial_i\partial_{i+1}\partial_i = \partial_{i+1}\partial_i\partial_{i+1}$.

For the long word $w_0 \in \Sigma_n$, set $\mathfrak{S}_{w_0}(x,y) = \prod_{i+j\le n}(x_i - y_j)$. In general, if $\ell(ws_i) = \ell(w) - 1$, we set $\mathfrak{S}_{ws_i}(x,y) = \partial_i \mathfrak{S}_w(x,y)$, where we interpret $\mathfrak{S}_w(x,y)$ as a polynomial in the variables $\{x_i\}_{i\ge 1}$ with coefficients in the ring $\mathbf{Z}[y_1, y_2, \dots]$. These polynomials are the **double Schubert polynomials**, and are well-defined thanks to the braid relations. They also enjoy the following stability property: if we embed Σ_n into Σ_{n+1} by identifying permutations of Σ_n with permutations of Σ_{n+1} which fix $n+1$, then the polynomial $\mathfrak{S}_w(x,y)$ is the same whether we regard w as an element of Σ_n or Σ_{n+1}.

Define the **single Schubert polynomials** by $\mathfrak{S}_w(x) = \mathfrak{S}_w(x,0)$. We will use the identity [Man, Proposition 2.4.7]

$$\mathfrak{S}_w(x,y) = \sum_{u \le_W w} \mathfrak{S}_u(x)\mathfrak{S}_{uw^{-1}}(-y). \tag{2.2}$$

2.2 Balanced super labelings.

For the rest of this article, we fix a totally ordered alphabet $\cdots < 3' < 2' < 1' < 1 < 2 < 3 < \cdots$.

For a permutation w, define its **diagram** $D(w) = \{(i, w(j)) \mid i < j,\ w(i) > w(j)\}$. Let T be a labeling of $D(w)$. The **hook** of a box $b \in D(w)$ is the set of boxes in the same column below it, and the set of boxes in the same row to the right of it (including itself). A hook is **balanced** (with respect to T) if it satisfies the following property: when the entries are rearranged so that they are weakly increasing going from the top right end to the bottom left end, the label in the corner stays the same. A labeling is **balanced** if all of the hooks are balanced. Call a labeling T of $D(w)$ with entries in our alphabet a **balanced super labeling (BSL)** if it is balanced, column-strict (no repetitions in any column) with respect to the unmarked alphabet, row-strict with respect to the marked alphabet, and satisfies $j' \le T(i,j) \le i$ for all i and j (this last condition will be referred to as the **flag conditions**).

Given a BSL T of $D(w)$, let $f_T(i)$, respectively $f_T(i')$, be the number of occurrences of i, respectively i'. Define a monomial

$$m(T) = x_1^{f_T(1)} \cdots x_{n-1}^{f_T(n-1)} (-y_1)^{f_T(1')} \cdots (-y_{n-1})^{f_T((n-1)')}. \tag{2.3}$$

Using [FGRS, Lemma 4.7, Theorem 4.8], we can prove the following.

Theorem 2.4 *For every permutation w,*

$$\mathfrak{S}_w(x,y) = \sum_T m(T),$$

where the sum is over all BSL T of $D(w)$.

Remark 2.5 Given a labeling T of $D(w)$, let T^* denote the labeling of $D(w^{-1})$ obtained by transposing T and performing the swap $i \leftrightarrow i'$. The operation $T \mapsto T^*$ gives a concrete realization of the symmetry $\mathfrak{S}_w(-y,-x) = \mathfrak{S}_{w^{-1}}(x,y)$ [Man, Corollary 2.4.2].

Example 2.6 We list the BSL for the permutation $w = 321$.

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 1' \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|}\hline 1 & 2' \\ \hline 2 \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|}\hline 1' & 2' \\ \hline 1 \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|}\hline 1 & 1' \\ \hline 2 \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|}\hline 1' & 1 \\ \hline 1' \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|}\hline 1' & 2' \\ \hline 2 \\ \cline{1-1} \end{array} \quad \begin{array}{|c|c|}\hline 1' & 2' \\ \hline 1' \\ \cline{1-1} \end{array}$$

In this case, $\mathfrak{S}_{321}(x,y) = (x_1 - y_1)(x_1 - y_2)(x_2 - y_1)$.

3 Double Schubert functors.

3.1 Super linear algebra preliminaries.

Let $V = V_0 \oplus V_1$ be a free super module over a commutative ring R with $V_0 = \langle e_1, \dots, e_n \rangle$ and $V_1 = \langle e'_1, \dots, e'_m \rangle$, and let $\mathfrak{gl}(m|n) = \mathfrak{gl}(V)$ be the Lie superalgebra of endomorphisms of V. Let $\mathfrak{b}(m|n) \subset \mathfrak{gl}(m|n)$ be the standard Borel subalgebra of upper triangular matrices with respect to the ordered basis $\langle e'_m, \dots, e'_1, e_1, \dots, e_n \rangle$. In the case $m = n$, we will write $\mathfrak{b}(n) = \mathfrak{b}(n|n)$, and if it is clear from context, we will drop the n and simply write $\mathfrak{b}$. Also, let $\mathfrak{b}(n)_0 = \mathfrak{gl}(V)_0 \cap \mathfrak{b}(n)$ be the even degree elements in $\mathfrak{b}(n)$, and again, we will usually denote this by simply $\mathfrak{b}_0$. We also write $\mathfrak{h}(n) \subset \mathfrak{b}(n)$ for the Cartan subalgebra of diagonal matrices (this is a Lie algebra concentrated in degree 0). Let $\varepsilon'_n, \dots, \varepsilon'_1, \varepsilon_1, \dots, \varepsilon_n$ be the dual basis vectors to the standard basis of $\mathfrak{h}(n)$. For notation, write $(a_n, \dots, a_1 | b_1, \dots, b_n)$ for $\sum_{i=1}^n (a_i \varepsilon'_i + b_i \varepsilon_i)$. The even and odd roots of $\mathfrak{b}(n)$ are $\Phi_0 = \{\varepsilon'_j - \varepsilon'_i,\ \varepsilon_i - \varepsilon_j \mid 1 \le i < j \le n\}$ and $\Phi_1 = \{\varepsilon'_i - \varepsilon_j \mid 1 \le i, j \le n\}$, respectively. The even and odd simple roots are $\Delta_0 = \{\varepsilon'_{i+1} - \varepsilon'_i,\ \varepsilon_i - \varepsilon_{i+1} \mid i = 1, \dots, n-1\}$ and $\Delta_1 = \{\varepsilon'_1 - \varepsilon_1\}$.

Given a highest weight representation W of $\mathfrak{b}(n)$, we have a weight decomposition $W = \bigoplus_\lambda W_\lambda$ as a representation of $\mathfrak{h}(n)$. Let Λ be the highest weight of W. Then every weight λ appearing in the weight decomposition can be written in the form $\Lambda - \sum n_\alpha \alpha$ where α ranges over the simple roots of $\mathfrak{b}(n)$ and $n_\alpha \in \mathbf{Z}_{\ge 0}$. For such a λ, set $\omega(\lambda) = (-1)^{\sum n_\alpha \deg \alpha}$. Then we define the **character** and **supercharacter** of W as

$$\operatorname{ch} W = \sum_\lambda (\dim W_\lambda) e^\lambda, \quad \operatorname{sch} W = \sum_\lambda \omega(\lambda)(\dim W_\lambda) e^\lambda. \tag{3.1}$$

We recall the $\mathbf{Z}/2$-graded analogues of the symmetric and exterior powers. Let $F = F_0 \oplus F_1$ be a free R-supermodule. Let D denote the divided power functor. Then $\bigwedge^i F$ and $\mathrm{D}^i F$ are $\mathbf{Z}$-graded modules with terms given by

$$(\bigwedge^i F)_d = \bigwedge^{i-d} F_0 \otimes \operatorname{Sym}^d F_1, \quad (\mathrm{D}^i F)_d = \mathrm{D}^{i-d} F_0 \otimes \bigwedge^d F_1. \tag{3.2}$$

We can define a coassociative $\mathbf{Z}$-graded comultiplication $\Delta\colon \mathrm{D}^{i+j}F \to \mathrm{D}^iF \otimes \mathrm{D}^jF$ as follows. On degree d, pick $0 \le a \le i$ and $0 \le b \le j$ such that $a+b=d$. Then we have the composition $\Delta_{a,b}$

$$\begin{aligned}
(\mathrm{D}^{i+j}F)_d &= \mathrm{D}^{i+j-a-b}F_0 \otimes \bigwedge^{a+b} F_1 \\
&\xrightarrow{\Delta\otimes\Delta} \mathrm{D}^{i-a}F_0 \otimes \mathrm{D}^{j-b}F_0 \otimes \bigwedge^{a} F_1 \otimes \bigwedge^{b} F_1 \\
&\cong \mathrm{D}^{i-a}F_0 \otimes \bigwedge^{a} F_1 \otimes \mathrm{D}^{j-b}F_0 \otimes \bigwedge^{b} F_1 = (\mathrm{D}^iF)_a \otimes (\mathrm{D}^jF)_b,
\end{aligned} \tag{3.3}$$

where Δ is the usual symmetrization map, and we define Δ on the degree d part to be $\sum_{a+b=d} \Delta_{a,b}$.

Similarly, we can define an associative $\mathbf{Z}$-graded multiplication $m\colon \bigwedge^i F \otimes \bigwedge^j F \to \bigwedge^{i+j} F$ as follows. For degrees a and b, we have

$$\begin{aligned}
(\bigwedge^i F)_a \otimes (\bigwedge^j F)_b &= \bigwedge^{i-a} F_0 \otimes \mathrm{Sym}^a F_1 \otimes \bigwedge^{j-b} F_0 \otimes \mathrm{Sym}^b F_1 \\
&\cong \bigwedge^{i-a} F_0 \otimes \bigwedge^{j-b} F_0 \otimes \mathrm{Sym}^a F_1 \otimes \mathrm{Sym}^b F_1 \\
&\xrightarrow{m\otimes m} \bigwedge^{i+j-a-b} F_0 \otimes \mathrm{Sym}^{a+b} F_1 = (\bigwedge^{i+j} F)_{a+b},
\end{aligned} \tag{3.4}$$

where m is the usual exterior multiplication.

3.2 Constructions.

Define a flag of $\mathbf{Z}/2$-graded submodules

$$V^\bullet : V^{-n} \subset \cdots \subset V^{-1} \subset V^1 \subset \cdots \subset V^n \tag{3.5}$$

such that V^{-1} consists of all of the odd elements of V^n. We will say that the flag is **split** if each term and each quotient is a free module. Fix a permutation $w \in \Sigma_n$. Let $r_k = r_k(w)$, respectively $c_j = c_j(w)$, be the number of boxes in the kth row, respectively jth column, of $D(w)$. Define $\chi_{k,j}$ to be 1 if $(k,j) \in D(w)$ and 0 otherwise. Consider the map

$$\begin{aligned}
\bigotimes_{k=1}^{n-1} \mathrm{D}^{r_k} V^k &\xrightarrow{\otimes\Delta} \bigotimes_{k=1}^{n-1} \bigotimes_{j=1}^{n-1} \mathrm{D}^{\chi_{k,j}} V^k \cong \bigotimes_{j=1}^{n-1} \bigotimes_{k=1}^{n-1} \mathrm{D}^{\chi_{k,j}} V^k \\
&\xrightarrow{\otimes m} \bigotimes_{j=1}^{n-1} \bigwedge^{c_j} V^{w^{-1}(j)} \xrightarrow{\otimes\pi} \bigotimes_{j=1}^{n-1} \bigwedge^{c_j} (V^{w^{-1}(j)}/V^{-j-1}),
\end{aligned} \tag{3.6}$$

where $\otimes\Delta$ denotes the product of symmetrization operations, $\otimes m$ denotes the product of exterior multiplications, and $\otimes\pi$ denotes the product of projection maps. Then its image $\mathscr{S}_w(V^\bullet)$ is the **$\mathbf{Z}/2$-graded Schubert functor**, or **double Schubert functor**. By convention, the empty tensor product is R, so that if w is the identity permutation, then $\mathscr{S}_w(V^\bullet) = R$ with the trivial action of $\mathfrak{b}(n)$.

This definition is clearly functorial: given an even map of flags $f\colon V^\bullet \to W^\bullet$, i.e., $f(V^k) \subset W^k$ for $-n \le k \le n$, we have an induced map $f\colon \mathscr{S}_w(V^\bullet) \to \mathscr{S}_w(W^\bullet)$.

We will focus on the case when $V^{-i} = \langle e'_n, e'_{n-1}, \cdots, e'_i \rangle$ and $V^i = V^{-1} + \langle e_1, e_2, \dots, e_i \rangle$, so that $\mathscr{S}_w = \mathscr{S}_w(V^\bullet)$ is a $\mathfrak{b}(n)$-module.

Remark 3.7 One could dually define the double Schubert functor as the image of (dual) exterior powers mapping to symmetric powers. The dual exterior powers are as in our definition, except that divided powers replace symmetric powers. However, we have chosen this definition to be consistent with [KP]. This will be especially convenient in Section 4.1.

Here is a combinatorial description of the map (3.6). The elements of $\bigotimes_{k=1}^{n-1} \mathrm{D}^{r_k} V^k$ can be thought of as labelings of $D = D(w)$ such that in row k, only the labels $n', (n-1)', \dots, 1', 1, \dots, k$ are used, such that there is at most one use of i' in a given row, and such that the entries in each row are ordered in the usual way (i.e., $n' < (n-1)' < \cdots < 1' < 1 < \cdots < k$). Let Σ_D be the permutation group of D. We say that $\sigma \in \Sigma_D$ is **row-preserving** if each box and its image under σ are in the same row. Denote the set of row-preserving permutations as $\mathrm{Row}(D)$. Let T be a labeling of D that is row-strict with respect to the marked letters. Let $\mathrm{Row}(D)_T$ be the subgroup of $\mathrm{Row}(D)$ that leaves T fixed, and let $\mathrm{Row}(D)^T$ be the set of cosets $\mathrm{Row}(D)/\mathrm{Row}(D)_T$. Given $\sigma \in \mathrm{Row}(D)^T$, and considering the boxes as ordered from left to right, let $\alpha(T,\sigma)_k$ be the number of inversions of σ among the *marked letters* in the kth row, and define $\alpha(T,\sigma) = \sum_{k=1}^{n-1} \alpha(T,\sigma)_k$. Note that this number is independent of the representative chosen since T is row strict with respect to the marked letters. Then the comultiplication sends T to $\sum_{\sigma \in \mathrm{Row}(D)^T} (-1)^{\alpha(T,\sigma)} \sigma T$ where σT is the result of permuting the labels of T according to σ.

For the multiplication map, we can interpret the columns as being alternating in the *unmarked letters* and symmetric in the *marked letters*. We write $m(T)$ for the image of T under this equivalence relation. Therefore, the map (3.6) can be defined as

$$T \mapsto \sum_{\sigma \in \mathrm{Row}(D)^T} (-1)^{\alpha(T,\alpha)} m(\sigma T). \tag{3.8}$$

3.3 A basis and a filtration.

Theorem 3.9 *Assume that the flag $V^\bullet$ is split. The images of the BSLs under* (3.6) *form a basis over R for $\mathscr{S}_w$. By convention, the empty diagram has exactly one labeling.*

Corollary 3.10 *Identify $x_i = -e^{\varepsilon_i}$ and $y_i = -e^{\varepsilon'_i}$ for $1 \le i \le n$. Then*

$$\operatorname{ch} \mathscr{S}_w = \mathfrak{S}_w(-x, y), \quad \operatorname{sch} \mathscr{S}_w = \mathfrak{S}_w(x, y).$$

Corollary 3.11 *Choose an ordering of the set of permutations below w in the weak Bruhat order: $1 = v_1 \prec v_2 \prec \cdots \prec v_N = w$ such that $v_i \prec v_{i+1}$ implies that $\ell(v_i) \le \ell(v_{i+1})$. Then there exists a $\mathfrak{b}$-equivariant filtration*

$$0 = F_0 \subset F_1 \subset \cdots \subset F_N = \mathscr{S}_w$$

such that

$$F_i/F_{i-1} \cong \mathscr{S}'_{v_i} \otimes \mathscr{S}''_{wv_i^{-1}}$$

as $\mathfrak{b}_0$-modules.

Our proof does not establish how one can write the image of an arbitrary labeling as a linear combination of the images of the BSLs. Such a straightening algorithm is preferred, but we have not been successful in finding one, so we leave this task as an open problem.

Problem 3.12 *Find an algorithm for writing the image of an arbitrary labeling of $D(w)$ as a linear combination of the images of the BSLs of $D(w)$.*

4 Schubert complexes.

Now we can use the above machinery to define Schubert complexes. We start with the data of two flags $F_0^\bullet : 0 = F_0^0 \subset F_0^1 \subset \cdots \subset F_0^n = F_0$ and $F_1^\bullet : 0 = F_1^{-n-1} \subset F_1^{-n} \subset F_1^{-n+1} \subset \cdots \subset F_1^{-1} = F_1$, and a map $\partial\colon F_0 \to F_1$ between them. Given the flag for F_0, we pick an ordered basis $\{e_1, \dots, e_n\}$ for it such that $e_i \in F_0^i \setminus F_0^{i-1}$. Similarly, we pick an ordered basis $\{e'_1, \dots, e'_n\}$ for F_1 such that $e'_i \in F_1^{-i} \setminus F_1^{-i-1}$. Given these bases, we can represent ∂ as a matrix. This matrix representation will be relevant for the definition of certain ideals later.

Equivalently, we can give $F_1^\bullet$ as a quotient flag $F_1 = G^n \twoheadrightarrow G^{n-1} \twoheadrightarrow \cdots \twoheadrightarrow G^1 \twoheadrightarrow G^0 = 0$, so that the correspondence is given by $F_1^{-i} = \ker(G^n \twoheadrightarrow G^{i-1})$. Note that $F_1^{-i}/F_1^{-i-1} = \ker(G^i \twoheadrightarrow G^{i-1})$. We assume that each quotient has rank 1. Then we form a flagged supermodule F with even part F_0 and odd part F_1. The formation of symmetric and exterior products commutes with the differential ∂, so we can form the **Schubert complex** $\mathscr{S}_w(F)$ for a permutation $w \in \Sigma_n$.

Proposition 4.1 *The ith term of $\mathscr{S}_w(F)$ has a natural filtration whose associated graded is*

$$\bigoplus_{\substack{v \leq_W w \\ \ell(v) = i}} \mathscr{S}_v(F_0) \otimes \mathscr{S}_{wv^{-1}}(F_1).$$

Proof: This is a consequence of Corollary 3.11. □

4.1 The Kraśkiewicz–Pragacz filtration.

In order to prove properties of $\mathscr{S}_w$, we will construct a filtration of subcomplexes, which is based on the filtration of the single Schubert functors introduced by Kraśkiewicz and Pragacz [KP].

Let $w \in \Sigma_n$ be a nonidentity permutation. Consider the set of pairs (α, β) such that $\alpha < \beta$ and $w(\alpha) > w(\beta)$. Choose (α, β) to be maximal with respect to the lexicographic ordering. Let $k_1 < \cdots < k_k$ be the numbers such that $k_t < \alpha$ and $w(k_t) < w(\beta)$, and such that $k_t < i < \alpha$ implies that $w(i) \notin \{w(k_t), w(k_t) + 1, \dots, w(\beta)\}$. Then we have the following identity of double Schubert polynomials

$$\mathfrak{S}_w = \mathfrak{S}_v \cdot (x_\alpha - y_{w(\beta)}) + \sum_{t=1}^{k} \mathfrak{S}_{\psi_t}, \tag{4.2}$$

where $v = wt_{\alpha,\beta}$ and $\psi_t = wt_{\alpha,\beta}t_{k_t,\alpha}$. Here $t_{i,j}$ denotes the transposition which switches i and j. See, for example, [Man, Exercise 2.7.3]. The formula in (4.2) will be called a **maximal transition** for w. Define the **index** of a permutation u to be the number $\sum (k-1)\#\{j > k \mid u(k) > u(j)\}$. Note that the index of ψ_t is smaller than the index of w.

Theorem 4.3 *Let $V^\bullet$ be a split flag as in* (3.5). *Given a permutation $w \in \Sigma_n$, let* (4.2) *be the maximal transition for w. Then there exists a functorial $\mathfrak{b}$-equivariant filtration*

$$0 = F_0 \subset F_1 \subset \cdots \subset F_k \subset F' \subset F = \mathscr{S}_w(V^\bullet)$$

such that $F/F' \cong \mathscr{S}_v(V^\bullet) \otimes V^\alpha / V^{\alpha-1}$, $F'/F_k \cong \mathscr{S}_v(V^\bullet) \otimes V^{-w(\beta)}/V^{-w(\beta)+1}$, and $F_t/F_{t-1} \cong \mathscr{S}_{\psi_t}(V^\bullet)$ for $t = 1, \dots, k$.

Corollary 4.4 *Let $\partial\colon F_0 \to F_1$ be a map. With the notation as in Theorem 4.3, there is a functorial $\mathfrak{b}$-equivariant filtration of complexes*

$$0 = C_0 \subset C_1 \subset \cdots \subset C_k \subset C' \subset C = \mathscr{S}_w(\partial)$$

such that $C/C' \cong \mathscr{S}_v(\partial)[-1] \otimes F_0^\alpha / F_0^{\alpha-1}$, $C'/C_k \cong \mathscr{S}_v(\partial) \otimes F_1^{-w(\beta)}/F_1^{-w(\beta)+1}$, and $C_t/C_{t-1} \cong \mathscr{S}_{\psi_t}(\partial)$ for $t = 1, \dots, k$.

Proof: The filtration of Theorem 4.3 respects the differentials since everything is defined in terms of multilinear operations. The grading shift of C/C' follows from the fact that the F_0 terms have homological degree 1. $\square$

4.2 Generic acyclicity of Schubert complexes.

Given a matrix ∂ and a permutation w, let $I_w(\partial)$ be the ideal generated by the $(r_w(p,q)+1)\times(r_w(p,q)+1)$ minors of the upper left $p \times q$ submatrix of ∂. It is clear that $I_v \subseteq I_w$ if and only if $v \leq w$. In the case that ∂ is a generic matrix of variables over some coefficient ring R, let $X(w)$ be the variety defined by $I_w(\partial) \subset R[\partial_{i,j}]$. We refer to the ideals $I_w(\partial)$ as **Schubert determinantal ideals**, and the varieties $X(w)$ as **matrix Schubert varieties**. These ideals are prime and have codimension $\ell(w)$ [MS, Chapter 15].

Our main result is the following.

Theorem 4.5 *Let $A = K[\partial_{i,j}]$ be a polynomial ring over a field K, and let $\partial\colon F_0 \to F_1$ be a generic map of variables between two free A-modules.*

(a) *The Schubert complex $\mathscr{S}_w(\partial)$ is acyclic, and resolves a Cohen–Macaulay module M of codimension $\ell(w)$ supported in $I_{w^{-1}}(\partial) \subseteq A$.*

(b) *The restriction of M to $X(w^{-1})$ is a line bundle outside of a certain codimension 2 subset.*

(c) *The Schubert complex defined over the integers is acyclic.*

Proof: (Sketch). From Corollary 4.4, we get a short exact sequence

$$0 \to \mathrm{H}_1(C) \to \mathrm{H}_0(\mathscr{S}_v(\partial)) \otimes F_0^\alpha / F_0^{\alpha-1} \xrightarrow{\delta} \mathrm{H}_0(C') \to \mathrm{H}_0(C) \to 0, \tag{4.6}$$

so we have to show that δ is injective, and that the support of $\mathrm{H}_0(C) = M$ is $P = I_{w^{-1}}(\partial)$.

The short exact sequence

$$0 \to C_k \to C' \to \mathscr{S}_v(\partial) \otimes \langle e'_{w(\beta)} \rangle \to 0$$

induces the sequence

$$0 \to \mathrm{H}_0(C_k) \to \mathrm{H}_0(C') \to \mathrm{H}_0(\mathscr{S}_v(\partial)) \otimes \langle e'_{w(\beta)} \rangle \to 0.$$

By induction on the filtration in Corollary 4.4, the support of $\mathrm{H}_0(C_k)$ is in the union of the $X(\psi_t^{-1})$, and hence does not contain $X(w^{-1})$. So localizing at P, we get an isomorphism

$$\mathrm{H}_0(C')_P \cong \mathrm{H}_0(\mathscr{S}_v(\partial))_P \otimes \langle e'_{w(\beta)} \rangle.$$

So we can restrict this isomorphism to $X(w^{-1})$. Localizing (4.6) at P and then restricting $X(w^{-1})$, we get a surjection

$$\mathrm{H}_0(\mathscr{S}_v(\partial))_P \otimes \langle e'_{w(\beta)} \rangle \to \mathrm{H}_0(C)_P \to 0.$$

By induction, the first term has length 1 over the generic point of $X(w^{-1})$, so the length of $\mathrm{H}_0(C)_P$ is either 0 or 1.

The idea for using this partial information is to carry our situation to the complete flag variety and to use its K-theory to show that $\mathrm{length}(\mathrm{H}_0(C)_P) - \mathrm{length}(\mathrm{H}_1(C)_P) = 1$. Some more analysis of the K-theory gives us the other statements which allow us to complete the induction step. □

Corollary 4.7 *Let X be an equidimensional Cohen–Macaulay variety, and let $\partial\colon E \to F$ be a map of vector bundles on X. Let $E_1 \subset \cdots \subset E_n = E$ and $F^{-n} \subset \cdots \subset F^{-1} = F$ be split flags of subbundles. Let $w \in \Sigma_n$ be a permutation, and define the degeneracy locus*

$$D_w(\partial) = \{x \in X \mid \mathrm{rank}(\partial_x\colon E_p(x) \to F/F^{-q-1}(x)) \le r_w(p,q)\},$$

where the ideal sheaf of $D_w(\partial)$ is locally generated by the minors given by the rank conditions. Suppose that $D_w(\partial)$ has codimension $\ell(w)$.

(a) *The Schubert complex $\mathscr{S}_w(\partial)$ is acyclic, and the support of its cokernel $\mathcal{L}$ is $D_w(\partial)$.*

(b) *The degeneracy locus $D_w(\partial)$ is Cohen–Macaulay.*

(c) *The restriction of $\mathcal{L}$ to $D_w(\partial)$ is a line bundle outside of a certain codimension 2 subset.*

4.3 Examples.

Here is a combinatorial description of the differentials in the Schubert complex for a flagged isomorphism. We will work with just the tensor product complex $\bigotimes_{k=1}^{n-1} \mathrm{D}^{r_k(w)}(F)$. Then the basis elements of its terms are row-strict labelings. The differential sends such a labeling to the signed sum of all possible ways to change a single unmarked letter to the corresponding marked letter. If T' is obtained from T by marking a letter in the ith row, then the sign on T' is $(-1)^n$, where n is the number of unmarked letters of T in the first $i-1$ rows.

Example 4.8 Consider the permutation $w = 1423$. Then $D(w) = \{(2,2),(2,3)\}$, and we denote the generic map by $e_1 \mapsto ae'_1 + be'_2 + ce'_3$ and $e_2 \mapsto de'_1 + ee'_2 + fe'_3$ (the images of e_3 and e_4 are irrelevant, and it is also irrelevant to map to e'_4) instead of a flagged isomorphism. The cokernel M is Cohen–Macaulay of codimension 2 over $A = K[a,b,c,d,e,f]$:

$$0 \to A^3 \xrightarrow{\begin{pmatrix} e & b & 0 \\ 0 & e & b \\ d & a & 0 \\ 0 & d & a \\ 0 & f & c \\ f & c & 0 \end{pmatrix}} A^6 \xrightarrow{\begin{pmatrix} d & a & -e & -b & 0 & 0 \\ 0 & 0 & -f & -c & a & d \\ -f & -c & 0 & 0 & b & e \end{pmatrix}} A^3 \to M \to 0$$

Example 4.9 Consider the permutation $w = 2413$. Then $D(w) = \{(1,1),(2,1),(2,3)\}$, and denote the generic matrix by $e_1 \mapsto ae'_1 + be'_2 + ce'_3$ and $e_2 \mapsto de'_1 + ee'_2 + fe'_3$ (the images of e_3 and e_4 are irrelevant, and it is also irrelevant to map to e'_4). The cokernel M is Cohen–Macaulay of codimension 3 over $A = K[a,b,c,d,e,f]$:

$$0 \to A^2 \xrightarrow{\begin{pmatrix} -d & -a \\ -e & -b \\ -f & -c \\ 0 & -d \\ a & 0 \\ 0 & a \end{pmatrix}} A^6 \xrightarrow{\begin{pmatrix} 0 & 0 & 0 & a & 0 & d \\ e & -d & 0 & b & 0 & e \\ f & 0 & -d & c & 0 & f \\ a & 0 & 0 & 0 & d & a \\ 0 & a & 0 & 0 & e & b \\ 0 & 0 & a & 0 & f & c \end{pmatrix}} A^6 \xrightarrow{\begin{pmatrix} -b & a & 0 & -e & d & 0 \\ -c & 0 & a & -f & 0 & d \end{pmatrix}} A^2 \to M \to 0$$

5 Degeneracy loci.

5.1 A formula of Fulton.

Suppose we are given a map $\partial\colon E \to F$ of vector bundles of rank n on a variety X, together with a flag of subbundles $E_1 \subset E_2 \subset \cdots \subset E_n = E$ and a flag of quotient bundles $F = F_n \twoheadrightarrow F_{n-1} \twoheadrightarrow \cdots \twoheadrightarrow F_1$ such that $\operatorname{rank} E_i = \operatorname{rank} F_i = i$. We assume that the quotient flags E_i/E_{i+1} are locally free. For a permutation w, define

$$D_w(\partial) = \{x \in X \mid \operatorname{rank}(\partial_x\colon E_p(x) \to F_q(x)) \le r_w(p,q)\}.$$

Then $\operatorname{codim} D_w(\partial) \le \ell(w)$. Define Chern classes $x_i = -c_1(E_i/E_{i-1})$ and $y_i = -c_1(\ker(F_i \twoheadrightarrow F_{i-1}))$.

Theorem 5.1 (Fulton) *Suppose that X is an equidimensional Cohen–Macaulay variety and $D_w(\partial)$ has codimension $\ell(w)$. Then the identity*

$$[D_w(\partial)] = \mathfrak{S}_w(x,y) \cap [X]$$

holds in the Chow group $\mathrm{A}_{\dim(D_w(\partial))}(X)$.

See [F1, §8] for a more general statement which does not enforce a codimension requirement on $D_w(\partial)$ or assume that X is Cohen–Macaulay.

We will only deal with the case when X is smooth. The general case can be reduced to this case using a "universal construction" (see [F2, Chapter 14]). So suppose that X is smooth. Let $\mathrm{A}_*(X) =$

$\bigoplus_{k\geq 0} \mathrm{A}_k(X)$ be the direct sum of its Chow groups, and $\mathrm{Gr}\,\mathrm{K}(X)$ be the associated graded of the topological filtration of its Grothendieck group (see [F2, Example 15.1.5]). Let $\varphi\colon \mathrm{A}_*(X) \to \mathrm{Gr}\,\mathrm{K}(X)$ be the functorial morphism of graded rings which for a subvariety $V \subseteq X$ sends $[V]$ to $[\mathcal{O}_V]$. If $\mathcal{F}$ is a coherent sheaf whose support has dimension at most k, then we have $\varphi(Z_k(\mathcal{F})) = [\mathcal{F}]$ where

$$Z_k(\mathcal{F}) = \sum_{\dim V = k} m_V(\mathcal{F})[V],$$

and $m_V(\mathcal{F})$ is the length of the stalk of $\mathcal{F}$ at the generic point of V. In order to state the connection between the Schubert complex and Fulton's formula, we will need the following lemma which was observed in [Pra, Appendix 6].

Lemma 5.2 *Let D be an irreducible closed subvariety of a smooth variety X. Let $C_\bullet$ be a finite complex of vector bundles on X and let $P \in \mathrm{A}^{\operatorname{codim} D}(X)$. If*

$$\operatorname{supp} C_\bullet = X \setminus \{x \in X \mid (C_\bullet)|_x \text{ is an exact complex}\}$$

is contained in D, and $\varphi(P \cap [X]) = [C_\bullet]$, then $c[D] = P \cap [X]$ for some $c \in \mathbf{Q}$.

We will use Lemma 5.2 with $D = D_w(\partial)$, $C_\bullet = \mathscr{S}_w(\partial)$, and $P = \mathfrak{S}_w(x,y)$ using the notation from the beginning of this section. We know that $\operatorname{supp} C_\bullet \subseteq D$ and that the codimension of D is $\ell(w) = \deg P$ by Corollary 4.7. So we need to check that $\varphi(P \cap [X]) = [C_\bullet]$.

For a line bundle L corresponding to an irreducible divisor D, we have $c_1(L) \cap [X] = [D]$, and hence

$$\varphi(c_1(L) \cap [X]) = 1 - [L^\vee]$$

by the short exact sequence

$$0 \to \mathscr{L}(-D) \to \mathcal{O}_X \to \mathcal{O}_D \to 0.$$

So the same formula holds for all L by linearity, and $\varphi(x_i) = 1-[E_i/E_{i-1}]$ while $\varphi(y_j) = 1-[\ker(F_j \twoheadrightarrow F_{j-1})]$. Let a and b be a new set of variables. We have $\mathfrak{S}_w(a,b) = \sum_{u \leq_W w} \mathfrak{S}_u(a)\mathfrak{S}_{uw^{-1}}(-b)$. Doing the transformation $a_i \mapsto x_i - 1$ and $b_j \mapsto y_j - 1$, we get $\varphi(\mathfrak{S}_w(a,b)) = \sum_{u \leq_W w} (-1)^{\ell(u)} \mathfrak{S}_u(E)\mathfrak{S}_{uw^{-1}}(F)$. By Proposition 4.1, this sum is $[\mathscr{S}_w(\partial)]$ (the change from uw^{-1} to wu^{-1} is a consequence of the fact that F_1 in Proposition 4.1 contains only odd elements). So it is enough to show that the substitution $a_i \mapsto a_i + 1$, $b_j \mapsto b_j + 1$ leaves the expression $\mathfrak{S}_w(a,b)$ invariant. This is clearly true for $\mathfrak{S}_{w_0}(x,y) = \prod_{i+j \leq n}(x_i - y_j)$, and holds for an arbitrary permutation because the divided difference operators (see (2.1)) applied to a substitution invariant function yield a substitution invariant function.

So it remains to show that the constant given by Lemma 5.2 is 1. This follows from Corollary 4.7(c).

5.2 Some remarks.

First we point out that the above can be applied to partial flags, but we have kept to complete flags for simplicity of notation.

A permutation $w \in \Sigma_n$ is **Grassmannian** if it has at most one descent, i.e., there exists r such that $w(1) < w(2) < \cdots < w(r) > w(r+1) < \cdots < w(n)$. Suppose that w is **bigrassmannian**, which means that w and w^{-1} are Grassmannian permutations. This is equivalent to saying that $D(w)$ is a rectangle. In this case, the double Schubert polynomial $\mathfrak{S}_w(x,y)$ is a super Schur polynomial for the partition $D(w)$. The degeneracy locus $D_w(\partial)$ can then be described by a single rank condition between the map $\partial E \to F$, so the degeneracy locus formula of Fulton specializes to the Thom–Porteous formula mentioned in the introduction. So in principle, the action of $\mathfrak{b}$ on $\mathscr{S}_w(\partial)$ should extend to an action of a general linear superalgebra, but it is not clear why this should be true without appealing to Schur polynomials.

We have seen that the modules which are the cokernels of generic Schubert complexes have linear minimal free resolutions. These modules can then be thought of as a sort of "linear approximation" to the ideal which defines the matrix Schubert varieties, which in general have rich and complicated minimal free resolutions.

Acknowledgements

The author thanks Jerzy Weyman for helpful discussions and for pointing out the connection between Schur complexes and the Thom–Porteous formula which was the starting point for this work.

References

[ABW] Kaan Akin, David A. Buchsbaum, and Jerzy Weyman, Schur functors and Schur complexes, *Adv. in Math.* **44** (1982), no. 3, 207–278.

[BE] David A. Buchsbaum and David Eisenbud, Generic free resolutions and a family of generically perfect ideals, *Advances in Math.* **18** (1975), no. 3, 245–301.

[Eis] David Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Graduate Texts in Mathematics **150**, Springer-Verlag, 1995.

[FGRS] Sergey Fomin, Curtis Greene, Victor Reiner, and Mark Shimozono, Balanced labellings and Schubert polynomials, *Europ. J. Combinatorics* **18** (1997), 373–389.

[F1] William Fulton, Schubert polynomials, degeneracy loci, and determinantal formulas, *Duke Math. J.* **65** (1992), no. 3, 381–420.

[F2] ——— , *Intersection Theory*, second edition, Springer-Verlag, Berlin, 1998.

[KP] Witold Kraśkiewicz and Piotr Pragacz, Schubert functors and Schubert polynomials, *Europ. J. Combinatorics* **25** (2004), 1327–1344.

[Man] Laurent Manivel, *Symmetric Functions, Schubert Polynomials and Degeneracy Loci*, translated by John R. Swallow, SMF/AMS Texts and Monographs **6**, American Mathematical Society (1998).

[MS] Ezra Miller and Bernd Sturmfels, *Combinatorial Commutative Algebra*, Graduate Texts in Mathematics **227**, Springer-Verlag, 2005.

[Pra] Piotr Pragacz, Symmetric polynomials and divided differences in formulas of intersection theory, *Parameter spaces* (Warsaw, 1994), 125–177, Banach Center Publ., 36, Polish Acad. Sci., Warsaw, 1996, `arXiv:alg-geom/9605014`.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 904–915

Counting RNA pseudoknotted structures (extended abstract)

Cédric Saule[1,4], Mireille Régnier[4,2], Jean-Marc Steyaert[2,4], Alain Denise[1,3,4]

[1] *LRI, Université Paris-Sud and CNRS, bât. 490, 91405 Orsay cedex, France*
[2] *LIX, Ecole Polytechnique and CNRS, 91128 Palaiseau cedex, France*
[3] *IGM, Université Paris-Sud and CNRS, bât. 400, 91405 Orsay cedex, France*
[4] *INRIA Saclay, Parc Orsay Université, 4 rue Jacques Monod, 91893 Orsay cedex, France*

Abstract. In 2004, Condon and coauthors gave a hierarchical classification of exact RNA structure prediction algorithms according to the generality of structure classes that they handle. We complete this classification by adding two recent prediction algorithms. More importantly, we precisely quantify the hierarchy by giving closed or asymptotic formulas for the theoretical number of structures of given size n in all the classes but one. This allows to assess the tradeoff between the expressiveness and the computational complexity of RNA structure prediction algorithms.

Résumé. En 2004, Condon et ses coauteurs ont défini une classification des algorithmes exacts de prédiction de structure d'ARN, selon le degré de généralité des classes de structures qu'ils sont capables de prédire. Nous complétons cette classification en y ajoutant deux algorithmes récents. Chose plus importante, nous quantifions la hiérarchie des algorithmes, en donnant des formules closes ou asymptotiques pour le nombre théorique de structures de taille donnée n dans chacune des classes, sauf une. Ceci fournit un moyen d'évaluer, pour chaque algorithme, le compromis entre son degré de généralité et sa complexité.

Keywords: bioinformatics, RNA structures, pseudoknots, enumeration, bijective combinatorics

1 Introduction

In bioinformatics, the RNA structure prediction problem consists, given a RNA sequence, in finding a conformation that the molecule is likely to take in the cell. In [3], Condon and coauthors classified RNA structure prediction algorithms according to the inclusion relations between their *classes of structures*. The class of structures of a given algorithm is the set of structures that can be, in theory, returned by the algorithm. Condon *et al.* focused only on *exact* algorithms, that is algorithms that guarantee to give an optimal solution to the structure prediction problem, stated as an optimisation problem. They considered the class of pseudoknot-free structures [13, 25] (PKF) , and the following classes for pseudoknotted structures: Lyngsø and Pedersen (L&P) [11], Dirks and Pierce (D&P) [4], Akutsu and Uemura (A&U) [1, 20], and Rivas and Eddy (R&E) [16]. They notably proved the following inclusion relations: $PKF \subset L\&P \subset D\&P \subset A\&U \subset R\&E$. Since then, two other exact prediction algorithms have been developed, involving new classes: Reeder and Giegerich (R&G) [15] and Cao and Chen (C&C) [2] algorithms.

In this paper, we aim to quantify the tradeoff between the computational complexity and the expressiveness of all these algorithms. For this purpose, we compare them from the double point of view of their computational complexity and the cardinality of their class of structures, for a given size n. And we give closed or asymptotic formulas for the theoretical number of structures of given size n except for the class $R\&E$. More precisely, we establish that, except for the $L\&P$ class whose asymptotic formula is simpler, the number of structures of size n is, asymptotically, $\frac{\alpha}{2\sqrt{\pi}\,n^{3/2}}\omega^n$, where α and ω are two constants which depend of the class. Additionally, we place the two new classes, R&G and C&C, in Condon *et al*'s hierarchy. The following table summarizes our results. We indicate by "*" the classes that had not been enumerated before. The class "All" denotes the whole set of pseudoknotted structures. The row "Compl" gives the complexity of each algorithm.

Class	asympt.	α	ω	Compl.	Remark
PKF	$\frac{\alpha}{2\sqrt{\pi}\,n^{3/2}}\omega^n$	2	4	$\mathcal{O}(n^3)$	Catalan numbers
L&P *	$\frac{1}{2}\omega^n$	-	4	$\mathcal{O}(n^5)$	Closed formula
C&C *	$\frac{\alpha}{2\sqrt{\pi}\,n^{3/2}}\omega^n$	1,6651	5,857	$\mathcal{O}(n^6)$	
R&G *	$\frac{\alpha}{2\sqrt{\pi}\,n^{3/2}}\omega^n$	0,1651	6,576	$\mathcal{O}(n^4)$	
D&P *	$\frac{\alpha}{2\sqrt{\pi}\,n^{3/2}}\omega^n$	0,7535	7,315	$\mathcal{O}(n^5)$	
A&U *	$\frac{\alpha}{2\sqrt{\pi}\,n^{3/2}}\omega^n$	0,6575	7,547	$\mathcal{O}(n^5)$	
R&E	open	-	-	$\mathcal{O}(n^6)$	
All	$\sqrt{2}\cdot 2^n\cdot\left(\frac{n}{e}\right)^n$	-	-	NPC	Involutions with no fixed points

A number of works have been done on combinatorial enumeration of RNA structures without pseudoknots, see e.g. [24, 21, 7, 12, 10] or, more recently, with pseudoknots, as in [22, 17, 8, 9] for instance. Our purpose is different, as our classes of structures are not defined *per se*, but correspond to given exact prediction algorithms.

The paper is organised as follows. In Section 2, we give some notation and definitions. In Section 3, we present a bijection between the L&P class and a class of planar maps, leading to a closed formula for the L&P class. In Section 4, we establish that each of the classes D&P, A&U, R&G, C&C, and L&P can be encoded by a context-free language. For each of them, we derive an equation for the generating function, leading to an asymptotic formula for the number of structures of size n. In Section 5, we conclude by giving some remarks on the expressiveness of the structure prediction algorithms compared to their complexity.

2 Definitions and notation.

A RNA secondary structure (possibly with pseudoknots) is given by a sequence of integers $(1, 2, \ldots, n)$ and a list of pairs (i, j), called *basepairs* or *arcs*, where $i < j$ and each number in $\{1, 2, \ldots, n\}$ appears exactly in one pair. Such a structure can be represented as in Figure 1, where each basepair (i, j) is represented by an edge between i and j. In real RNA structures there are unpaired bases, but we do not consider them.

Definition 1 (Crossing arcs) *Let (i, j) and (k, l) two arcs such that $i < k$. We say that (i, j) and (k, l) are* crossing *if $i < k < j < l$.*

Definition 2 (Crossing graph) *The* crossing graph *of an RNA structure is a graph G defined as follows: the vertices of G are the arcs of the structure, and two vertices of G are connected by an edge if and only if their two corresponding arcs are crossing.*

Definition 3 (Pseudoknot) *A* pseudoknot *is a set of arcs that is not a singleton and that corresponds to a maximal connected component in the crossing graph.*

Definition 4 (Simple pseudoknot [1]) *A pseudoknot P is* simple *if there exist two numbers j_1 and j_2, with $j_1 < j_2$, such that: (i) each arc (i, j) in P satisfies either $i < j_1 < j \leq j_2$ or $j_1 \leq i < j_2 < j$, (ii) and if two arcs (i, j) and (i', j') satisfy $i < i' < j_1$ or $j_1 \leq i < i'$, then $j > j'$.*

The first property ensures that, for each arc of P, one of its ends exactly is between j_1 and j_2. And the arcs are divided in two sets: those having their other end smaller than j_1, and those having their other end greater than j_2. We call these two sets, respectively, the *left part* and the *right part* of the pseudoknot. The second property of the definition ensures that two arcs in the same set cannot intersect each other. Figure 1 shows a simple pseudoknot.

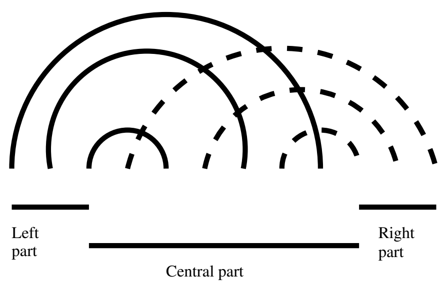


Fig. 1: A pseudoknot given by the sequence $(1, 2, \ldots, 12)$ and the arcs $(1, 9)$, $(2, 7)$, $(3, 5)$, $(4, 12)$, $(6, 11)$, $(8, 10)$. This pseudoknot is simple, with $j_1 = 4$ and $j_2 = 9$.

Definition 5 (H-type Pseudoknot) *A* H-type pseudoknot *is a simple pseudoknot having the following additional property: each arc in one of the two above sets crosses all the arcs of the other set.*

3 A bijection between the L&P structures and a class of planar maps.

The Lyngsø-Pedersen (L&P) class is the simplest class of pseudoknotted structures. According to [11] and [3], a structure is in the L&P class if and only if it contains either no pseudoknot or a unique H-type pseudoknot, and this pseudoknot is not embedded under any arc. Between any two consecutive ends of the arcs of the pseudoknots, there can be a nested structure. Theorem 1, and its straightforward Corollary 1, give the closed formula and the asymptotic formula for the number of such structures, respectively.

Theorem 1 *The number of* L&P structures with n arcs is:

$$LP(n) = \frac{1}{2} \cdot 4^n - \binom{2n+1}{n} + \binom{2n-1}{n-1} + \frac{1}{n+1}\binom{2n}{n}.$$

Corollary 1

$$LP(n) \sim \frac{1}{2} \cdot 4^n.$$

Proof of Theorem 1:

The proof is bijective: we establish a bijection between the set of L&P structures of any size n and the set of rooted isthmusless planar maps with n edges and one or two vertices. The first two terms of the formula count the number of such maps with two vertices [18, 23], while the last term, a Catalan number, counts the number of such maps with one vertex [19]. Hence the theorem.

A *planar map* is a proper embedding of a connected planar graph. It is said *isthmusless* if the deletion of any edge does not split the graph. A *rooted* planar map is a planar map where a vertex and an edge adjacent to it are distinguished. Any planar map with n edges can be represented by two permutations σ and τ on $\{+1, -1, +2, -2, \dots, +(n-1), -(n-1), +n, -n\}$, in the following way: The edges of the map are numbered from 1 to n. Then, for any edge i, one labels its extremities with $+i$ and $-i$, respectively. By convention, the root edge is labelled with $+1$ and -1, in such a way that -1 labels the extremity adjacent to the root vertex. Now, the two permutations are as follows:

- the permutation σ is an involution without fixed points that represents the edges of the map. Each cycle of σ is of size two and contains both ends of one edge: $\sigma = (+1, -1), (+2, -2), \dots, (+n, -n)$.

- the permutation τ has as many cycles as vertices in the map. Each cycle is given by the sequence of labellings around the corresponding vertex, clockwise.

Let us consider a L&P structure S with n edges, and let us label the left extremities of its arcs with $+1, +2, \dots, +n$ from left to right, and give to each right foot the label $-i$ if the corresponding left foot has label $+i$. Let $w = [w_1, w_2, \dots w_{2n}]$ be the sequence of labels of S, from left to right. From any w we can now construct two permutations σ and τ that represent an isthmusless rooted planar map with one or two vertices. Regarding σ, we just set $\sigma = (+1, -1) \dots (+n, -n)$.

Let us first consider the simple case where there is no crossing in the structure. It is known for a long time that such nested structures are counted by Catalan numbers. This can be established, for example by a folkloric bijection with planar maps having one vertex, by setting σ as above, and $\tau = (w)$.

Now suppose that there is a pseudoknot in the structure, and let us present a bijection between the set of such structures and the set of rooted ithmusless planar maps with two vertices. Start from w. Since τ must have two cycles, we have to split w in two parts that will be the two cycles. Let us define the left set (resp. the right set) of arcs of the pseudoknot, respectively, as the set of arcs whose left (resp. right) extremities are in the left (resp. right) part of the pseudoknot, where left and right parts are defined as in Section 2. There are two cases:

Case 1. There is only one arc in the right set. In this case, let ℓ be the position of the first right extremity of an arc in the left set. We cut w between positions $\ell - 1$ and ℓ. Each part corresponds to a cycle of τ: $\tau = (w_1, \dots, w_{\ell-1})(w_\ell, \dots, w_{2n})$. See Figure 2 for an illustration.

Case 2. There are at least two arcs in the right set. We cut w just before the first right extremity of an arc in the right set. See Figure 3.

Let us show that, in both cases, the resulting map is planar and isthmusless. At first, remark that if the map is not planar or has an isthmus, necessarily it comes from arcs that are involved in the pseudoknot. Indeed, by construction, non crossing arcs in the structure give non crossing loops in the map. So, without loss of generality, we can consider only structures where all the arcs are involved in the pseudoknot. Consider such a structure with n arcs. In the case 1, we have $w = [+1, +2, \dots, +(n-1), +n, -(n-$

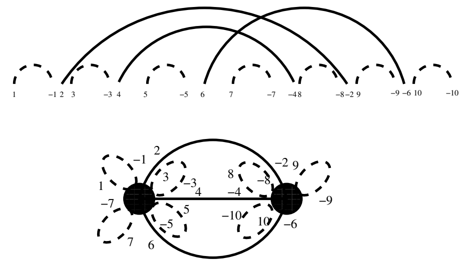


Fig. 2: Top, a L&P structure corresponding to case 1. Bottom, the corresponding planar map. Arcs not involved in the pseudoknot are drawn in dotted lines.

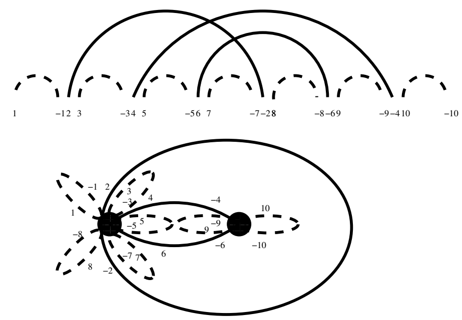


Fig. 3: Top, a L&P structure corresponding to case 2. Bottom, the corresponding planar map. Arcs not involved in the pseudoknot are drawn in dotted lines.

$1), -(n-2), \ldots, -1, -n]$, hence $\tau = (+1, +2, \ldots, +n)(-(n-1), -(n-2), \ldots, -1, -n)$. Clearly, this gives a planar map, since the two cycles of τ are in opposite order. And there is no isthmus because all edges go from one vertex to the other. In the case 2, we have $w = [+1, +2, \ldots, +(\ell-1), +\ell, +(\ell+1), \ldots, +n, -(\ell-1), \ldots, -2, -1, -n, -(n-1), \ldots, -\ell]$, hence $\tau = (+1, +2, \ldots, +(\ell-1), +\ell, +(\ell+1), \ldots, +n, -(\ell-1), \ldots, -2, -1)(-n, -(n-1), \ldots, -\ell)$. Again, this gives a planar map: edges $1, 2, \ldots, \ell-1$ are nested loops, and edges $\ell, \ldots, n$ go from one vertex to the other, without any crossing. And there is no isthmus because the number of edges going from one vertex to the other, $n-\ell+1$, is greater or equal to 2.

Now let us present the converse transformation. Consider an isthmusless rooted planar map with two vertices, given by $\sigma = (+1, -1), (+2, -2), \ldots, (+n, -n)$ and τ having two cycles. We aim to construct the sequence w that represents the corresponding pseudoknotted structure. Let us consider the cycle of τ which contains 1, and write it in such a way that it begins with 1. Let us call u this sequence of labels. This gives the first part of the sequence w. We are now searching for the second part of w, that is the sequence v such that $uv = w$. For that purpose, consider the set of *isolated labels*, that is the labels in u that have not their opposite label in u. We have the two following cases:

Case 1. There is no pair $(+i, -i)$ in u such that the isolated labels are located between $+i$ and $-i$. Let $+j$ the penultimate isolated label in u. Write the second cycle of τ in such a way that it begins with $-j$. This gives v, and there is exactly one edge in the second part of the pseudoknot.

Case 2. There is a pair of labels $(+i, -i)$ in u such that all isolated labels are located between $+i$ and $-i$. Let $+j$ the last isolated label in u. Write the second cycle of τ in such a way that it begins with $-j$. This gives v. In this case, there are at least two edges in the second part of the pseudoknot. □

4 Asymptotic enumeration of pseudoknotted structures.

4.1 A context-free encoding for simple and H-type pseudoknots

As will be seen farther, all the classes that are involved in exact prediction algorithms but one involve either H-type pseudoknots or simple pseudoknots. The only exception is the R&E class. Here we define a transformation that allow to encode any class of pseudoknotted structures where all pseudoknots are simple by a context-free language.

Let us first recall some definitions. Let L be a language on a given alphabet A, and $w = w_1 w_2 \dots w_n$ a word of L, where the w_i's are the letters of w. A word v is a *subword* of w if $v = w_{i_1} w_{i_2} \dots w_{i_k}$, where $1 \leq i_1 < i_2 < \dots < i_k \leq n$. The *projection* of w onto an alphabet $A' \in A$ is the subword w' obtained by erasing in w all letters that do not belong to A'. The projection of L onto A' is the set of projections of the words of L onto A'. Finally, let us recall that the Dyck language on any two-letter alphabet $\{d, \bar{d}\}$ is the language of balanced parentheses strings, where d and $\bar{d}$ stand, respectively, for opening and closing parentheses. Now we can state twe two following straightforward lemmas:

Lemma 1 *Any class of pseudoknotted structures where all pseudoknots are simple can be encoded by the words of a language L on the alphabet $\{d, \bar{d}, x, \bar{x}, y, \bar{y}\}$ where (i) d and $\bar{d}$ encode, respectively, the left and right ends of arcs that are not involved in pseudoknots; (ii) x and $\bar{x}$ encode, respectively, the left and right ends of arcs that are involved in the left parts of pseudoknots; (iii) y and $\bar{y}$ encode, respectively, the left and right ends of arcs that are involved in the right parts of pseudoknots. Additionally, the projection of the language to the alphabet $\{d, \bar{d}\}$ (resp. $\{x, \bar{x}\}$, $\{y, \bar{y}\}$) is a sublanguage of the Dyck language on the same alphabet.*

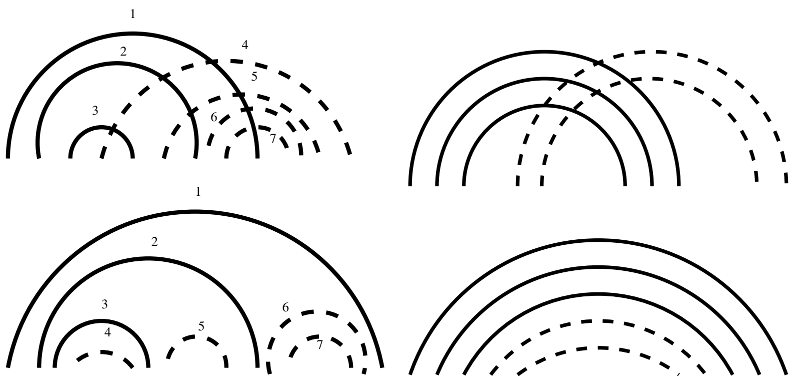
Lemma 2 *Let S be a pseudoknotted structure, and w be the word on $\{d, \bar{d}, x, \bar{x}, y, \bar{y}\}$ that encodes S. Then every simple pseudoknot in S is encoded by a subword v of w, such that $v = x^n\, y^{m_1}\, \bar{x}^{n_1}\, y^{m_2}\, \bar{x}^{n_2}\, \dots\, y^{m_k}\, \bar{x}^{n_k}\, \bar{y}^m$, where $n_1 + n_2 + \dots + n_k = n$ and $m_1 + m_2 + \dots + m_k = m$.*

Remark that a H-type pseudoknots is a simple pseudoknot where $k = 1$. Thus every H-type pseudoknot in S is encoded by a subword $v = x^n\, y^m\, \bar{x}^n\, \bar{y}^m$. Finally, the following Proposition gives a way to encode any pseudoknotted structure where all pseudoknots are simple by a subset of the Dyck language with three kinds of pairs of parentheses, that is on the alphabet $\{d, \bar{d}, x, \bar{x}, y, \bar{y}\}$.

Proposition 1 *Let S be a pseudoknotted structure, and w be the word on $\{d, \bar{d}, x, \bar{x}, y, \bar{y}\}$ that encodes S. Then w can be encoded by a word on the alphabet $\{d, \bar{d}, x, \bar{x}, y, \bar{y}\} \cup \{p, \bar{p}\}$ where every subword $v = x^n\, y^{m_1}\, \bar{x}^{n_1}\, y^{m_2}\, \bar{x}^{n_2}\, \dots\, y^{m_k}\, \bar{x}^{n_k}\, \bar{y}^m$, corresponding to a H-type pseudoknot is replaced with $v' = px^{n-1}\, y^{m_1} \bar{y}^{m_1}\, \bar{x}^{n_1}\, y^{m_2} \bar{y}^{m_2}\, \bar{x}^{n_2}\, \dots\, y^{m_k} \bar{y}^{m_k}\, \bar{x}^{n_k - 1} \bar{p}$.*

In particular , every subword $v = x^n\, y^m\, \bar{x}^n\, \bar{y}^m$ corresponding to a simple pseudoknot is replaced with $v' = px^{n-1}\, y^m\, \bar{y}^m\, \bar{x}^{n-1} \bar{p}$.

Proof (sketch): The proof is straightforward, as there is an immediate one-to-one correspondance between the two kinds of words below. The transformation is illustrated in Figure 4(a) and Figure 4(b), respectively, for simple pseudoknots and for the particular case of H-type pseudoknots. □



(a) Simple pseudoknot (b) H-type pseudoknot

Fig. 4: Top: two pseudoknots. Bottom: their encodings by Proposition 1.

4.2 Asymptotic results.

For each of the D&P, A&U, R&G, and C&C classes, we give an asymptotic equivalent for the number of structures of size n. In each case, the proof is in three steps: *(i)* We design an unambiguous context-free grammar which generates the language that encodes the considered structures, according to Proposition 1. *(ii)* From the grammar, we deduce an algebraic equation satisfied by the ordinary generating function (o.g.f.) of the language. *(iii)* From this equation, we compute an asymptotic formula for the number of structures of size n. For any class $X\&Y$, we write $X\&Y(n)$ for its number of structures having n arcs.

4.2.1 The Akutsu-Uemura class (A&U).

Following [1, 3], the A&U structures are composed of non crossing edges and of any number of simple pseudoknots (Fig. 1). As these pseudoknot can embed other substructures which can be pseudoknotted in turn, they are said to be *recursive* [1].

Theorem 2

$$A\&U(n) \;=\; \frac{\alpha_1}{2\sqrt{\pi}}{\omega_1}^n n^{-3/2}(1+O(1/n)),$$

where $\alpha_1 = 0.6575407644...$, $\omega_1 = 7.547308334...$, *are algebraic constants.*

Proof: Let $L_{A\&U}$ be the language that encodes the A&U class, according to Proposition 1. The following unambiguous context-free grammar generates $L_{A\&U}$:

$$S \to dS\bar{d}S|P \; ; \; P \to pSX\bar{p}S|\epsilon \; ; \; X \to xSX\bar{x}SY|yYS\bar{y}S \; ; \; Y \to ySY\bar{y}S|\epsilon$$

The two rules in the first line allow to generate non crossing arcs and to put pseudoknots anywhere. The other rules generate words that correspond to the code for a simple pseudoknot.

Given the grammar, we obtain the set of recursive equations for the o.g.f. of the various sets defined in the 1-to-1 encoding. Letting the formal symbol z denote an arc, we thus have through a straightforward translation:

$$S(z) = zS^2(z){+}P(z)\;;\; P(z) = zS^2(z)X(z){+}1\;;\; X(z) = zS^2(z)Y(z)(X(z){+}1)\;;\; Y(z) = zS^2(z)Y(z){+}1$$

By iterated bottom-up substitutions, we ultimately get that the o.g.f. $S(z)$ is solution of the algebraic equation

$$F(z,S) \;=\; z^2S^4 - 2zS^3 + zS^2 + S - 1 = 0, \tag{1}$$

from which we can derive the number of structures of size n.

For this proof we present in some details the main steps of the computations that have to be performed in order to get the asymptotics for an o.g.f. given by the algebraic implicit equation $F(z,S) = 0$ satisfied by the o.g.f. $S(z)$.

Since $\partial F/\partial z|_{z=0,S=1} = 1$ is defined and $\partial F/\partial S|_{z=0,S=1} = 1$ is non vanishing, $z = 0$ is not a singular point for S; by the implicit function theorem, $S(z)$ exists as a regular function in a circular neighborhood of $z = 0$ until $\partial F/\partial S$ vanishes. The radius of convergence ρ_1 of the o.g.f. $S(z)$ is thus a solution of the system $\{F(z,S) = 0, \partial F/\partial S(z,S) = 0\}$. At such a point the local holomorphic solution $z = \zeta(S)$ is no longer invertible, which implies that this point is a singular point for the o.g.f. $S(z)$. The Darboux method allows to get accurate information and precise asymptotics for the Taylor expansion of $S(z)$.

Let $(z = \rho_1, S = \sigma_1)$ be the point of the Riemann surface of the solution located on the fold issued from $(z = 0, S = 1)$, which is closest to $(z = 0, S = 1)$ and for which $\partial F/\partial S = 0$. This point is usually unique and located on the positive real axis, since the o.g.f. is indeed a function of z with all coefficients being positive. At this point, the local expansion of z with respect to S writes:

$$z = \rho_1 + \frac{1}{2}\frac{d^2z}{dS^2}(S-\sigma_1)^2 + \frac{1}{3!}\frac{d^3z}{dS^3}(S-\sigma_1)^3 + \ldots, \tag{2}$$

since the first derivative, $\frac{dz}{dS} = -\frac{\partial F/\partial S}{\partial F/\partial z}$ vanishes at $(z = \rho_1, S = \sigma_1)$ and the second derivative $\frac{d^2z}{dS^2} = -\frac{\partial^2 F/\partial S^2}{\partial F/\partial z}$ does not.

Hence taking the square root of the previous equation we get the Taylor expansion at $(z = \rho_1, S = \sigma_1)$:

$$\sqrt{1 - z/\rho_1} = \beta_1(S-\sigma_1) + \beta_2(S-\sigma_1)^2 + \ldots, \tag{3}$$

with $\beta_1 = -\sqrt{\frac{1}{2}\frac{\partial^2 F/\partial S^2}{\partial F/\partial z}}$, which can now be inverted locally giving:

$$S = \sigma_1 - \sqrt{\frac{2\rho_1 \partial F/\partial z|_{z=\rho_1,S=\sigma_1}}{\partial^2 F/\partial S^2|_{z=\rho_1,S=\sigma_1}}}\sqrt{1 - z/\rho_1} + O(1 - z/\rho_1). \tag{4}$$

The expansion can be calculated at any order, so that we obtain for the coefficients $A\&U(n)$ an infinite asymptotic development whose dominant term is given by the first square root in the previous expansion, since it is well-known that $[z^n] - \sqrt{1 - z/\rho} = \frac{1}{2\sqrt{\pi}}\rho^{-n}n^{-3/2}(1 + O(1/n))$:

$$[z^n]S(z) = \sqrt{\frac{2\rho_1 \partial F/\partial z|_{z=\rho_1,S=\sigma_1}}{\partial^2 F/\partial S^2|_{z=\rho_1,S=\sigma_1}}}\frac{1}{2\sqrt{\pi}}\rho_1^{-n}n^{-3/2}(1 + O(1/n)). \tag{5}$$

We thus get the general form of the solution, as stated in the theorem, with $\alpha_1 = \sqrt{\frac{2\rho_1 \partial F/\partial z|_{z=\rho_1,S=\sigma_1}}{\partial^2 F/\partial S^2|_{z=\rho_1,S=\sigma_1}}}$ and $\omega_1 = 1/\rho_1$. In order to get the values for the constants in the expansions and for the radius of convergence, we used Maple. From Equation 1, we compute the partial derivatives $\partial F/\partial z = 2zS^4 -$

$2S^3+S^2$ and $\partial F/\partial S = 4{*}z^2{*}S^3-6{*}z{*}S^2+2{*}z{*}S+1$. The system is too complex to be solved formally; so we lower the degree in S by considering the combination $R = 4F-S\partial F/\partial S = -2zS^3+2zS^2+3S-4$ which has to vanish at the points where F and $\partial F/\partial S$ do. Since R is of degree 1 in z, it is easy to get an expression for z that we substitute into $\partial F/\partial S$, obtaining that $8S^3 - 31S^2 + 42S - 20$ should equivalently be zero. Hence we obtain 3 possible algebraic roots, one being real σ_1 and the other two conjugate complex numbers. Only $\sigma_1 = 1.403556586...$ and the associated real value of z for which $F(z,S) = 0$ — $\rho_1 = 0.1324975681...$ — are of interest. A direct approximate solution using the floating point solver of Maple confirms this situation and a more involved study or the Riemann surface also yields $\rho_1 = 0.1324975681...$ to be the radius of convergence of the series. Further computations provide all the constants encountered in the proof and stated in the theorem. □

4.2.2 The Dirks and Pierce class (D&P).

Structures of D&P class are characterized by the presence of non crossing edges and any number of H-type pseudoknots [4, 3].

Theorem 3

$$D\&P(n) \;=\; \frac{\alpha_2}{2\sqrt{\pi}}{\omega_2}^n n^{-3/2}(1+O(1/n)),$$

where $\alpha_2 = 0.7534777262...$, $\omega_2 = 7.3148684640...$, *are algebraic constants.*

Proof (sketch): The following unambiguous grammar generates the language that encodes the D&P structures, according to Proposition 1:

$$S \to dS\bar{d}S|P \;;\; P \to pXS\bar{p}S|\epsilon \;;\; X \to xSX\bar{x}S|ySY\bar{y}S \;;\; Y \to ySY\bar{y}S|\epsilon$$

ù From this grammar, we get the following algebraic equation:

$$F(z,S) = z^3S^6 - z^2S^5 + 2zS^3 - zS^2 - S + 1 = 0 \tag{6}$$

which is very similar to the equation satisfied by the o.g.f. for the $A\&U$ family. We solve it in the same way, and find out the dominant singularity in $z = \rho_2 = 0.1367078581...$, $S = \sigma_2 = 1.439796009...$, with the same local behaviour, implying similar asymptotics for the coefficients. The only problem encountered in finding this dominant singularity comes from the fact that there exists another singularity closer to the origin in $z = \mu = 0.08794976637...$, $S = \tau = 7.169944393...$, but which is not on the same fold of the Riemann surface and which therefore does not have to be taken into consideration. □

4.2.3 The Reeder ang Giegerich class (R&G).

It corresponds to the structures of Reeder and Giegerich's algorithm [15]. It has a $\mathcal{O}(n^4)$ complexity.

Theorem 4

$$R\&G(n) \;=\; \frac{\alpha_3}{2\sqrt{\pi}}{\omega_3}^n n^{-3/2}(1+O(1/n)),$$

where $\alpha_3 = 1.165192913...$, $\omega_3 = 6.576040092...$, *are algebraic constants.*

Proof: In [15], the following grammar is given (we removed the unpaired bases):

$$S \to SS|dS\bar{d}|x^kSy^lS\bar{x}^kS\bar{y}^l|\epsilon.$$

This grammar is not context-free. However, we remark that the pseudoknot defined here is a particular case of a H-Type pseudoknot. So by applying Proposition 1 again, we define the following context free grammar :

$$S \to dS\bar{d}S|P\ ;\ P \to pX\bar{p}S|\epsilon\ ;\ X \to xX\bar{x}|SyY\bar{y}S\ ;\ Y \to yY\bar{y}|S$$

Computations as above lead to the result. □

Additionally, the following theorem places this new class into Condon *et al.*'s classification.

Theorem 5 *$R\&G \subset D\&P$, $L\&P \cap R\&G \neq \emptyset$ and $R\&G \not\subset L\&P$*

Proof (sketch): The grammar wich describes the pseudoknots in R&G is less general than the grammar for H-type pseudoknots. So $R\&G \subset D\&P$ and $L\&P \cap R\&G \neq \emptyset$. As R&G structures can contain several pseudoknots, we have $L\&P \cap R\&G \neq L\&P$. □

4.2.4 The Cao and Chen class (C&C).

It corresponds to the structures of Cao and Chen's algorithm [2], whose complexity is $\mathcal{O}(n^6)$.

Theorem 6

$$C\&C(n) \ = \ \frac{\alpha_4}{2\sqrt{\pi}}{\omega_4}^n n^{-3/2}(1 + O(1/n)),$$

where $\alpha_4 = 1.665071176...$, $\omega_4 = 5.856765093...$, are algebraic constants.

Proof (sketch): The following non context-free grammar generates the C&C structures:

$$S \to SS|dS\bar{d}|x^kSy^l\bar{x}^kS\bar{y}^l|\epsilon.$$

It can be translated into a context-free grammar which is a restriction of the R&G grammar:

$$S \to dS\bar{d}S|P\ ;\ P \to pX\bar{p}S|\epsilon\ ;\ X \to xX\bar{x}|SyY\bar{y}S\ ;\ Y \to yY\bar{y}|\epsilon$$

Computations as above lead to the result. □

Additionally, we easily state that

Theorem 7 *$C\&C \subset D\&P$, $L\&P \cap C\&C \neq 0$, $C\&C \not\subset L\&P$ and $C\&C \subset R\&G$*

4.2.5 The Lyngsø and Pedersen class (L&P).

We already gave a closed formula and an asymptotic equivalent for this class in Section 3. It can be shown that its generating series can also be found in a very simple way by designing a context-free grammar. This will not be developed in this extended abstract.

5 Conclusion

We proved that most classes of pseudoknotted structures that can be predicted by exact algorithms (all but R&E for which the problem remains open) can be encoded by context-free languages. We extended Condon *et al.*'s hierarchy by adding two more classes, and we computed closed or asymptotic formulas for the cardinality of all classes but one.

These results, summarized in Table 1, allow us to quantify the relationship between the complexity of an algorithm and the generality of the class that it can handle. Notably, from a strict quantitative point of view, the growth of complexity by a factor n^2 between the PKF and L&P classes seems not to be justified compared to the very small increase in cardinality. The situation is even worse for the C&C class, whose related algorithm has a stronger complexity than the R&G one, while $C\&C \subset R\&G$ and the ratio of their cardinalitues is exponential. On the other hand, the linear increasing between PKF andR&G complexities seems very reasonable compared to the exponential increasing of the cardinalities.

Besides, the fact that most of the classes are encoded by context-free languages gives an easy way to perform uniform or controlled non uniform random generation of pseudoknotted RNA structures, with standard methods and tools (see *e.g.* [6, 5, 14]).

Acknowledgements. This research was supported in part by the ANR project BRASERO ANR-06-BLAN-0045, and by the Digiteo project "RNAomics".

References

[1] T. Akutsu. Dynamic programming algorithms for RNA secondary structure prediction with pseudoknots. *Discrete Applied Mathematics*, 104:45–62, 2000.

[2] S. Cao and S-J Chen. Predicting structured and stabilities for h-type pseudoknots with interhelix loop. *RNA*, 15:696–706, 2009.

[3] A. Condon, B. Davy, B. Rastegari, S. Zhao, and F. Tarrant. Classifying RNA pseuknotted structures. *Theorical computer science*, 320:35–50, 2004.

[4] N.A. Dirks, R.M. Pierce. A partition function algorithm for nucleic acid secondary structure including pseudoknots. *J Comput Chem*, 24:1664–1677, 2003.

[5] P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer. Boltzmann samplers for the random generation of combinatorial structures. *Combinatorics, Probability, and Computing*, 13(4–5):577–625, 2004. Special issue on Analysis of Algorithms.

[6] Ph. Flajolet, P. Zimmermann, and B. Van Cutsem. A calculus for the random generation of labelled combinatorial structures. *Theoretical Computer Science*, 132:1–35, 1994.

[7] I. L. Hofacker, P. Schuster, and P. F. Stadler. Combinatorics of RNA secondary structures. *Discr. Appl. Math*, 89, 1996.

[8] F. W. D. Huang and M. Reidys. Statistics of canonical RNA pseudoknot structures. *Journal of Theoretical Biology*, 253(3):570–578, 2008.

[9] E. Y. Jin and C. M. Reidys. RNA pseudoknot structures with arc-length ≥ 3 and stack-length $\geq \sigma$. *Discrete Appl. Math.*, 158(1):25–36, 2010.

[10] W.A. Lorenz, Y. Ponty, and P. Clote. Asymptotics of RNA shapes. *Journal of Computational Biology*, 15(1):31–63, Jan–Feb 2008.

[11] R. B. Lyngsø and Pedersen C. N. RNA pseudoknot prediction in energy based models. *Journal of computational biology*, 7:409–428, 2000.

[12] M. E. Nebel. Combinatorial properties of RNA secondary structures. *Journal of Computational Biology*, 9(3):541–574, 2003.

[13] R. Nussinov, G. Pieczenik, J. R. Griggs, and Kleitman D. J. Algorithms for loop matching. *SIAM J. Appl. Math.*, 35:68–82, 1978.

[14] Y. Ponty, M. Termier, and A. Denise. GenRGenS: Software for generating random genomic sequences and structures. *Bioinformatics*, 22(12):1534–1535, 2006.

[15] J. Reeder and R. Giegerich. Design, implementation and evaluation of a practical pseudoknot folding algorithm based on thermodynamics. *BMC Bioinformatics*, 5:104, 2004.

[16] E. Rivas and S. R. Eddy. A dynamic programming algorithm for RNA structure prediction including pseudoknots. *Journal of molecular biology*, 285:2053–2068, 1999.

[17] E. A. Rødland. Pseudoknots in RNA secondary structures: representation, enumeration, and prevalence. *Journal of Computational Biology*, 13(6):1197–1213, 2006.

[18] N. J. A. Sloane and Simon Plouffe. *The Encyclopedia of Integer Sequences*. Academic Press, 1995.

[19] W.T. Tutte. A census of planar maps. *Canadian Journal of Mathematics*, 15:249–271, 1963.

[20] Y. Uemura, A. Hasegawa, S. Kobayashi, and T. Yokomori. Tree adjoining grammars for RNA structures prediction. *Theorical computer science*, 210:277–303, 1999.

[21] M. Vauchaussade de Chaumont and X.G. Viennot. Enumeration of RNA's secondary structures by complexity. In V. Capasso, E. Grosso, and S.L. Paven-Fontana, editors, *Mathematics in Medecine and Biology*, volume 57 of *Lecture Notes in Biomathematics*, pages 360–365, 1985.

[22] G. Vernizzi, H. Orland, and A. Zee. Enumeration of RNA structures by matrix models. *Phys. Rev. Lett.*, 94:168103, 2005.

[23] T. R. S. Walsh and A. B. Lehman. Counting rooted maps by genus. iii: Nonseparable maps. *J. Combinatorial Theory Ser. B*, 18:222–259, 1975.

[24] M. S. Waterman. Secondary structure of single-stranded nucleic acids. *Advances in Mathematics Supplementary Studies*, 1(1):167–212, 1978.

[25] M. Zucker and P. Stiegler. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acid Research*, 9:133–148, 1981.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 916–923

Boolean complexes and boolean numbers

Bridget Eileen Tenner

Department of Mathematical Sciences, DePaul University, Chicago, Illinois

Abstract. The Bruhat order gives a poset structure to any Coxeter group. The ideal of elements in this poset having boolean principal order ideals forms a simplicial poset. This simplicial poset defines the boolean complex for the group. In a Coxeter system of rank n, we show that the boolean complex is homotopy equivalent to a wedge of $(n-1)$-dimensional spheres. The number of these spheres is the boolean number, which can be computed inductively from the unlabeled Coxeter system, thus defining a graph invariant. For certain families of graphs, the boolean numbers have intriguing combinatorial properties. This work involves joint efforts with Claesson, Kitaev, and Ragnarsson.

Résumé. L'ordre de Bruhat munit tout groupe de Coxeter d'une structure de poset. L'idéal composé des éléments de ce poset engendrant des idéaux principaux ordonnés booléens, forme un poset simplicial. Ce poset simplicial définit le complexe booléen pour le groupe. Dans un système de Coxeter de rang n, nous montrons que le complexe booléen est homotopiquement équivalent à un bouquet de sphères de dimension $(n-1)$. Le nombre de ces sphères est le nombre booléen, qui peut être calculé inductivement à partir du système de Coxeter non-étiquetté; définissant ainsi un invariant de graphe. Pour certaines familles de graphes, les nombres booléens satisfont des propriétés combinatoires intriguantes. Ce travail est une collaboration entre Claesson, Kitaev, et Ragnarsson.

Keywords: Coxeter system, Bruhat order, boolean, boolean number, cell complex, homotopy

1 Boolean complexes

The boolean complex of a finitely generated Coxeter system (W, S) was introduced by the author and Ragnarsson in [RT], and the first sections here describe this work. The boolean Complex arises from the (strong) Bruhat order on W. This ordering gives a poset structure to W, and the *boolean ideal* $\mathbb{B}(W, S) \subseteq W$ is the subposet comprised of those elements of W whose principal order ideals are boolean. Such elements are *boolean elements*. This order ideal $\mathbb{B}(W, S)$ is necessarily a simplicial poset.

Definition 1.1 *Let* (W, S) *be a Coxeter system. The* boolean complex *of* (W, S) *is the regular cell complex* $\Delta(W, S)$ *whose face poset is the simplicial poset* $\mathbb{B}(W, S)$.

The existence of the complex $\Delta(W, S)$ follows from a result of Björner about simplicial posets, and in fact about CW-posets (see [Bjö]).

The following lemma is immediate from the definition of the Bruhat order, and gives a useful characterization of boolean elements.

Lemma 1.2 *Let* (W, S) *be a Coxeter system. An element of* W *is boolean if and only if it has no repeated letters in its reduced expressions.*

It follows from Lemma 1.2 that the boolean complex is pure, and each maximal face has dimension $|S| - 1$.

The relevance of boolean elements arises from their properties related to R-polynomials, Kazhdan-Lusztig polynomials, and g-polynomials (see [Bre]).

The fact that the boolean elements are exactly those which are the products of distinct elements of the generating set S means that these elements are described by the commutativity of elements of S. This commutativity information is encoded in the unlabeled Coxeter graph of (W, S), and thus the boolean objects described above can be reformulated in terms of graphs. This graph-theoretic description is given in the following paragraphs, and also appears in work of Jonsson and Welker [JW].

Let G be a finite simple graph, and let S be the vertex set of G. Set $\mathbb{W}(S)$ to be the set of words on S with no repeated letters, ordered by the subword order relation. Elements of $\mathbb{W}(S)$ are thus of the form $s_1 s_2 \cdots s_\ell$, where $s_1, s_2, \ldots, s_\ell \in S$ are distinct elements. We generate an equivalence relation by the requirement that

$$s_1 \cdots s_{i-1} s_i s_{i+1} s_{i+2} \cdots s_\ell \sim s_1 \cdots s_{i-1} s_{i+1} s_i s_{i+2} \cdots s_\ell$$

if $\{s_i, s_{i+1}\}$ is not an edge in G.

The equivalence relation generated by this $\sim$ defines a set of equivalence classes on $\mathbb{W}(S)$, which we call $\mathbb{B}(G)$. The preimages of an element $\sigma \in \mathbb{B}(G)$ are *word representatives*. Note that the set of letters occurring in each word representative of σ is the same. A poset structure on $\mathbb{B}(G)$ arises from the subword order: $\sigma \leq \tau$ in $\mathbb{B}(G)$ if some word representative of σ is a subword of a word representative of τ.

A special case worthy of its own mention is the complete graph K_n. For this graph, the ideal $\mathbb{B}(K_n)$ describes the complex of injective words. This complex has been studied by Farmer [Far], Björner and Wachs [BW], and Reiner and Webb [RW]. The complete graph is treated here in Corollary 4.5.

Example 1.3 *For the complete graph K_2, the poset $\mathbb{B}(K_2)$ and the boolean complex $\Delta(K_2)$ are depicted in Figure 1.*

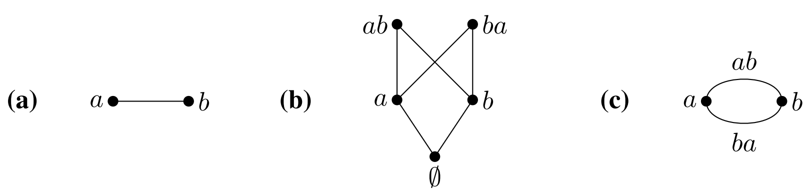


Fig. 1: **(a)** The graph K_2. **(b)** The poset $\mathbb{B}(K_2)$. **(c)** The boolean complex $\Delta(K_2)$, whose geometric realization $|\Delta(K_2)|$ is homotopy equivalent to S^1.

The unlabeled Coxeter graphs of the Coxeter groups $A_2, B_2/C_2, G_2$ and $I_2(m)$ are all the same as K_2. Thus Example 1.3 demonstrates that the geometric realization of each of these boolean complexes is homotopy equivalent to S^1. More generally, as shown in joint work with Ragnarsson in [RT], and discussed in Section 3, the boolean complex of any finite simple graph, and hence of any Coxeter system, has a similar property: its geometric realization is homotopy equivalent to a wedge of top dimensional spheres, and we give a recursive formula for calculating the number of these spheres.

2 Motivation and definitions

This analysis of boolean complexes and their topological and enumerative properties is motivated by the field of Coxeter systems and the relevance of their boolean elements.

Throughout this work, we assume that all Coxeter systems are finitely generated.

The Bruhat order makes a Coxeter group W into a ranked poset, with rank determined by the number of letters in a reduced expression of an element. We use the convention that the minimal element in the poset $\mathbb{B}(W, S)$ has rank -1, because this element corresponds to the empty face in the geometric realization of that poset. This reinforces the fact that the face data of the complex is described by the non-negative ranks in the poset. More precisely, an element of rank $k \geq 0$ represents a k-dimensional cell.

Although one can view the cells in $\mathbb{B}(W, S)$ as simplices, because the minimal subcomplex containing each cell is isomorphic to a simplex of the same dimension, it must be noted that the boolean complex itself is not a simplicial complex. This is because the cells in the complex are not uniquely determined by the vertices they contain. One case of this is depicted in Example 1.3, where two 1-cells have the same vertices.

A geometric realization $|\Delta(W, S)|$ of the boolean complex $\Delta(W, S)$ is obtained in the standard way: we take one geometric simplex of dimension k for each cell of dimension k, and glue them together according to the data encoded in the face poset. The homotopy type of this geometric realization is what we mean when we say the *homotopy type* of a boolean complex.

Theorem 3.4 states the main result in this area, which is that $|\Delta(W, S)|$ has the homotopy type of a wedge of spheres of dimension $|S| - 1$. Moreover, the number of spheres in the wedge can be computed by a recursive formula, also given in the theorem. This recursion is given in terms of graph operations, and uses the alternative construction of the boolean complex, in terms of the unlabeled Coxeter graph of (W, S). The motivation for the construction of $\mathbb{B}(G)$ is the following obvious fact: if G is the unlabeled Coxeter graph of the Coxeter system (W, S), then $\mathbb{B}(G) \cong \mathbb{B}(W, S)$.

Definition 2.1 *The* boolean complex *of a finite simple graph G is the regular cell complex $\Delta(G)$ associated to $\mathbb{B}(G)$.*

The main result about boolean complexes for Coxeter systems can be restated as follows: for any finite simple graph G with vertex set S, the geometric realization $|\Delta(G)|$ is homotopy equivalent to a wedge of $(|S| - 1)$-dimensional spheres. It is this version of the result which we shall state, and, as promised, the recursive formula for the number of spheres is given in terms of basic graph operations on G.

Before discussing the main results, we foreshadow one of them for the Coxeter group A_n. The Euler characteristic χ of a regular cell complex Δ, and likewise the Euler characteristic of its geometric realization $|\Delta|$, is the alternating sum of the number of faces of each non-negative rank in Δ. Given that we consider the rank of the minimal element in a Coxeter group to be -1, this can be computed by an alternating sum of the number of elements in each non-negative rank in the corresponding simplicial poset. In particular, an enumeration from [Ten] enables the calculation of the Euler characteristic of the boolean complex for the Coxeter group A_n.

Corollary 2.2 *For all $n \geq 1$,*

$$\chi(\Delta(A_n)) = (-1)^{n-1} f(n-1) + 1,$$

where $\{f(n)\}$ are the Fibonacci numbers.

It is convenient to use the notation $b \cdot S^r$ for a wedge sum of b spheres of dimension r. Since the wedge sum is the coproduct in the category of pointed spaces, $0 \cdot S^r$ then denotes a single point. Corollary 2.2 foreshadows the fact that $|\Delta(A_n)|$ is homotopy equivalent to the wedge sum $f(n-1) \cdot S^{n-1}$. Although the statistic discussed in this work is equal to the absolute value of the reduced Euler characteristic, Corollary 2.2 references the unreduced Euler characteristic because it is valuable to point out the relationship between $\{\chi(\Delta(A_n))\}$ and sequences A008346 and A119282 in [Slo]: sequence A008346 is equal to $\{|\chi(\Delta(A_n))|\}$, while sequence A119282 is equal to $\{-\chi(\Delta(A_n))\}$.

3 Computing the homotopy type

We now state the main result about the homotopy type of the boolean complex. As mentioned previously, this is stated in terms of the graph formulation (thus it is a statement about $\Delta(G)$) because of a recursive formula involving graph operations.

Definition 3.1 *For a finite graph G, let $|G|$ denote the number of vertices in G.*

Definition 3.2 *Let G be a finite simple graph and e an edge in G. Define three operations on G as follows.*

- Deletion: *$G - e$ is the graph obtained by deleting the edge e.*
- Simple contraction: *G/e is the graph obtained by contracting the edge e and then removing all loops and redundant edges.*
- Extraction: *$G - [e]$ is the graph obtained by removing the edge e and its incident vertices.*

Definition 3.3 *For $n \geq 1$, let δ_n be the graph consisting of n disconnected vertices.*

The symbol $\simeq$ will denote homotopy equivalence in the following theorem and subsequent discussion.

Theorem 3.4 ([RT]) *For every nonempty, finite simple graph G, there is an integer $\beta(G)$ so that*

$$|\Delta(G)| \simeq \beta(G) \cdot S^{|G|-1}.$$

Moreover, the values $\beta(G)$ can be computed using the recursive formula

$$\beta(G) = \beta(G - e) + \beta(G/e) + \beta(G - [e]), \tag{1}$$

if e is an edge in G such that $G - [e]$ is nonempty, with initial conditions

$$\beta(K_2) = 1 \text{ and } \beta(\delta_n) = 0.$$

The integer $\beta(G)$ is the *boolean number* of the graph G.

Because $\mathbb{B}(G) \cong \mathbb{B}(W,S)$ when G is the unlabeled Coxeter graph of the Coxeter system (W,S), Theorem 3.4 implies that the geometric realization of the boolean complex of a Coxeter system (W,S) is homotopy equivalent to a wedge of spheres of dimension $|S|-1$. The number of spheres occurring in the wedge can be calculated recursively using equation (1). In fact, this calculation can be aided by the next result, which states that the boolean number is multiplicative with respect to connected components of a graph.

Proposition 3.5 ([RT]) *If $G = H_1 \sqcup H_2$ for graphs H_1 and H_2, then*

$$\Delta(G) = \Delta(H_1) * \Delta(H_2),$$

where $*$ denotes simplicial join. Thus

$$|\Delta(G)| \simeq \beta(H_1)\beta(H_2) \cdot S^{|H_1|+|H_2|-1}.$$

In particular, $\beta(G) = \beta(H_1)\beta(H_2)$.

4 Homotopy type in special cases

For some classes of graphs, we can obtain more specific data regarding the boolean number β. A selection of these results are given below.

If the graph G has any leaves, then the recursive equation (1) can be simplified by Proposition 3.5.

Corollary 4.1 *If G has a vertex of degree one, then the computation of $\beta(G)$ is simplified according to the identities in Figure 2.*

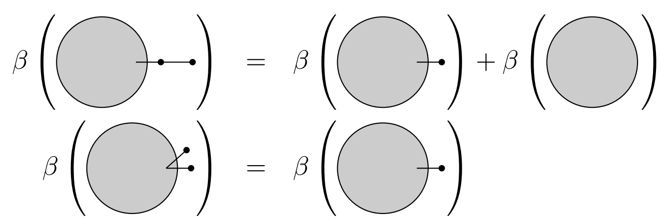


Fig. 2: Simplification for recursively calculating $\beta(G)$ when the graph G has a leaf.

Combining the results of Corollary 4.1 and Theorem 3.4 allows for the efficient calculation of the boolean number of trees. Moreover, it shows that if T is a tree with more than one vertex, then $\beta(T) > 0$. The recursive formula for calculating β shows that adding edges does not decrease the boolean number, which leads to the following conclusion.

Corollary 4.2 ([RT]) *A finite simple graph G has an isolated vertex if and only if $\beta(G) = 0$.*

In other words, the boolean complex of a Coxeter group is contractible if and only if the center of the group contains a generator of the group.

Definition 4.3 *For $n \geq 1$, let S_n be the star graph on n vertices.*

Corollary 4.4 *For $n \geq 1$, $\beta(S_n) = 1$. That is, $\Delta(S_n) \simeq S^{n-1}$.*

Recursively computing the boolean numbers of complete graphs is similarly easy, and gives a recurrence indicating the following corollary.

Corollary 4.5 *For $n \geq 1$, $\beta(K_n)$ is equal to the number of derangements of $[n]$.*

Corollary 4.5 recovers a result of Reiner and Webb, which was proved in the context of the complex of injective words, using character theory [RW].

Equation (1) indicates that the function β is monotonically increasing with respect to edge addition. More precisely, if $H \subseteq G$ is obtained by deleting some edges from the graph G, then $\beta(H) \leq \beta(G)$. One could ask when this inequality is strict, and when there is equality.

Corollary 4.6 ([RT]) *Fix a finite simple graph G. Obtain $H \subset G$ by removing an edge of G. Then $\beta(H) = \beta(G)$ if and only if G has an isolated vertex, in which case $\beta(G) = \beta(H) = 0$. Otherwise $\beta(H) < \beta(G)$.*

As suggested earlier, the boolean number for the Coxeter group A_n is the Fibonacci number$f(n-1)$. Given Corollaries 4.1, 4.4, and 4.6, we see that among all trees on n vertices, the smallest boolean number is obtained by the star S_n, and the largest boolean number is obtained by the path A_n.

5 Boolean numbers: enumerative directions

Given a finite simple graph G, the boolean number $\beta(G)$ is a graph invariant. There are a variety of enumerative aspects of this object that one can explore, several of which are mentioned here.

Corollary 4.6 gives some perspective to the image of the function β. If we analyze the function β, we quickly find that it is neither surjective nor injective, in the following senses. Let $\mathcal{G}_n$ be the set of finite simple graphs having n vertices. The set $\{\beta(G) : G \in \mathcal{G}_n\}$ is not necessarily a contiguous interval of values (necessarily beginning at 0), and it may be that $\beta(G) = \beta(H)$ for distinct graphs $G, H \in \mathcal{G}_n$.

Example 5.1 *No graph in $\mathcal{G}_4$ has boolean number 4, although the boolean number of a 4-cycle is 5.*

Example 5.2 *The two graphs in Figure 3 each have boolean number 3, and thus are each homotopy equivalent to $S^4 \vee S^4 \vee S^4$.*

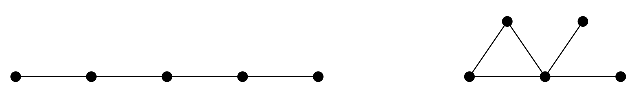


Fig. 3: Two elements of $\mathcal{G}_5$, each having boolean number 3.

As suggested already, there are several families of graphs whose boolean numbers give well-known sequences. Three such families already mentioned are

- graphs with disjoint vertices (boolean numbers are always 0),
- paths (boolean numbers are the Fibonacci numbers), and
- complete graphs (boolean numbers are the derangement numbers).

Other interesting connections to integer sequences have arisen in joint work with Claesson, Kitaev, and Ragnarsson [CKRT]. These results are related to Ferrers graphs, particularly for rectangular and staircase shapes, and are described below.

As described in [CN], Ferrers shapes, or Young shapes or partitions, arise in a variety of contexts including Schubert varieties, symmetric functions, hypergeometric series, permutation statistics, quantum mechanical operators, and inverse rook problems. One can relate a bipartite graph known as a *Ferrers graph* to any Ferrers shape, as introduced in [EvW].

Definition 5.3 *Let $\lambda = (\lambda_1, \ldots, \lambda_r)$ be a partition, where $\lambda_1 \geq \cdots \geq \lambda_r \geq 0$. The associated bipartite Ferrers graph has vertices $\{x_1, \ldots, x_r\} \sqcup \{y_1, \ldots, y_{\lambda_1}\}$, and edges $\big\{\{x_i, y_j\} : \lambda_i \geq j\big\}$.*

In particular, vertex x_i has degree λ_i. A Ferrers graph and its associated Ferrers shape are depicted in Figure 4.

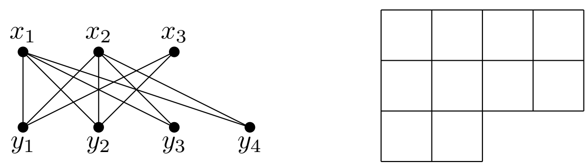


Fig. 4: The Ferrers graph and shape for the partition $\lambda = (4, 4, 2)$.

The Ferrers graph of an m-by-n rectangular shape is the complete bipartite graph $K_{m,n}$. Computation of the boolean number of such a graph invokes the Stirling numbers of the second kind.

Corollary 5.4 ([CKRT]) *For $m, n \geq 1$,*

$$\beta(K_{m,n}) = \sum_{k=1}^{m} (-1)^{m-k} k! \left\{ {m+1 \atop k+1} \right\} k^n.$$

Even more directly, the boolean numbers of the Ferrers graph of staircase shapes are exactly the median Genocchi numbers, which are sequence A005439 of [Slo]. As shown in [ES], the median Genocchi number g_n is equal to the number of permutations of $2n$ letters having alternating excedances.

Definition 5.5 *For $n \geq 1$, the* staircase shape of height n *is the Ferrers shape*

$$\Sigma_n = (n, n-1, \ldots, 2, 1).$$

Let F_n denote the Ferrers graph for the Ferrers shape Σ_n.

Corollary 5.6 ([CKRT]) *For $n \geq 1$, $\beta(F_n) = g_n$.*

References

[Bjö] A. Björner, Posets, regular CW complexes and Bruhat order, *European J. Combin.* **5** (1984), 7–16.

[BW] A. Björner and M. Wachs, On lexicographically shellable posets, *Trans. Amer. Math. Soc.* **277** (1983), 323–341.

[Bre] F. Brenti, A combinatorial formula for Kazhdan-Lusztig polynomials, *Invent. Math.* **118** (1994), 371–394.

[CKRT] A. Claesson, S. Kitaev, K. Ragnarsson, and B. E. Tenner, Boolean complexes for Ferrers graphs, preprint.

[CN] A. Corso and U. Nagel, Monomial and toric ideals associated to Ferrers graphs, *Trans. Amer. Math. Soc.*, **361** (2009), 1371–1395.

[ES] R. Ehrenborg and E. Steingrímsson, The excedance set of a permutation, *Adv. Appl. Math.* **24** (2000), 284–299.

[EvW] R. Ehrenborg and S. van Willigenburg, Enumerative properties of Ferrers graphs, *Discrete Comput. Geom. special issue in honor of Louis J. Billera*, **32** (2004), 481–492.

[Far] F. D. Farmer, Cellular homology for posets, *Math. Japon.* **23** (1978/79), 607–613.

[JW] J. Jonsson and V. Welker, Complexes of injective words and their commutation classes, *Pacific J. Math.* **243** (2009), 313–329.

[RT] K. Ragnarsson and B. E. Tenner, Homotopy type of the boolean complex of a Coxeter system, *Adv. Math.* **222** (2009), 409–430.

[RW] V. Reiner and P. Webb, The combinatorics of the bar resolution in group cohomology, *J. Pure Appl. Algebra* **190** (2004), 291–327.

[Slo] N. J. A. Sloane, The on-line encyclopedia of integer sequences, published electronically at http://www.research.att.com/˜njas/sequences/.

[Sta] R. P. Stanley, *Enumerative Combinatorics, vol. 1*, Cambridge Studies in Advanced Mathematics, no. 49, Cambridge University Press, Cambridge, 1997

[Ten] B. E. Tenner, Pattern avoidance and the Bruhat order, *J. Combin. Theory, Ser. A* **114** (2007), 888–905.

FPSAC 2010, San Francisco, USA DMTCS proc. **AN**, 2010, 924–935

A unification of permutation patterns related to Schubert varieties

Henning A. Úlfarsson[†]

Mathematics Institute, School of Computer Science, Reykjavik University, Reykjavík, Iceland

Abstract. We prove new connections between permutation patterns and singularities of Schubert varieties, by giving a new characterization of factorial and Gorenstein varieties in terms of so called *bivincular patterns*. These are generalizations of classical patterns where conditions are placed on the location of an occurrence in a permutation, as well as on the values in the occurrence. This clarifies what happens when the requirement of smoothness is weakened to factoriality and further to Gorensteinness, extending work of Bousquet-Mélou and Butler (2007), and Woo and Yong (2006). We also prove results that translate some known patterns in the literature into bivincular patterns.

Résumé. Nous démontrons de nouveaux liens entre les motifs de permutation et les singularités des variétés de Schubert, par la méthode de donner une nouvelle caractérisation des variétés factorielles et de Gorenstein par rapport à les *motifs bivinculaires*. Ces motifs sont généralisations des motifs classiques où des conditions se posent sur la position d'une occurrence dans une permutation, aussi bien que sur les valeurs qui se présentent dans l'occurrence. Ceci éclaircit les phénomènes où la condition de nonsingularité s'affaiblit á factorialité et même à Gorensteinité, et augmente les travaux de Bousquet-Mélou et Butler (2007), et de Woo et Yong (2006). Nous démontrons également des résultats qui traduisent quelques motifs connus en la littérature en motifs bivinculaires.

Keywords: Patterns, Permutations, Schubert varieties, Singularities

1 Introduction

We prove new connections between permutation patterns and singularities of Schubert varieties X_π in the complete flag variety $\mathrm{Flags}(\mathbb{C}^n)$, by giving a new characterization of factorial and Gorenstein varieties in terms of which *bivincular patterns* the permutation π avoids. Bivincular patterns, defined by Bousquet-Mélou et al. (2010), are generalizations of classical patterns where conditions are placed on the location of an occurrence in a permutation, as well as on the values in the occurrence. This clarifies what happens when the requirement of smoothness is weakened to factoriality and further to Gorensteinness, extending work of Bousquet-Mélou and Butler (2007), and Woo and Yong (2006). We also prove results that translate some known patterns in the literature into bivincular patterns. In particular we will give a characterization of the Baxter permutations.

In section 2 we recall the definitions of classical, vincular (also called generalized patterns, Babson-Steingrímsson patterns or dashed patterns), bivincular and barred patterns. We will also recall Bruhat embeddings of patterns, as defined by Woo and Yong (2006).

[†]This work is supported by grant no. 090038011 from the Icelandic Research Fund.

In section 3 we will show how these different types of patterns and their avoidances are related to each other. We will then use patterns that come up in the connections between permutations and Schubert varieties as motivation. In particular, recall the theorem of Ryan (1987), Wolper (1989) and Lakshmibai and Sandhya (1990) that the Schubert variety X_π is non-singular (or smooth) if and only if π avoids the patterns 1324 and 2143. Saying that the variety X_π is non-singular means that every local ring is regular. A weakening of this condition is the requirement that every local ring only be a unique factorization domain; a variety satisfying this is a *factorial* variety.

Bousquet-Mélou and Butler (2007) proved a conjecture stated by Yong and Woo (Bousquet-Mélou et al., 2005, Personal communication) that factorial Schubert varieties are those that correspond to permutations avoiding 1324 and bar-avoiding $21\bar{3}54$. In the terminology of Woo and Yong (2006) the bar-avoidance of the latter pattern corresponds to avoiding 2143 with Bruhat condition $(1 \leftrightarrow 4)$, or equivalently, interval avoiding $[2413, 2143]$ in the terminology of Woo and Yong (2008). However, as remarked by Steingrímsson (2007), bar-avoiding $21\bar{3}54$ is equivalent to avoiding the vincular pattern $2\underline{14}3$. If we summarize this in terms of vincular patterns a striking thing becomes apparent; see the lines corresponding to smoothness and factoriality in Table 1, and Theorem 3.1.

X_π is	The permutation π avoids the patterns
smooth	2143 and 1324
factorial	$2\underline{14}3$ and 1324
Gorenstein	$\begin{smallmatrix}12\overline{34}5\\3\underline{15}24\end{smallmatrix}$, $\begin{smallmatrix}1\overline{23}45\\24\underline{15}3\end{smallmatrix}$; and satisfies a condition on descents involving two infinite families of bivincular patterns

Tab. 1: Connections between singularity properties and bivincular patterns. See Theorem 3.1 for the second line and see Theorem 3.13 for the third line.

We see that requiring 1 and 4 to be adjacent in the first pattern (and thus turning it into a vincular pattern) corresponds to weakening smoothness to factoriality.

A further weakening is to only require that the local rings of X_π be Gorenstein local rings, in which case we say that X_π is a *Gorenstein* variety. Woo and Yong (2006) showed that X_π is Gorenstein if and only if it avoids two patterns with two Bruhat restrictions each, as well as satisfying a certain condition on descents. We will translate their results into avoidance of bivincular patterns; see Theorem 3.13 in subsection 3.2 and the line corresponding to Gorensteinness in Table 1. Essentially the two bivincular patterns that are shown there are certain *upgrades* of the pattern $2\underline{14}3$ while the avoidance of the pattern 1324 is weakened to avoidance of two infinite families of bivincular patterns.

In section 3 we will also prove a proposition that leads to a characterization of the Baxter permutations in terms of vincular patterns, see Example 3.5.

2 Three Types of Pattern Avoidance

Here we recall definitions of different types of patterns. We will use one-line notation for all permutations, *e.g.*, write $\pi = 312$ for the permutation in S_3 that satisfies $\pi(1) = 3$, $\pi(2) = 1$ and $\pi(3) = 2$.

The three types correspond to:

- Bivincular patterns, subsuming vincular patterns and classical patterns.

- Barred patterns.
- Bruhat restricted patterns.

2.1 Bivincular Patterns

We denote the symmetric group on n letters by S_n, whose elements are *permutations*. We write permutations as words $\pi = a_1 a_2 \cdots a_n$, where the letters are distinct and come from the set $\{1, 2, \ldots, n\}$. A *pattern* p is also a permutation, but we are interested in when a pattern is *contained* in a permutation π as described below.

An *occurrence* (or *embedding*) of a pattern p in a permutation π is classically defined as a subsequence in π, of the same length as p, whose letters are in the same relative order (with respect to size) as those in p. For example, the pattern 123 corresponds to a increasing subsequence of three letters in a permutation. If we use the notation 1_π to denote the first, 2_π for the second and 3_π for the third letter in an occurrence, then we are simply requiring that $1_\pi < 2_\pi < 3_\pi$. If a permutation has no occurrence of a pattern p we say that π *avoids* p.

Example 2.1 *The permutation* 32415 *contains two occurrences of the pattern* 123 *corresponding to the sub-words* 345 *and* 245*. It avoids the pattern* 132*.*

In a *vincular pattern* (also called a *generalized pattern*, *Babson-Steingrímsson pattern* or *dashed pattern*), two adjacent letters may or may not be underlined. If they are underlined it means that the corresponding letters in the permutation π must be adjacent.

Example 2.2 *The permutation* 32415 *contains one occurrence of the pattern* $\underline{12}3$ *corresponding to the sub-word* 245*. It avoids the pattern* $1\underline{23}$*. The permutation* $\pi = 324615$ *has one occurrence of the pattern* 2143*, namely the sub-word* 3265*, but no occurrence of* $2\underline{14}3$*, since* 2 *and* 6 *are not adjacent in* π*.*

These types of patterns have been studied sporadically for a very long time but were not defined in full generality until Babson and Steingrímsson (2000).

This notion was generalized further in Bousquet-Mélou et al. (2010): In a *bivincular pattern* we are also allowed to put restrictions on the values that occur in an embedding of a pattern. We use two-line notation to describe these patterns. If there is a line over the letters i, $i+1$ in the top row, it means that the corresponding letters in an occurrence must be adjacent in values. This is best described by an example:

Example 2.3 *An occurrence of the pattern* $\begin{smallmatrix}\overline{12}3\\123\end{smallmatrix}$ *in a permutation* π *is an increasing subsequence of three letters, such that the second one is larger than the first by exactly* 1*, or more simply* $2_\pi = 1_\pi + 1$*. The permutation* 32415 *contains one occurrence of this bivincular pattern corresponding to the sub-word* 345*. This is also an occurrence of* $\begin{smallmatrix}\overline{123}\\123\end{smallmatrix}$*. The permutation avoids the bivincular pattern* $\begin{smallmatrix}\overline{123}\\1\underline{23}\end{smallmatrix}$*.*

We will also use the notation of Bousquet-Mélou et al. (2010) to write bivincular patterns: A bivincular pattern consists of a triple (p, X, Y) where p is a permutation in S_k and X, Y are subsets of $[\![0, k]\!]$. An occurrence of this bivincular pattern in a permutation $\pi = \pi_1 \cdots \pi_n$ in S_n is a subsequence $\pi_{i_1} \cdots \pi_{i_k}$ such that the letters in the subsequence are in the same relative order as the letters of p and

- for all x in X, $i_{x+1} = i_x + 1$; and
- for all y in Y, $j_{y+1} = j_y + 1$, where $\{\pi_{i_1}, \ldots, \pi_{i_k}\} = \{j_1, \ldots, j_k\}$ and $j_1 < j_2 < \cdots < j_k$.

By convention we put $i_0 = 0 = j_0$ and $i_{k+1} = n + 1 = j_{k+1}$.

Example 2.4 *We can translate all of the patterns we have discussed above into this notation:*

$$
\begin{array}{lll}
123 = (123, \varnothing, \varnothing), & 132 = (132, \varnothing, \varnothing), & \underline{12}3 = (123, \{1\}, \varnothing), \\
1\underline{23} = (123, \{2\}, \varnothing), & 2143 = (2143, \varnothing, \varnothing), & 2\underline{14}3 = (2143, \{2\}, \varnothing), \\
\begin{matrix}\overline{12}3\\123\end{matrix} = (123, \varnothing, \{1\}), & \begin{matrix}\overline{123}\\123\end{matrix} = (123, \varnothing, \{1, 2\}), & \begin{matrix}\overline{123}\\1\underline{23}\end{matrix} = (123, \{2\}, \{1, 2\}).
\end{array}
$$

We have not considered the case when 0 or k are elements of X or Y, as we will not need those cases. We just remark that if $0 \in X$ then an occurrence of (p, X, Y) must start at the beginning of a permutation π, in other words, $\pi_{i_1} = \pi_1$. The other cases are similar.

The bivincular patterns behave well with respect to the operations reverse, complement and inverse: Given a bivincular pattern (p, X, Y) we define

$$(p, X, Y)^{\mathrm{r}} = (p^{\mathrm{r}}, k - X, Y), \qquad (p, X, Y)^{\mathrm{c}} = (p^{\mathrm{c}}, X, k - Y), \qquad (p, X, Y)^{\mathrm{i}} = (p^{\mathrm{i}}, Y, X),$$

where p^{r} is the usual reverse of the permutation of p, p^{c} is the usual complement of the permutation of p, and p^{i} is the usual inverse of the permutation of p. Here $k - M = \{k - m \mid m \in M\}$.

We get a very simple but useful lemma:

Lemma 2.5 *Let* a *denote one of the operations above (or their compositions). Then a permutation* π *avoids the bivincular pattern* p *if and only if the permutation* π^{a} *avoids the bivincular pattern* p^{a}. □

2.2 Barred Patterns

We will only consider a single pattern of this type, but the general definition is easily inferred from this special case. We say that a permutation π *avoids the barred pattern* $21\bar{3}54$ if π avoids the pattern 2143 (corresponding to the unbarred elements) *except* where that pattern is a part of the pattern 21354. This notation for barred patterns was introduced by West (1990). It turns out that avoiding this barred pattern is equivalent to avoiding $2\underline{14}3$, see section 3.

Example 2.6 *The permutation* $\pi = 425761$ *avoids the barred pattern* $21\bar{3}54$ *since the unique occurrence of* 2143*, as the sub-word* 4276*, is contained in the sub-word* 42576 *which is an occurrence of* 21354*.*

2.3 Bruhat Restricted Patterns

Here we recall the definition of Bruhat restricted patterns from Woo and Yong (2006). First we need the *Bruhat order* on permutations in S_n, defined as follows: Given integers $i < j$ in $[\![1, n]\!]$ and a permutation $\pi \in S_n$ we define $w(i \leftrightarrow j)$ as the permutation that we get from π by swapping $\pi(i)$ and $\pi(j)$. For example $24153(1 \leftrightarrow 4) = 54123$. We then say that $\pi(i \leftrightarrow j)$ *covers* π if $\pi(i) < \pi(j)$ and for every k with $i < k < j$ we have either $\pi(k) < \pi(i)$ or $\pi(k) > \pi(j)$. We then define the Bruhat order as the transitive closure of the above covering relation. This definition should be compared to the construction of the graph G_π in subsection 3.1. We see that in our example above that $24153(1 \leftrightarrow 4)$ does not cover 24153 since we have $\pi(2) = 4$. Now, given a pattern p with a set of transpositions $\mathcal{T} = \{(i_\ell \leftrightarrow j_\ell)\}$ we say that a permutation π *contains* $(p, \mathcal{T})$, or that π *contains the Bruhat restricted pattern* p, if $\mathcal{T}$ is understood from the context, if there is an embedding of p in π such that if any of the transpositions in $\mathcal{T}$ are carried out on the embedding the resulting permutation covers π.

We should note that Bruhat restricted patterns were further generalized to *intervals of patterns* in Woo and Yong (2008). We will not consider this generalization here.

In the next section we will show how these three types of patterns are related to one another.

3 Connections between the Three Types

3.1 Factorial Schubert Varieties and Forest-like Permutations

Bousquet-Mélou and Butler (2007) defined and studied *forest-like* permutations. Here we recall their definition: Given a permutation π in S_n, construct a graph G_π on the vertex set $\{1, 2, \ldots, n\}$ by joining i and j if

1. $i < j$ and $\pi(i) < \pi(j)$; and

2. there is no k such that $i < k < j$ and $\pi(i) < \pi(k) < \pi(j)$.

The permutation π is *forest-like* if the graph G_π is a forest. In light of the definition of Bruhat covering above we see that the vertices i and j are connected in the graph of G_π if and only if $\pi(i \leftrightarrow j)$ covers π.

They then show that a permutation is forest-like if and only if it avoids the classical pattern 1324 and the barred pattern $p_{\text{bar}} = 21\bar{3}54$. This barred pattern can be described in terms of Bruhat restricted embeddings and in terms of bivincular patterns, as we now show.

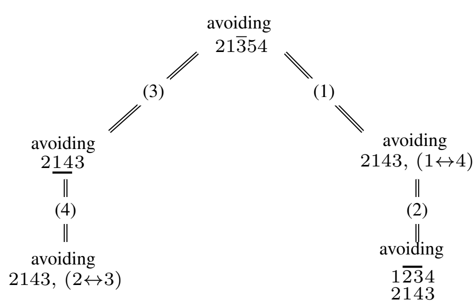


Fig. 1: The barred pattern $21\bar{3}54$ gives a connection between two bivincular patterns. The labels on the edges correspond to the enumerated list below.

1. Bousquet-Mélou and Butler (2007) remark that forest-like permutations π correspond to factorial Schubert varieties X_π and avoiding the barred pattern is equivalent to avoiding $p_{\text{Br}} = 2143$ with Bruhat restriction $(1 \leftrightarrow 4)$. This last part is easily verified.

2. Avoiding $p_{\text{Br}} = 2143$ with Bruhat restriction $(1 \leftrightarrow 4)$ is equivalent to avoiding the bivincular pattern $p_{\text{bi}} = \begin{smallmatrix}1\overline{23}4\\2143\end{smallmatrix}$, as we will now show:

   Assume π contains the bivincular pattern p_{bi}, so we can find an embedding of it in π such that $3_\pi = 2_\pi + 1$. This embedding clearly satisfies the Bruhat restriction.

Now assume that π has an embedding of p_{Br}. If $3_\pi = 2_\pi + 1$ we are done. Otherwise $2_\pi + 1$ is either to the right of 3_π or to the left of 2_π (because of the Bruhat restriction). In the first case change 3_π to $2_\pi + 1$ and we are done. In the second case replace 2_π with $2_\pi + 1$, thus reducing the distance in values to 3_π, then repeat.

This will be generalized in Proposition 3.8.

3. The barred pattern $p_{\text{bar}} = 21\bar{3}54$ has another connection to bivincular patterns: avoiding it is equivalent to avoiding the bivincular pattern $q_{\text{biv}} = 2\underline{14}3$, as remarked in the survey by Steingrímsson (2007).

4. We can translate this into Bruhat restricted embeddings as well: Avoiding the bivincular pattern $q_{\text{bi}} = 2\underline{14}3$ is equivalent to avoiding $q_{\text{Br}} = 2143$ with Bruhat restriction $(2 \leftrightarrow 3)$:

   Assume π has an embedding of q_{Br}. If 1_π and 4_π are adjacent then we are done. Otherwise look at the letter to right of 1_π. If this letter is larger than 4_w we can replace 4_w by it and we are done. Otherwise this letter must be less than 4_w, which implies by the Bruhat restriction, that it must also be less than 1_w. In this case we replace 1_w by this letter, and repeat.

   Now assume π has an embedding of the bivincular pattern q_{bi}. If 1_π and 4_π are adjacent we are done. Otherwise look at the letter to the right of 1_π. This letter is either smaller than 1_π or larger than 4_π. In the first case, replace 1_π with this letter; in the second case, replace 4_π with this letter. Then repeat if necessary.

   This will be generalized in Proposition 3.8.

This gives us:

Theorem 3.1 *Let $\pi \in S_n$. The Schubert variety X_π is factorial if and only if π avoids the patterns $2\underline{14}3$ and 1324.* □

From the equivalence of the patterns in Figure 1 we also get that a permutation π avoids the bivincular pattern

$$2\underline{14}3 = (2143, \{2\}, \varnothing)$$

if and only if it avoids

$$\begin{smallmatrix}1\overline{23}4\\2143\end{smallmatrix} = (2143, \varnothing, \{2\}).$$

We will prove this without going through the barred pattern, and then generalize the proof, but first of all we should note that these bivincular patterns are inverses of one another, and that will simplify the proof.

Assume π contains $\begin{smallmatrix}1\overline{23}4\\2143\end{smallmatrix}$. If 1_π and 4_π are adjacent in π we are done. Otherwise consider the element immediately to the right of 1_π. If this element is less than 2_π then replace 1_π by it and we will have reduced the distance between 1_π and 4_π. If this element is larger than 2_π it must also be larger than 3_π, since $3_\pi = 2_\pi + 1$, so replace 4_π by it. This will (immediately, or after several steps) produce an occurrence of $2\underline{14}3$.

Now assume π contains $2\underline{14}3$. Then π^{i} contains the inverse pattern

$$(2\underline{14}3)^{\text{i}} = \begin{smallmatrix}1\overline{23}4\\2143\end{smallmatrix}.$$

Then by the above, π^{i} contains $2\underline{14}3$, so $\pi = (\pi^{\mathrm{i}})^{\mathrm{i}}$ contains $(2\underline{14}3)^{\mathrm{i}} = \genfrac{}{}{0pt}{}{1\overline{23}4}{2143}$.

This generalizes to:

Proposition 3.2 *Let p be the pattern*

$$\cdots \underline{1k} \cdots \ = (\cdots 1k \cdots, \{j\}, \varnothing)$$

in S_k, where $j = p^{\mathrm{i}}(1)$ is the index *of 1 in p, so $j+1$ is the index of k in p. A permutation π in S_n that avoids the pattern p must also avoid the bivincular pattern*

$$\genfrac{}{}{0pt}{}{1\overline{2\cdot\cdots\cdot}k}{\cdots 1k \cdots} = (\cdots 1k \cdots, \varnothing, \{2, 3, \ldots, k-2\}).$$

Proof: Assume a permutation π contains the latter pattern in the proposition. If 1_π and k_π are adjacent in π we are done. Otherwise consider the element immediately to the right of 1_π. If this element is larger than $(k-1)_\pi$ we replace k_π by it and are done. Otherwise this element must me less than $(k-1)_\pi$ and therefore less than 2_π, so we can replace 1_π by it, and repeat. □

By applying the reverse to everything in sight in Proposition 3.2 we get:

Corollary 3.3 *Let p be the pattern*

$$\cdots \underline{k1} \cdots \ = (\cdots k1 \cdots, \{j\}, \varnothing)$$

in S_k, where $j = p^{\mathrm{i}}(k)$ is the index *of k in p, so $j+1$ is the index of 1 in p. A permutation π in S_n that avoids the pattern p must also avoid the bivincular pattern*

$$\genfrac{}{}{0pt}{}{1\overline{2\cdot\cdots\cdot}k}{\cdots k1 \cdots} = (\cdots k1 \cdots, \varnothing, \{2, 3, \ldots, k-2\}).$$

By repeatedly applying the operations of inverse, reverse and complement we can generate six other corollaries. We will not need them here so we will not write them down.

Example 3.4 *Let's look at some simple applications:*

1. *Consider the bivincular pattern $p_1 = 3\underline{14}2$. Proposition 3.2 shows a permutation π that avoids p_1 must also avoid $\genfrac{}{}{0pt}{}{1\overline{23}4}{3142}$. In fact, the converse can be shown to be true, by taking inverses and applying the proposition. We will say more about the pattern p_1 in Example 3.5.*
2. *Consider the bivincular pattern $p_2 = 3\underline{15}24$. The proposition shows that a permutation π that avoids p_2 must also avoid $\genfrac{}{}{0pt}{}{1\overline{234}5}{31524}$. We will say more about the pattern p_2 in subsection 3.2.*

Example 3.5 *The* Baxter permutations *were originally defined and studied in relation to the "commuting function conjecture" of Dyer, see Baxter (1964), and were enumerated in Chung et al. (1978). Gire (1993) showed that these permutations can also be described as those avoiding the barred patterns $41\bar{3}52$ and $25\bar{1}35$. It was then pointed out by Ouchterlony (2005) that this is equivalent to avoiding the vincular patterns $3\underline{14}2$ and $2\underline{41}3$.*

Similarly to what we did above we can show that the Baxter permutations can also be characterized as those avoiding the bivincular patterns $\genfrac{}{}{0pt}{}{1\overline{23}4}{3142}$ *and* $\genfrac{}{}{0pt}{}{1\overline{23}4}{2413}$*, and this is essentially a translation of the description in Chung et al. (1978) into bivincular patterns.*

Finally, here is an example that shows the converse of Proposition 3.2 is not true.

Example 3.6 *The permutation* $\pi = 423165$ *avoids the pattern* $\genfrac{}{}{0pt}{}{1\overline{234}5}{23154}$ *but contains the pattern* $23\underline{15}4$*, as the sub-word* 23165*.*

3.2 Gorenstein Schubert Varieties in terms of Bivincular Patterns

Woo and Yong (2006) classify those permutations π that correspond to Gorenstein Schubert varieties X_w. They do this using embeddings of patterns with Bruhat restrictions, which we have described above, and with a certain condition on the associated Grassmannian permutations of w, which we will describe presently:

First, a *descent* in a permutation π is an integer d such that $\pi(d) > \pi(d+1)$. A *Grassmannian permutation* is a permutation with a unique descent. Given any permutation π we can associate a Grassmannian permutation to each of its descents, as follows: Given a particular descent d of π we construct the sub-word $\gamma_d(\pi)$ by concatenating the right-to-left minima of the segment strictly to the left of $d+1$ with the left-to-right maxima of the segment strictly to the right of d. More intuitively we start with the descent $\pi(d)\pi(d+1)$ and enlarge it to the left by adding increasing elements without creating another descent and similarly enlarge it to the right by adding decreasing elements without creating another descent. We then denote the *flattening* (or *standardization*) of $\gamma_d(\pi)$ by $\tilde{\gamma}_d(\pi)$, which is the unique permutation whose letters are in the same relative order as $\gamma_d(\pi)$.

Example 3.7 *Consider the permutation* $\pi = 11|6|12|94153728|10$ *where we have used the symbol* $|$ *to separate two digit numbers from other numbers. For the descent at* $d = 4$ *we get* $\gamma_4(\pi) = 694578|10$ *and* $\tilde{\gamma}_4(\pi) = 3612457$*.*

Now, given a Grassmannian permutation π in S_n with its unique descent at d we construct its *associated partition* $\lambda(\pi)$ as the partition inside a bounding box $d \times (n-d)$, with d rows and $n-d$ columns, whose lower border is the lattice path that starts at the lower left corner of the bounding box and whose i-th step, for $i \in [\![1, n]\!]$, is vertical if i is weakly to the left of the position d, and horizontal otherwise. We are interested in the *inner corner distance* of this partition, *i.e.*, for every inner corner we add its distance from the left side and the distance from the top of the bounding box. If all these inner corner distances are the same then the inner corners all lie on the same anti-diagonal.

In Theorem 1 of Woo and Yong (2006) they show that a permutation $\pi \in S_n$ corresponds to a Gorenstein Schubert variety X_π if and only if

1. for each descent d of π, $\lambda(\tilde{\gamma}_d(\pi))$ has all of its inner corners on the same anti-diagonal; and
2. the permutation π avoids both 31524 and 24153 with Bruhat restrictions $\{(1 \leftrightarrow 5), (2 \leftrightarrow 3)\}$ and $\{(1 \leftrightarrow 5), (3 \leftrightarrow 4)\}$, respectively.

Let's take a closer look at condition 2: Proposition 3.8 below shows that avoiding 31524 with Bruhat restrictions $\{(1 \leftrightarrow 5), (2 \leftrightarrow 3)\}$ is equivalent to avoiding the bivincular pattern

$$\genfrac{}{}{0pt}{}{12\overline{34}5}{3\underline{15}24} = (31524, \{2\}, \{3\}).$$

Similarly, avoiding 24153 with Bruhat restrictions $\{(1 \leftrightarrow 5), (3 \leftrightarrow 4)\}$ is equivalent to avoiding the bivincular pattern

$$\begin{smallmatrix} 1\overline{23}45 \\ 24\underline{15}3 \end{smallmatrix} = (24153, \{3\}, \{2\}).$$

Proposition 3.8 *1. Let p be the pattern*

$$\cdots 1k \cdots$$

in S_k. Let $j = p^{\mathrm{i}}(1)$ be the index of 1 in p, so $j+1$ is the index of k in p. A permutation π in S_n avoids p with Bruhat restriction $(j \leftrightarrow j+1)$ if and only if π avoids the vincular pattern

$$\cdots \underline{1k} \cdots \quad = (\cdots 1k \cdots, \{j\}, \varnothing).$$

2. Let $\ell \in [\![1, k-1]\!]$ and p be the pattern

$$\ell \cdots (\ell+1)$$

in S_k. A permutation π in S_n avoids p with Bruhat restriction $(1 \leftrightarrow k)$ if and only if π avoids the bivincular pattern

$$\begin{smallmatrix} 1 \cdots \overline{\ell\,\ell{+}1} \cdots k \\ \ell \cdots\cdots\cdots \ell{+}1 \end{smallmatrix} = (\ell \cdots (\ell+1), \varnothing, \{\ell\}).$$

Proof: We consider each case separately.

1. Assume π contains the vincular pattern mentioned. Then it clearly also contains an embedding satisfying the Bruhat restriction.

   Conversely assume π contains an embedding satisfying the Bruhat restriction. If 1_π and k_π are adjacent then we are done. Otherwise look at the element immediately to the right of 1_π. This element must be either larger than k_π, in which case we can replace k_π by it and are done, or smaller, in which case we replace 1_π by it, and repeat.

2. Assume π contains the bivincular pattern mentioned. Then it clearly also contains an embedding satisfying the Bruhat restriction.

   Conversely assume π contains an embedding satisfying the Bruhat restriction. If $(\ell+1)_\pi = \ell_\pi + 1$ then we are done. Otherwise consider the element $\ell_\pi + 1$. It must either be to the right of $(\ell+1)_\pi$ or to the left of ℓ_π. In the first case we can replace $(\ell+1)_\pi$ by $\ell_\pi + 1$ and be done. In the second case replace ℓ_π with $\ell_\pi + 1$ and repeat. □

As a consequence we get:

Corollary 3.9 *A permutation π in S_n avoids*

$$\cdots 1k \cdots, \ (j \leftrightarrow j+1),$$

where j is the index of 1, if and only if the inverse π^{-1} avoids

$$j \cdots (j+1), \ (1 \leftrightarrow k).$$

Note that we could have proved the statement of the corollary with out going through bivincular patterns and then used that to prove part 2 of Proposition 3.8, as part 2 is the inverse statement of the statement in part 1.

Translating condition 1 of Theorem 1 of Woo and Yong (2006) is a bit more work. The failure of this condition is easily seen to be equivalent to some partition λ of an associated Grassmannian permutation $\tilde{\gamma}_d(\pi)$ having an outer corner that is either "too wide" or "too deep". Precisely, given a Grassmannian permutation π and an outer corner of $\lambda(\pi)$, we say that it is *too wide* if the distance upward from it to the next inner corner is smaller than the distance to the left from it to the next inner corner. Conversely we say that an outer corner is *too deep* if the distance upward from it to the next inner corner is larger than the distance to the left from it to the next inner corner. We say that an outer corner is *unbalanced* if it is either too wide or too deep. We say that an outer corner is *balanced* if it is not *unbalanced*.

If a permutation has an associated Grassmannian permutation with an outer corner that is too wide we say that the permutation itself is *too wide* and similarly for *too deep*. If the permutation is either too wide or too deep we say that it is *unbalanced*, otherwise it is *balanced*. It is time to see some examples.

Example 3.10

1. *Consider the permutation* $\pi = 14235$ *with a unique descent at* $d = 2$*. It corresponds to the partition* $(2) \subseteq 2 \times 3$ *and has just one outer corner. This outer corner is too wide.*

2. *Consider the permutation* $\pi = 13425$ *with a unique descent at* $d = 3$*. It corresponds to the partition* $(1, 1) \subseteq 3 \times 2$ *and has just one outer corner. This outer corner is too deep.*

3. *Consider the permutation* $\pi = 134892567|10$ *with a unique descent at* $d = 5$*. It corresponds to the partition* $(4, 4, 1, 1) \subseteq 5 \times 5$ *and has two outer corners. The first outer corner is too deep and the second is too wide.*

4. *Consider the permutation* $\pi = 13672458$ *with a unique descent at* $d = 4$*. It corresponds to the partition* $(3, 3, 1) \subseteq 4 \times 4$ *and has two outer corners that are both balanced.*

These properties of Grassmannian permutations can be detected with bivincular patterns, as we now show.

Lemma 3.11 *Let* π *be a Grassmannian permutation.*

1. *The permutation* π *is too wide if and only if it contains at least one of the bivincular patterns from the infinite family*

$$\mathcal{F} = \left(\begin{array}{c} 1\overline{2345} \\ 14235 \end{array}, \begin{array}{c} 1\overline{234567} \\ 1562347 \end{array}, \begin{array}{c} 1\overline{23456789} \\ 167823459 \end{array}, \dots \right).$$

The general member of this family is of the form

$$\begin{array}{c} 1\overline{2\cdot\cdots\cdots\cdot k} \\ 1\,\ell{+}1\cdot\cdot 2\cdot\cdot \ell\, k \end{array},$$

where $\ell = (k-3)/2$.

2. *The permutation π is too deep if and only if it contains at least one of the bivincular patterns from the infinite family*

$$\mathcal{G} = \left(\begin{smallmatrix} \overline{1234}5 \\ 13425 \end{smallmatrix}, \begin{smallmatrix} \overline{123456}7 \\ 1456237 \end{smallmatrix}, \begin{smallmatrix} \overline{12345678}9 \\ 156782349 \end{smallmatrix}, \cdots \right).$$

The general member of this family is of the form

$$\begin{smallmatrix} \overline{12 \cdots\cdots\cdots} k \\ 1\,\ell{+}1 \cdots 2 \cdots \ell\, k \end{smallmatrix},$$

where $\ell = (k-1)/2$.

Proof: We only consider part 1, as part 2 is proved analogously. Assume that π is a Grassmannian permutation that is too wide, so it has an outer corner that is too wide. Let ℓ be the distance from this outer corner to the next inner corner above. Then the distance from this outer corner to the next inner corner to the left is at least $\ell + 1$. This allows us to construct an increasing sequence t of length ℓ in π, starting at a distance at least two to the right of the descent. We can also choose t so that every element in it is adjacent both in location and values. Similarly we can construct an increasing sequence s of length ℓ in π, located strictly to the left of the descent. We can also choose s so that every element in it is adjacent both in location and values. This produces the required member of the family $\mathcal{F}$.

Conversely, assume π contains a particular member of the family $\mathcal{F}$. Then π clearly has at least one outer corner that is too wide. □

It should be noticed that these two infinite families are obtained from one another by reverse complement.

We have now shown that

Proposition 3.12 *A permutation π is balanced if and only if every associated Grassmannian permutation avoids every bivincular pattern in the two infinite families $\mathcal{F}$ and $\mathcal{G}$ in Lemma 3.11.* □

This gives us:

Theorem 3.13 *Let $\pi \in S_n$. The Schubert variety X_π is Gorenstein if and only if*

1. *π is balanced; and*

2. *the permutation π avoids the bivincular patterns*

$$\begin{smallmatrix} 12\overline{34}5 \\ 3\underline{15}24 \end{smallmatrix} \text{ and } \begin{smallmatrix} 1\overline{23}45 \\ 24\underline{15}3 \end{smallmatrix}.$$

□

I should note that with the descriptions of factorial and Gorenstein Schubert varieties given above it is easy to verify that smoothness implies factoriality implies Gorensteinness.

References

E. Babson and E. Steingrímsson. Generalized permutation patterns and a classification of the Mahonian statistics. *Sém. Lothar. Combin.*, 44:Art. B44b, 18 pp. (electronic), 2000. ISSN 1286-4889.

G. Baxter. On fixed points of the composite of commuting functions. *Proc. Amer. Math. Soc.*, 15:851–855, 1964. ISSN 0002-9939.

M. Bousquet-Mélou and S. Butler. Forest-like permutations. *Ann. Comb.*, 11(3-4):335–354, 2007. ISSN 0218-0006.

M. Bousquet-Mélou, S. Butler, A. Woo, and A. Yong. Personal communication between Bousquet-Mélou and Butler, and Woo and Yong, January 2005.

M. Bousquet-Mélou, A. Claesson, M. Dukes, and S. Kitaev. (2+2)-free posets, ascent sequences and pattern avoiding permutations. *Journal of Combinatorial Theory Series A*, 117:884–909, 2010.

F. R. K. Chung, R. L. Graham, V. E. Hoggatt, Jr., and M. Kleiman. The number of Baxter permutations. *J. Combin. Theory Ser. A*, 24(3):382–394, 1978. ISSN 0097-3165. doi: 10.1016/0097-3165(78)90068-7. URL `http://dx.doi.org/10.1016/0097-3165(78)90068-7`.

S. Gire. *Arbres, permutations à motifs exclus et cartes planaires: quelques problèmes algorithmique et combinatoire*. PhD thesis, Université Bordeaux I, 1993.

V. Lakshmibai and B. Sandhya. Criterion for smoothness of Schubert varieties in $\mathrm{Sl}(n)/B$. *Proc. Indian Acad. Sci. Math. Sci.*, 100(1):45–52, 1990. ISSN 0253-4142.

E. Ouchterlony. *On Young tableaux involutions and patterns in permutations*. PhD thesis, Matematiska institutionen Linköpings universitet, Linköping, Sweden, 2005.

K. M. Ryan. On Schubert varieties in the flag manifold of $\mathrm{Sl}(n, \mathbf{C})$. *Math. Ann.*, 276(2):205–224, 1987. ISSN 0025-5831. doi: 10.1007/BF01450738. URL `http://dx.doi.org/10.1007/BF01450738`.

E. Steingrímsson. Generalized permutation patterns – a short survey. *To appear in "Permutation Patterns, St Andrews 2007", S.A. Linton, N. Ruskuc, V. Vatter (eds.), LMS Lecture Note Series, Cambridge University Press*, 2007. URL `http://arxiv.org/abs/0801.2412v2`.

J. West. *Permutations with forbidden subsequences and stack-sortable permutations*. PhD thesis, MIT, 1990.

J. S. Wolper. A combinatorial approach to the singularities of Schubert varieties. *Adv. Math.*, 76(2):184–193, 1989. ISSN 0001-8708. doi: 10.1016/0001-8708(89)90048-0. URL `http://dx.doi.org/10.1016/0001-8708(89)90048-0`.

A. Woo and A. Yong. When is a Schubert variety Gorenstein? *Adv. Math.*, 207(1):205–220, 2006. ISSN 0001-8708.

A. Woo and A. Yong. Governing singularities of Schubert varieties. *J. Algebra*, 320(2):495–520, 2008. ISSN 0021-8693.