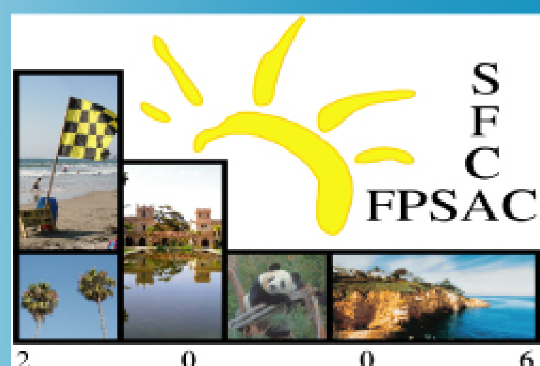
FPSAC 2006
SFCA 2006

Proceedings
Actes

18th Annual
International Conference on
Formal Power Series and
Algebraic Combinatorics
June 19 - 23, 2006

18-ième
Conférence Internationale sur les
Séries Formelles et
la Combinatoire Algébrique
19 -- 23 juin 2006

San Diego, California
Catamaran Resort Hotel,
Pacific Beach, USA



FPSAC 2006 – SFCA 2006 San Diego

Programme Committee – Comité de programme

François Bergeron (UQAM) *(chair/ président)*
Jeff Remmel (UC San Diego)*(co-chair /co-président)*
Nantel Bergeron (York University)
Miklós Bóna (University of Florida)
Bill Chen (Nankai University)
Joseph Kung (University of North Texas)
Roberto Mantaci (Université Paris 7)
Jennifer Morse (University of Miami)
Rosa Orellana (Dartmouth College)
Andrew Rechnitzer (University of Melbourne)
Manfred Schocker (University of Wales Swansea)
Hugh Thomas (University of New Brunswick)
Mike Zabrocki (York University)

Organizing Committee – Comité d'organisation

Jeff Remmel (UC San Diego)*(co-chair/co-président)*
Mike Zabrocki (York University)*(co-chair/co-président)*
Nantel Bergeron (York University)
Sara Faridi (Dalhousie University)
Christophe Hohlweg (The Fields Institute)
Isaiah Lankham (UC Davis)
Marni Mishna (Simon Fraser University)
Rosa Orellana (Dartmouth College)
Andrew Rechnitzer (University of Melbourne)
Tom Roby (University of Connecticut)
Mercedes Rosas (Universidad de Sevilla)

Sponsors – Commanditaires

York University

University of California

National Security Agency

Mathematical Sciences Research Institute

US National Science Foundation

Proceedings prepared by Marni Mishna, with the help of translators Cedric Chauve, Christophe Hohlweg and Philippe Choquette.

Welcome to San Diego!

Dear friends,

We are looking forward to a successful conference and the organizers are happy to be hosting FPSAC in San Diego this year. San Diego is a major tourist destination with such attractions as the San Diego Zoo, Wild Animal Park, Sea World, the Stephen Birch Aquarium, beaches and a proximity to mountains and deserts. We hope that many of you who are traveling here will take some time before or after the conference and see a little of southern California.

The organizers are very grateful to the National Security Agency and the National Science Foundation who have both provided generous support to qualified applicants.

We would also like to acknowledge the financial support provided by the Mathematical Sciences Research Institute (MSRI) and the Center for Communications Research as well as from York University and University of California, San Diego. We would also like to thank Anita McKee and Natalie Powell, staff at UCSD, who have worked to provide logistic support to the organizing committee.

Finally, we would also like to thank the members of the organizing committee who have donated many hours of work towards the tasks which are required to put on an event like this.

Sincerely, on behalf of the organizing committee,

Jeff Remmel	Mike Zabrocki
University of California, San Diego	York University

FPSAC at 18

For Canadians, 18 is the age of majority and, more importantly, 18 is the age at which one has the right to buy alcohol. Even though, this year's conference is being in the held in the USA (where the alcohol buying majority is at 21), we feel that we can safely say that 2006 FPSAC Conference has reached maturity as this is its 18th incarnation. This is certainly reflected in the depth and variety of subjects that are going to be discussed at the 2006 FPSAC in San Diego.

The invited speakers, M. Bayer, F. Chung, J. Haglund, S.J. Kang, T. Koornwinder, N. Ray, B. Sagan and M. Wachs, represent expertise a great variety of well established research areas, spanning important and interesting aspects of combinatorics, algebra, and geometry, among others. It is also the case that an exceptionally broad list of subjects are represented in the contributed papers and posters. These contributions clearly reflect the richness and liveliness of our subject. We note that many of the submissions are from young researchers which reflects well on the vitality of algebraic combinatorics. It was a conscious "bias" of the program committee to give preference to submissions from young researchers in its selection process.

More then half of the contributed talks are by young researchers and this proportion is even higher in the contributed posters. Despite this bias, the main criteria used by the Program Committee to select papers for presentations or for posters was the overall excellence of the contribution. Clear trends appear to emerge from even a cursory study of the list of titles of all contributions.

Our combinatorial community is showing a marked interest in interactions between its subject and other areas of mathematics and physics. This is a clear indication of a subject that has both maturity and scope. It also spells out the need for continuing the FPSAC series, and we look forward to next year's meeting in Nankai University, Tianjin, China.

François Bergeron (chair) and Jeff Remmel (co-chair), for the Program Committee.

Table of Contents

Invited Speakers – Conférenciers Invités

Presentations – Présentations

Posters – Affiches

Software – Logiciels

Part I

Invited Presentations
Conférenciers Invités

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Flag Vectors of Polytopes: An Overview

Margaret M. Bayer

A convex polytope is the convex hull of a finite set of points in $\mathbf{R}^d$. A d-dimensional polytope has faces of dimension 0 through $d-1$; each face is itself a convex polytope. The faces (along with $\emptyset$ and P itself), ordered by inclusion, form a lattice. This talk is concerned with a study of the face lattices of convex polytopes.

Of historical importance is the problem of characterizing the face vectors of polytopes; these vectors give the number of faces of each dimension. The characterization of face vectors of 3-dimensional polytopes was done by Steinitz a century ago. For 4-dimensional polytopes the problem is still open. The biggest advance since Steinitz was the characterization of face vectors of simplicial polytopes (where all faces are simplices) by Stanley, and Billera and Lee in 1980.

The face vector is apparently not robust enough for attempts at characterization by combinatorial and algebraic techniques. We turn instead to the *flag vector* of a polytope. For a d-dimensional polytope P, and $S = \{s_1, s_2, \ldots, s_k\} \subseteq \{0, 1, \ldots, d-1\}$, $f_S(P)$ is the number of chains of faces $\emptyset \subset F_1 \subset F_2 \subset \cdots \subset F_k \subset P$ with $\dim F_i = s_i$. The *flag vector* of P is the length 2^d vector $(f_S(P))_{S\subset\{0,1,\ldots,d-1\}}$. In the cases of 3-dimensional polytopes and simplicial polytopes the flag vector is determined linearly by the face vector; in general it can be viewed as an extension of the face vector.

Richard Stanley (1979) studied flag vectors of Cohen-Macaulay posets, a class that contains face lattices of convex polytopes. Bayer and Billera (1985) proved the generalized Dehn-Sommerville equations, the complete set of linear equations satisfied by the flag vectors of all convex polytopes. Since then a wide variety of approaches have been used in the study of flag vectors.

A crucial ingredient in the characterization of face vectors of simplicial polytopes is the connection with toric varieties. In the nonsimplicial case, the middle perversity intersection homology of the toric variety gives an h-vector, linearly dependent on the flag vector. Results from algebraic geometry translate into linear inequalities on the flag vector (Stanley 1987). Another main source of linear inequalities is the cd-index of a polytope, discovered by Jonathan Fine (1985). The cd-index is a vector linearly equivalent to the flag vector; it can be viewed as a reduction of the flag vector by the generalized Dehn-Sommerville equations.

Rigidity theory, shellings, and co-algebras have been used to generate inequalities on flag vectors of polytopes. The talk will survey results and highlight techniques. Some results pertain to special classes of polytopes, such as cubical polytopes and zonotopes. Others hold for more general classes of combinatorial objects, such as general graded posets, Eulerian posets, and Gorenstein* lattices.

We are still, apparently, far from a characterization of flag vectors of polytopes. In fact, we do not even know if the closed convex cone of flag vectors is finitely generated. This is an area of active research. It has exposed interesting connections with other areas of combinatorics, algebra and geometry.

Department of Mathematics, University of Kansas, 405 Snow Hall, 1460 Jayhawk Blvd, Lawrence, KS 66045-7523

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The diameter and Laplacian of directed graphs

Fan Chung Graham

We consider Laplacians for directed graphs. The spectral gap of the Laplacian can be used to establish an upper bound for the diameter of a directed graph. In addition, the Laplacian eigenvalues of a directed graph capture various isoperimetric properties of the directed graph. For example, we will discuss several versions of the Cheeger inequalities and derive bounds for mixing time for random walks on directed graphs or non-reversible Markov chains.

As to related links, there are some relavant papers at my homepage: http://www.math.ucsd.edu/ fan

Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0112

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Recent Combinatorial Results Involving Macdonald Polynomials and Diagonal Harmonics

Jim Haglund

The theory of nonsymmetric Macdonald polynomials was developed by Cherednik, Macdonald and Opdam. Like their symmetric counterparts, they have versions for arbitrary root systems, which satisfy an orthogonality relation and a norm evaluation (generalizing Macdonald's constant term conjecture), and which feature in a generalization of Selberg's integral. Their construction of these polynomials was existential, and up to now no particularly nice expressions for them were known. In this talk we overview some of this history, and then present an explicit combinatorial formula for the type A versions of these polynomials, which is recent joint work with M. Haiman and N. Loehr. We then discuss connections of our formula to the theory of symmetric functions and earlier conjectures involving the character of diagonal harmonics. Time permitting, some recent results on diagonal harmonics will be highlighted.

Department of Mathematics, University of Pennsylvania, 209 South 33rd, St.David Rittenhouse Laboratory, Philadelphia, PA 19104-6395

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Crystal bases for quantum generalized Kac-Moody algebras

Seok-Jin Kang

We develop the crystal basis theory for quantum generalized Kac-Moody algebras. We define the notion of crystal bases for $U_q(\mathfrak{g})$-modules in the category $\mathcal{O}_{int}$ and prove the standard properties of crystal bases including the tensor product rule. We then prove that there exist crystal bases (and global bases) for $V(\lambda)$ ($\lambda \in P^+$) and $U_q^-(\mathfrak{g})$.

We also introduce the notion of abstract crystals for quantum generalized Kac-Moody algebras and study their fundamental properties. Finally, we prove the crystal embedding theorem and give a characterization of the crystals $B(\infty)$ and $B(\lambda)$.

Department of Mathematics, Seoul National University, Seoul 151-742, Korea

Structure relation and raising/lowering operators for orthogonal polynomials

Tom H. Koornwinder

The structure relation for classical orthogonal polynomials (OP's), is traditionally defined as a fixed polynomial times the derivative of the n-th degree OP being equal to some explicit linear combination of the OP's of degree n-1, n and n+1, with coefficients depending on n. By substitution of the three-term recurrence relation, the structure relation gives rise to a relation with a raising of lowering operator. A variant of the structure relation can be obtained, for all OP's in the Askey scheme and the q-Askey scheme, by taking the commutator of the second order operator having the OP's as eigenfunctions and the operator of multiplication by x. The lecture will survey past approaches and results on structure relations etc. for OP's in the (q-)Askey scheme and for multivariable OP's associated with root systems. The so-called string equation also pops up here. Then some new results, in particular in the multivariable case will be presented.

Some references:

(1) W.A. Al-Salam and T.S. Chihara,
Another characterization of the classical orthogonal polynomials,
SIAM J. Math. Anal. 3 (1972), 65-70.
(2) A.S. Zhedanov, "Hidden symmetry" of Askey-Wilson polynomials,
Theoret. and Math. Phys. 89 (1991), 1146-1157.
(3) T.H. Koornwinder, Lowering and raising operators for some special orthogonal polynomials,
arXiv:math.CA/0505378.
(4) T.H. Koornwinder The structure relation for Askey-Wilson polynomials,
arXiv:math.CA/0601303.

University of Amsterdam,Plantage Muidergracht 24 1018 TV Amsterdam The Netherlands

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Hopf algebroids: can combinatorialists help?

Nigel Ray

In this talk I shall attempt to explain certain algebraic concepts which seem ripe for combinatorial modelling. I shall avoid technical details by focusing on the underlying ideas, and will concentrate on a few basic examples that I hope convey the flavour of the challenge to those who may not be algebraic experts. I shall certainly not presuppose any familiarity with algebraic topology!

The study of Hopf algebras was initiated by algebraic topologists in the 1930s, and has been permeating other areas of mathematics and theoretical physics ever since. Thanks to the vision of Gian-Carlo Rota and his associates, the theory entered combinatorics during the 1960s, and their viewpoint has now begun to enjoy modest feedback into topology. During the 1970s, however, topologists had already discovered that certain generalisations of Hopf algebras arise rather naturally in stable homotopy theory, and the resulting structures came of age when their status as cogroupoid objects was properly understood.

I shall describe these ideas in terms of examples of two main types. First are those which are particularly straightforward, and therefore illustrate the basic principles rather well to a general audience, and second are those which are relevant to the study of formal power series, and have close ties with algebraic topology. In some of these cases the structures in question are merely Hopf algebras; in others, the full power of algebroids is required. It would be exciting for topologists if combinatorial models could be constructed in these situations, and I shall outline a couple of situations where some success has been achieved in this direction.

Topics I hope to mention include partitions of finite sets, composition of formal power series, and groupoids as graphs.

School of Mathematics University of Manchester Oxford Road Manchester M13 9PL England

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The Incidence Algebra of a Composition Poset

Bruce E. Sagan

A *composition* is just a sequence $w = k_1 k_2 \dots k_r$ of positive integers. A number of partial orders on the set of all compositions have been studied recently. For example, Björner and Stanley have defined a poset on compositions which has many of the properties of Young's lattice for partitions. We consider a ordering that was first defined by Bergeron, Bousquet-Mélou, and Dulucq: Given $u = k_1 \dots k_r$ and $w = l_1 \dots l_s$ then we have $u \leq w$ if there is a subsequence $l_{i_1} \dots l_{i_r}$ of w which is componentwise bigger than u, i.e.,

$$k_j \leq l_{i_j} \quad \text{for } 1 \leq j \leq r. \tag{1}$$

Call this poset C. It is interesting, in part, because it is related to the poset of all permutations ordered by pattern containment.

In the first half of the talk, we will study the zeta function, ζ, of the incidence algebra $I(C)$. This is joint work with Anders Björner and full details can be found in the paper at

http://www.math.msu.edu/~sagan/Papers/rmf.pdf

If $w = k_1 \dots k_r$ satisfies $\sum_i k_i = N$, then w is said to be a *composition of* N and we write $|w| = N$. Let c_N be the number of compositions of N. It is well known (and easy to prove) that

$$\sum_{N \geq 0} c_N x^N = \frac{1-x}{1-2x} \tag{2}$$

which is a rational function of x. Now given $u \in C$, consider the generating function

$$Z(u) = \sum_{w \geq u} x^{|w|} = \sum_{w \in C} \zeta(u, w) x^{|w|}.$$

So equation 2 is just $Z(\epsilon)$, where ϵ is the empty composition. We show that $Z(u)$ is always a rational function by using techniques from the theory of formal languages. We also investigate similar generating function for powers of ζ. Surprisingly, to evaluate the sums, hypergeometric series identities are needed.

In the second half of the talk, we will study the Möbius function, $\mu = \zeta^{-1}$, in $I(C)$. This is joint work with Vincent Vatter and full details can be found in the paper at

http://www.math.msu.edu/~sagan/Papers/mfc.pdf

A set of indices $I = \{i_1, \dots, i_r\}$ such that 1 holds is called an *embedding* of u into w. We show that $\mu(u, w)$ gives a signed counting of certain embeddings of u into w. In fact, there are three proofs of this result: one combinatorial via an involution, one topological using discrete Morse theory, and one using the machinery discussed in the previous paragraph (this last being work with Björner). We will present the topological proof, giving an introduction to discrete Morse theory in the process.

The results above have analogues in the work Björner, some of it with Reutenauer, on subword order. We will show that both their results and ours are part of a more general framework. In particular, let P be any poset and consider the *Kleene closure* of all words over P:

$$P^* = \{w = k_1 k_2 \dots k_r \ : \ k_i \in P \text{ for all } i \text{ and } r \geq 0\}.$$

Then 1 defines a partial order $u \leq w$ on elements of P^*, where the inequalities $k_j \leq l_{i_j}$ are taken in P. With little extra effort, one can prove theorems about P^* which specialize to those about composition order or about subword order by taking P to be a chain or an antichain, respectively.

We will end with a list of intriguing conjectures and open questions concerning these ideas. As another surprise, the Tchebyshev polynomials of the first kind enter into a conjectured formula for the Möbius function of certain intervals of P^* for a particular 3-element poset, P.

Department of Mathematics, Michigan State University, East Lansing, MI 48824-1027

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Poset topology and permutation statistics

Michelle Wachs

Various connections between permutation statistics and poset topology have been explored in the literature over the past three decades originating with the work of Stanley. In this talk I will present a connection, recently discovered with John Shareshian. I will discuss how a study of the topology of a certain interesting class of posets has led to results and conjectures on a new q-analog of the Eulerian polynomials. These new q-Eulerian polynomials are the enumerators for the joint distribution of the excedance number and the major index. One of our conjectures is a formula for their q-exponential generating function, which is a nice q-analog of a well-known formula for the exponential generating function of the Eulerian polynomials. A more general version of this conjecture involves an intriguing new class of quasisymmetric functions and a representation of the symmetric group on the cohomology of the toric variety associated with the Coxeter complex of the symmetric group, studied by Procesi, Stanley, and Stembridge.

Department of Mathematics, University of Miami, Coral Gables, FL 33124

Part II

Presentations
Présentations

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Flag arrangements and triangulations of products of simplices.

Federico Ardila and Sara Billey

ABSTRACT. We investigate the line arrangement that results from intersecting d complete flags in $\mathbb{C}^n$. We give a combinatorial description of the matroid $\mathcal{T}_{n,d}$ that keeps track of the linear dependence relations among these lines.

We prove that the bases of the matroid $\mathcal{T}_{n,3}$ characterize the triangles with holes which can be tiled with unit rhombi. More generally, we provide evidence for a conjectural connection between the matroid $\mathcal{T}_{n,d}$, the triangulations of the product of simplices $\Delta_{n-1} \times \Delta_{d-1}$, and the arrangements of d tropical hyperplanes in tropical $(n-1)$-space.

Our work provides a simple and effective criterion to ensure the vanishing of many Schubert structure constants in the flag manifold, and a new perspective on Billey and Vakil's method for computing the non-vanishing ones.

RÉSUMÉ. Nous étudions l'arrangement de droites qui résulte de l'intersection de d drapeaux complets dans $\mathbb{C}^n$. Nous donnons une description combinatoire du matroide $\mathcal{T}_{n,d}$ défini par les dépendances linéaires entre ces droites.

Nous démontrons que les bases du matroide $\mathcal{T}_{n,3}$ caractérisent les triangles sans trou qui peuvent être pavés par des losanges unitaires. Plus généralement, nous étayons une relation conjecturale entre le matroide $\mathcal{T}_{n,d}$, les triangulations du produit de simplexes $\Delta_{n-1} \times \Delta_{d-1}$ et les arrangements de d hyperplans tropicaux dans l'espace tropical de dimension $n-1$.

Nos travaux produisent un critère simple et efficace pour déterminer quand de nombreuses constantes de structure de Schubert sont nulles, et une nouvelle façon de voir la méthode de Billey et Vakil pour calculer celles qui sont non-nulles..

1. Introduction.

Let $E_\bullet^1, \ldots, E_\bullet^d$ be d generically chosen complete flags in $\mathbb{C}^n$. Write

$$E_\bullet^k = \{\{0\} = E_0^k \subset E_1^k \subset \cdots \subset E_n^k = \mathbb{C}^n\},$$

where E_i^k is a vector space of dimension i. Consider the set $\mathbf{E}_{n,d}$ of one-dimensional intersections determined by the flags; that is, all lines of the form $E_{a_1}^1 \cap E_{a_2}^2 \cap \cdots \cap E_{a_d}^d$.

The initial goal of this paper is to characterize the line arrangements $\mathbb{C}^n$ which arise in this way from d generically chosen complete flags. We will then show an unexpected connection between these line arrangements and an important and ubiquitous family of subdivisions of polytopes: the triangulations of the product of simplices $\Delta_{n-1} \times \Delta_{d-1}$. These triangulations appear naturally in studying the geometry of the product of all minors of a matrix [**1**], tropical geometry [**4**], and transportation problems [**17**]. To finish, we will illustrate some of the consequences that the combinatorics of these line arrangements have on the Schubert calculus of the flag variety.

The results of the paper are roughly divided into four parts as follows. First of all, Section 2 is devoted to studying the line arrangement determined by the intersections of a generic arrangement of hyperplanes. This will serve as a warmup before we investigate generic arrangements of complete flags, and the results we obtain will be useful in that investigation.

2000 *Mathematics Subject Classification.* Primary 52C35; Secondary 05B35,14N15, 52C22.

Key words and phrases. flag variety, matroids, permutation arrays, tilings, triangulations, Schubert calculus.

The second part consists of Sections 3 and 4, where we will characterize the line arrangements that arise as intersections of a "matroid-generic" arrangement of d flags in $\mathbb{C}^n$. Section 3 is a short discussion of the combinatorial setup that we will use to encode these geometric objects. In Section 4, we propose a combinatorial definition of a matroid $\mathcal{T}_{n,d}$, and show that it is the matroid of the line arrangement of any d flags in $\mathbb{C}^n$ which are generic enough. Finally, we show that these line arrangements are completely characterized combinatorially: any line arrangement in $\mathbb{C}^n$ whose matroid is $\mathcal{T}_{n,d}$ arises as an intersection of d flags.

The third part establishes a surprising connection between these line arrangements and an important class of subdivisions of polytopes. The bases of $\mathcal{T}_{n,3}$ exactly describe the ways of punching n triangular holes into the equilateral triangle of size n, so that the resulting holey triangle can be tiled with unit rhombi. A consequence of this is a very explicit geometric representation of $\mathcal{T}_{n,3}$. We show these results in Section 5. We then pursue a higher-dimensional generalization of this result. In Section 6, we suggest that the fine mixed subdivisions of the Minkowski sum $n\Delta_{d-1}$ are an adequate $(d-1)$-dimensional generalization of the rhombus tilings of holey triangles. We give a completely combinatorial description of these subdivisions. Finally, in Section 7, we prove that each pure mixed subdivision of the Minkowski sum $n\Delta_{d-1}$ (or equivalently, each triangulation of the product of simplices $\Delta_{n-1} \times \Delta_{d-1}$) gives rise to a basis of $\mathcal{T}_{n,d}$. We conjecture that every basis of $\mathcal{T}_{n,d}$ arises in this way. In fact, we conjecture that every basis of $\mathcal{T}_{n,d}$ arises from a *coherent* subdivision or, equivalently, from an arrangement of d tropical hyperplanes in tropical $(n-1)$-space.

The fourth and last part of the paper, Section 8, presents some of the consequences of our work in the Schubert calculus of the flag variety. We start by recalling Eriksson and Linusson's permutation arrays, and Billey and Vakil's related method for explicitly intersecting Schubert varieties. In Section 8.1 we show how the geometric representation of the matroid $\mathcal{T}_{n,3}$ of Section 5 gives us a new perspective on Billey and Vakil's method for computing the structure constants c_{uvw} of the cohomology ring of the flag variety. Finally, Section 8.2 presents a simple and effective criterion for guaranteeing that many Schubert structure constants are equal to zero.

2. The lines in a generic hyperplane arrangement.

Before thinking about flags, let us start by studying the slightly easier problem of understanding the matroid of lines of a generic arrangement of m hyperplanes in $\mathbb{C}^n$. We will start by presenting, in Proposition 2.1, a combinatorial definition of this matroid $\mathcal{H}_{n,m}$. Theorem 2.2 then shows that this is, indeed, the right matroid. As it turns out, this warmup exercise will play an important role in Section 4.

Throughout this section, we will consider an arrangement of m generically chosen hyperplanes $H_1, \ldots, H_m$ in $\mathbb{C}^n$ passing through the origin. For each subset A of $[m]$, let

$$H_A = \bigcap_{a \in A} H_a.$$

By genericity,

$$\dim H_A = \begin{cases} n - |A| & \text{if } |A| \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the set $L_{n,m}$ of one-dimensional intersections of the H_is consists of the $\binom{m}{n-1}$ lines H_A for $|A| = n-1$.

There are several "combinatorial" dependence relations among the lines in $L_{n,m}$, as follows. Each t-dimensional intersection H_B (where B is an $(n-t)$-subset of $[m]$) contains the lines H_A with $B \subseteq A$. Therefore, in an independent set $H_{A_1}, \ldots, H_{A_k}$ of $L_{n,m}$, we cannot have $t+1$ A_is which contain a fixed $(n-t)$-set B.

At first sight, it seems intuitively clear that, in a generic hyperplane arrangement, these will be the only dependence relations among the lines in $L_{n,m}$. This is not as obvious as it may seem: let us illustrate a situation in $L_{4,5}$ which is surprisingly close to a counterexample to this statement. For simplicity, we will draw the projective picture, and denote hyperplanes $H_1, \ldots, H_5$ simply by $1, \ldots, 5$, and an intersection like H_{124} simply by 124.

In Figure 1, we have started by drawing the triangles T and T' with vertices $124, 234, 134$ and $125, 235, 135$, respectively. The three lines connecting the pairs $(124, 125)$, $(234, 235)$ and $(134, 135)$, are the lines 12, 23, and 13, respectively. They intersect at the point 123, so that the triangles T and T' are perspective with respect to this point.

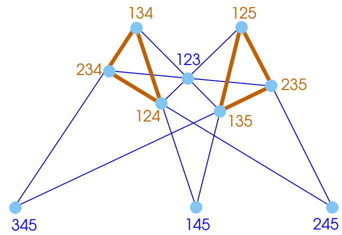


FIGURE 1. The Desargues configuration in $L_{4,5}$.

Now, Desargues' theorem applies, and it predicts an unexpected dependence relation. It tells us that the three points of intersection of the corresponding sides of T and T' are collinear. The lines 14 (which connects 124 and 134) and 15 (which connects 125 and 135) intersect at the point 145. Similarly, 24 and 25 intersect at 245, and 34 and 35 intersect at 345. Desargues' theorem says that the points 145, 245, and 345 are collinear. In principle, this new dependence relation does not seem to be one of our predicted "combinatorial relations". Somewhat surprisingly, it is: it simply states that these three points are on the line 45.

The previous discussion illustrates two points. First, it shows that Desargues' theorem is really a combinatorial statement about incidence structures, rather than a geometric statement about points on the Euclidean plane. Second, and more important to us, it shows that even five generic hyperplanes in $\mathbb{C}^4$ give rise to interesting geometric configurations. It is not unreasonable to think that larger arrangements $L_{n,m}$ will contain other configurations, such as the Pappus configuration, which have nontrivial and honestly geometric dependence relations that we may not have predicted.

Having told our readers what they might need to worry about, we now intend to convince them not to worry about it.

First we show that the combinatorial dependence relations in $L_{n,m}$ are consistent, in the sense that they define a matroid.

PROPOSITION 2.1. *Let $\mathcal{I}$ consist of the collections I of subsets of $[m]$, each containing $n-1$ elements, such that no $t+1$ of the sets in I contain an $(n-t)$-set. In symbols,*

$$\mathcal{I} := \left\{ I \subseteq \binom{[m]}{n-1} \text{ such that for all } S \subseteq I,\ \Big|\bigcap_{A\in S} A\Big| \leq n - |S| \right\}.$$

Then $\mathcal{I}$ is the collection of independent sets of a matroid $\mathcal{H}_{n,m}$.

PROOF. Omitted. □

Then we show that this matroid $\mathcal{H}_{n,m}$ is the one determined by the lines in a generic hyperplane arrangement.

THEOREM 2.2. *If a central*[1] *hyperplane arrangement $\mathcal{A} = \{H_1, \ldots, H_m\}$ in $\mathbb{C}^n$ is generic enough, then the matroid of the $\binom{m}{n-1}$ lines H_A is isomorphic to $\mathcal{H}_{n,m}$.*

PROOF. We already observed that the one-dimensional intersections of $\mathcal{A}$ satisfy all the dependence relations of $\mathcal{H}_{n,m}$. Now we wish to show that, if $\mathcal{A}$ is "generic enough", these are the only relations.

It is enough to construct one "generic enough" hyperplane arrangement, and we do it as follows. Consider the m coordinate hyperplanes in $\mathbb{C}^m$, numbered $J_1, \ldots, J_m$. Pick a sufficiently generic n-dimensional subspace V of $\mathbb{C}^m$, and consider the ($(n-1)$-dimensional) hyperplanes $H_1 = J_1 \cap V, \ldots, H_m = J_m \cap V$ in V. The theory of Dilworth truncations of matroids precisely guarantees that V can be chosen in such a way that the lines determined by the H_is satisfy no new relations. We omit the details. □

3. From lines in a flag arrangement to lattice points in a simplex.

Having understood the matroid of lines in a generic hyperplane arrangement, we proceed to study the case of complete flags. In the following two sections, we will describe the matroid of lines of a generic arrangement of d complete flags in $\mathbb{C}^n$. We start, in this section, with a short discussion of the combinatorial

[1] A hyperplane arrangement is *central* if all its hyperplanes go through the origin.

setup that we will use to encode these geometric objects. We then propose, in Section 4, a combinatorial definition of the matroid $\mathcal{T}_{n,d}$, and show that this is, indeed, the matroid we are looking for.

Let $E^1_\bullet, \ldots, E^d_\bullet$ be d generically chosen complete flags in $\mathbb{C}^n$. Write

$$E^k_\bullet = \{\{0\} = E^k_0 \subset E^k_1 \subset \cdots \subset E^k_n = \mathbb{C}^n\},$$

where E^k_i is a vector space of dimension i.

These d flags determine a line arrangement $\mathbf{E}_{n,d}$ in $\mathbb{C}^n$ as follows. Look at all the possible intersections of the subspaces under consideration; they are of the form $E_{a_1,\ldots,a_d} = E^1_{a_1} \cap E^2_{a_2} \cap \cdots \cap E^d_{a_d}$. We are interested in the one-dimensional intersections. Since the $E^k_\bullet$s were chosen generically, $E_{a_1,\ldots,a_d}$ has codimension $(n - a_1) + \ldots + (n - a_d)$ (or n if this sum exceeds n). Therefore, the one-dimensional intersections are the lines $E_{a_1,\ldots,a_d}$ for $a_1 + \cdots + a_d = (d-1)n + 1$. There are $\binom{n+d-2}{d-1}$ such lines, corresponding to the ways of writing $n-1$ as a sum of d nonnegative integers $n - a_1, \ldots, n - a_d$.

Let $T_{n,d}$ be the set of lattice points in the following $(d-1)$-dimensional simplex in $\mathbb{R}^d$:

$$\{\, (x_1, \ldots, x_d) \in \mathbb{R}^d \mid x_1 + \cdots + x_d = n - 1 \text{ and } x_i \geq 0 \text{ for all } i\}.$$

The d vertices of this simplex are $(n-1, 0, 0, \ldots, 0), (0, n-1, 0, \ldots, 0), \ldots, (0, 0, \ldots, n-1)$.

For example, $T_{n,3}$ is simply a triangular array of dots *of size* n; that is, with n dots on each side. We will call $T_{n,d}$ the $(d-1)$-*simplex of size* n.

It will be convenient to identify the line $E_{a_1,\ldots,a_d}$ (where $a_1 + \cdots + a_d = (d-1)n+1$ and $1 \leq a_i \leq n$) with the vector of codimensions $(n - a_1, \ldots, n - a_d)$. This clearly gives us a one-to-one correspondence between the set $T_{n,d}$ and the lines in our line arrangement $\mathbf{E}_{n,d}$.

We illustrate this correspondence for $d = 3$ and $n = 4$ in Figure 2. This picture is easier to visualize in real projective 3-space. Now each one of the flags $E_\bullet$, $F_\bullet$, and $G_\bullet$ is represented by a point in a line in a plane. The lines in our line arrangement are now the 10 intersection points we see in the picture.

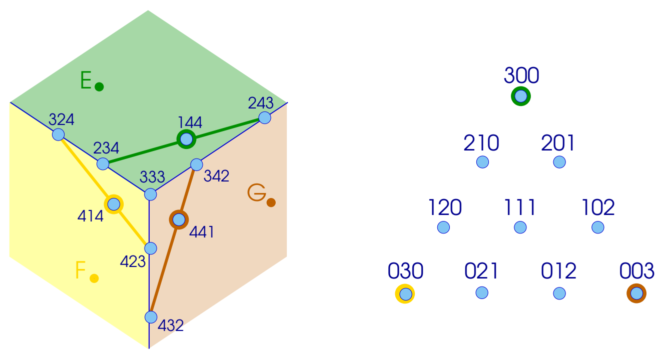


Figure 2. The lines determined by three flags in $\mathbb{C}^4$, and the array $T_{4,3}$.

We are interested in the dependence relations among the lines in the line arrangement $\mathbf{E}_{n,d}$. As in the case of hyperplane arrangements, there are several *combinatorial relations* which arise as follows. Consider a k-dimensional subspace $E_{b_1,\ldots,b_d}$ with $b_1 + \cdots + b_d = (d-1)n + k$. Every line of the form $E_{a_1,\ldots,a_d}$ with $a_i \leq b_i$ is in this subspace, so no $k+1$ of them can be independent. The corresponding points $(n - a_1, \ldots, n - a_d)$ are the lattice points inside a parallel translate of $T_{k,d}$, the simplex of size k, in $T_{n,d}$. In other words, in a set of independent lines of our arrangement, we cannot have more than k lines whose corresponding dots are in a simplex of size k in $T_{n,d}$.

For example, no four of the lines $E_{144}, E_{234}, E_{243}, E_{324}, E_{333}$, and E_{342} are independent, because they are in the 3-dimensional hyperplane E_{344}. The dots corresponding to these six lines form the upper $T_{3,3}$ found in our $T_{4,3}$.

In principle, there could be other hidden dependence relations among the lines in $\mathbf{E}_{n,d}$. The goal of the next section is to show that this is not the case. In fact, these combinatorial relations are the only dependence relations of the line arrangement associated to d generically chosen flags in $\mathbb{C}^n$.

We will proceed as in the case of hyperplane arrangements. We will start by showing that the combinatorial relations do give rise to a matroid $\mathcal{T}_{n,d}$. We will then show that this is, indeed, the matroid we are looking for.

4. The lines in a generic flag arrangement.

We first show that the combinatorial dependence relations defined in Section 3 do determine a matroid.

THEOREM 4.1. *Let $\mathcal{I}_{n,d}$ be the collection of subsets I of $T_{n,d}$ such that every parallel translate of $T_{k,d}$ contains at most k points of I, for every $k \leq n$.*

Then $\mathcal{I}_{n,d}$ is the collection of independent sets of a matroid $\mathcal{T}_{n,d}$ on the ground set $T_{n,d}$.

PROOF. Omitted. □

We now show that the matroid $\mathcal{T}_{n,d}$ of Section 4 is, indeed, the matroid that arises from intersecting d flags in $\mathbb{C}^n$ which are generic enough.

THEOREM 4.2. *If d complete flags $E^1_\bullet, \dots, E^d_\bullet$ in $\mathbb{C}^n$ are generic enough, then the matroid of the $\binom{n+d-2}{d-1}$ lines $E_{a_1,\dots,a_d}$ is isomorphic to $\mathcal{T}_{n,d}$.*

PROOF. As mentioned in Section 3, the one-dimensional intersections of the $E^i_\bullet$s satisfy the following combinatorial relations: each k dimensional subspace $E_{b_1\dots b_d}$ with $b_1 + \cdots + b_d = (d-1)n + k$, contains the lines $E_{a_1\dots a_d}$ with $a_i \leq b_i$; therefore, it is impossible for $k+1$ of these lines to be independent. The subspace $E_{b_1\dots b_d}$ corresponds to the simplex of dots which is labelled $T_{n-b_1,\dots,n-b_d}$, and has size $n - \sum(n-b_i) = k$. The lines $E_{a_1\dots a_d}$ with $a_i \leq b_i$ correspond precisely the dots in this copy of $T_{k,d}$. So these "combinatorial relations" are precisely the dependence relations of $\mathcal{T}_{n,d}$.

Now we need to show that, if the flags are "generic enough", these are the only linear relations among these lines. It is enough to construct one set of flags which satisfies no other relations.

Consider a set $\mathcal{H}$ of $d(n-1)$ hyperplanes H^i_j in $\mathbb{C}^n$ (for $1 \leq i \leq d$ and $1 \leq j \leq n-1$) which are generic in the sense of Theorem 2.2, so the only dependence relations among their one-dimensional intersections are the combinatorial ones. Now, for $i = 1, \dots, d$, define the flag $E^i_\bullet$ by:

$$\begin{aligned} E^i_{n-1} &= H^i_{n-1} \\ E^i_{n-2} &= H^i_{n-1} \cap H^i_{n-2} \\ &\vdots \\ E^i_1 &= H^i_{n-1} \cap H^i_{n-2} \cap \cdots \cap H^i_1, \end{aligned}$$

We show that these d flags are generic enough; in other words, the matroid of their one-dimensional intersections is $\mathcal{T}_{n,d}$. We omit the details. □

With Theorem 4.2 in mind, we will say that the complete flags $E^1_\bullet, \dots, E^d_\bullet$ in $\mathbb{C}^n$ are *matroid-generic* if the matroid of the $\binom{n+d-2}{d-1}$ lines $E_{a_1,\dots,a_d}$ is isomorphic to $\mathcal{T}_{n,d}$.

We conclude this section by showing that the one-dimensional intersections of matroid-generic flag arrangements are completely characterized by their combinatorial properties.

PROPOSITION 4.3. *If a line arrangement $\mathcal{L}$ in $\mathbb{C}^n$ has matroid $\mathcal{T}_{n,d}$, then it can be realized as the arrangement of one-dimensional intersections of d complete flags in $\mathbb{C}^n$.*

PROOF. Omitted. □

5. Rhombus tilings of holey triangles and the matroid $\mathcal{T}_{n,3}$.

Let us change the subject for a moment.

Let $T(n)$ be an equilateral triangle with side length n. Suppose we wanted to tile $T(n)$ using unit rhombi with angles equal to 60° and 120°. It is easy to see that this task is impossible, for the following reason. Cut $T(n)$ into n^2 unit equilateral triangles; $n(n+1)/2$ of these triangles point upward, and $n(n-1)/2$ of them point downward. Since a rhombus always covers one upward and one downward triangle, we cannot use them to tile $T(n)$.

Suppose, then, that we make n holes in the triangle $T(n)$, by cutting out n of the upward triangles. Now we have an equal number of upward and downward triangles, and it may or may not be possible to tile the remaining shape with rhombi.

The main question we address in this section is the following:

Question 5.1. *Given n holes in $T(n)$, is there a simple criterion to determine whether there exists a rhombus tiling of the holey triangle that remains?*

A rhombus tiling is equivalent to a perfect matching between the upward triangles and the downward triangles. Hall's theorem then gives us an answer to Question 5.1: It is necessary and sufficient that any k downward triangles have a total of at least k upward triangles to match to.

However, the geometry of $T(n)$ allows us to give a simpler criterion. Furthermore, in view of Theorem 4.1, this criterion reveals an unexpected connection between these rhombus tilings and the line arrangement determined by 3 generically chosen flags in $\mathbb{C}^n$.

Theorem 5.2. *Let S be a set of n holes in $T(n)$. The triangle $T(n)$ with holes at S can be tiled with rhombi if and only if every $T(k)$ in $T(n)$ contains at most k holes, for all $k \leq n$.*

Proof. Omitted. ☐

Corollary 5.3. *The possible locations of n holes for which a rhombus tiling of the holey triangle $T(n)$ exists correspond to the bases of the matroid $\mathcal{T}_{n,3}$.*

Proof. This is just a restatement of Theorem 5.2. ☐

Corollary 5.3 allows us to say more about the structure of the matroid $\mathcal{T}_{n,3}$. We first remind the reader of the definition of an important family of matroids, called *cotransversal* matroids. For more information, we refer the reader to [**13**].

Let G be a directed graph with vertex set V, and let $A = \{v_1, \ldots, v_r\}$ be a subset of V. We say that an r-subset B of V *can be linked to* A if there exist r vertex-disjoint directed paths whose initial vertex is in B and whose final vertex is in A. We will call these r paths a *routing* from B to A. The collection of r-subsets which can be linked to A are the bases of a matroid denoted $L(G, A)$. Such a matroid is called a *strict gammoid* or a *cotransversal* matroid.

Theorem 5.4. *The matroid $\mathcal{T}_{n,3}$ is cotransversal.*

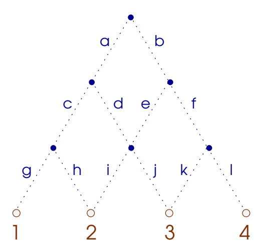


Figure 3. The graph G_4.

Proof. Let G_n be the directed graph whose set of vertices is the triangular array $T_{n,3}$, where each dot not on the bottom row is connected to the two dots directly below it. Label the dots on the bottom row $1, 2, \ldots, n$. Figure 3 shows G_4; all the edges of the graph point down.

There is a bijection between the rhombus tilings of the holey triangles of size n, and the routings (sets of n non-intersecting paths) in the graph G_n which end at vertices $1, 2, \ldots, n$. This correspondence is best understood in an example; see Figure 4. We leave it to the reader to check the details.

In this correspondence, the holes of the holey triangle correspond to the starting points of the n paths in the graph. From Corollary 5.3, it follows that $\mathcal{T}_{n,3}$ is the cotransversal matroid $L(G_n, [n])$. ☐

Theorem 5.5. *Assign algebraically independent weights to the edges of G_n.*[2] *For each dot D in the triangular array $T_{n,3}$ and each $1 \leq i \leq n$, let $v_{D,i}$ be the sum of the weights of all paths*[3] *from dot D to dot i on the bottom row.*

Then the path vectors $v_D = (v_{D,1}, \ldots, v_{D,n})$ *are a geometric representation of the matroid $T_{n,3}$.*

[2]Integer weights which increase extremely quickly will also work.

[3]The weight of a path is defined to be the product of the weights of its edges.

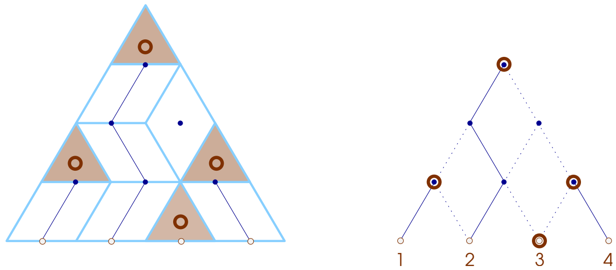


FIGURE 4. A tiling of a holey $T(4)$ and the corresponding routing of G_4.

For example, the top dot of $T_{4,3}$ in Figure 3 would be assigned the *path vector* $(acg, ach+adi+bei, adj+bej+bfk, bfl)$ Similarly, focusing our attention on the top three rows, the representation we obtain for the matroid $\mathcal{T}_{3,3}$ is given by the columns of the following matrix:

$$\begin{pmatrix} 1 & 0 & 0 & c & 0 & ac \\ 0 & 1 & 0 & d & e & ad+be \\ 0 & 0 & 1 & 0 & f & bf \end{pmatrix}$$

PROOF OF THEOREM 5.5. This is a consequence of the Lindström-Gessel-Viennot lemma [**7, 8, 10, 12**]; we omit the details. $\square$

The very simple and explicit representation of $\mathcal{T}_{n,3}$ of Theorem 5.5 will be shown in Section 8 to have an unexpected consequence in the Schubert calculus: it provides us with a reasonably efficient method for computing Schubert structure constants in the flag variety.

6. Fine mixed subdivisions of $n\Delta_{d-1}$ and triangulations of $\Delta_{n-1} \times \Delta_{d-1}$.

The surprising relationship between the geometry of three flags in $\mathbb{C}^n$ and the rhombus tilings of holey triangles is useful to us in two ways: it explains the structure of the matroid $\mathcal{T}_{n,3}$, and it clarifies the conditions for a rhombus tiling of such a region to exist. We now investigate a similar connection between the geometry of d flags in $\mathbb{C}^n$, and certain well-studied $(d-1)$-dimensional analogs of these tilings.

Instead of thinking of rhombus tilings of a holey triangle, it will be slightly more convenient to think of them as *lozenge tilings* of the triangle: these are the tilings of the triangle using unit rhombi and upward unit triangles. A good high-dimensional analogue of the lozenge tilings of the triangle $n\Delta_2$ are the *fine mixed subdivisions* of the simplex $n\Delta_{d-1}$; we briefly recall their definition. Define a *fine mixed cell* of the simplex Δ_{d-1} to be a Minkowski sum $B_1 + \cdots + B_n$, where the B_is are faces of Δ_{d-1} which lie in independent affine subspaces, and whose dimensions add up to $d-1$. A *fine mixed subdivision* of $n\Delta_{d-1}$ is a subdivision of $n\Delta_{d-1}$ into fine mixed cells[**15**, Theorem 2.6].

In the same way that we identified arrays of triangles with triangular arrays of dots in Section 5, we can identify the array of possible locations of the simplices in $n\Delta_{d-1}$ with the array of dots $T_{n,d}$ defined in Section 3. A conjectural generalization of Corollary 5.3, which we now state, would show that fine mixed subdivisions of $n\Delta_{d-1}$ are also closely connected to the matroid $\mathcal{T}_{n,d}$.

CONJECTURE 6.1. *The possible locations of the simplices in a fine mixed subdivision of $n\Delta_{d-1}$ are precisely the bases of the matroid $\mathcal{T}_{n,d}$.*

In the remainder of this section, we will give a completely combinatorial description of the fine mixed subdivisions of $n\Delta_{d-1}$. We will use this description to prove one direction of this conjecture in Section 7.

We start by recalling the one-to-one correspondence between the fine mixed subdivisions of $n\Delta_{d-1}$ and the triangulations of the polytope $\Delta_{n-1} \times \Delta_{d-1}$. This equivalent point of view has the drawback of bringing us to a higher-dimensional picture. Its advantage is that it simplifies greatly the combinatorics of the tiles, which are now just simplices.

Let $v_1, \ldots, v_n$ and $w_1, \ldots, w_d$ be the vertices of Δ_{n-1} and Δ_{d-1}, so that the vertices of $\Delta_{n-1} \times \Delta_{d-1}$ are of the form $v_i \times w_j$. A triangulation T of $\Delta_{n-1} \times \Delta_{d-1}$ is given by a collection of simplices. For each simplex t in T, consider the fine mixed cell whose i-th summand is $w_a w_b \ldots w_c$, where $a, b, \ldots, c$ are the indexes j such that $v_i \times w_j$ is a vertex of t. These fine mixed cells constitute the fine mixed subdivision of

$n\Delta_{d-1}$ corresponding to T. (This bijection is only a special case of the more general Cayley trick, which is discussed in detail in [**15**].)

For instance, Figure 5 shows a triangulation of the triangular prism $\Delta_1 \times \Delta_2 = 12 \times ABC$, and the corresponding fine mixed subdivision of $2\Delta_2$, whose three tiles are $ABC + B$, $AC + AB$, and $C + ABC$.

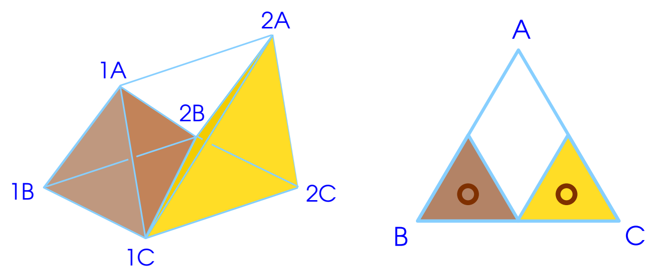


Figure 5. The Cayley trick.

Consider the complete bipartite graph $K_{n,d}$ whose vertices are $v_1, \ldots, v_n$ and $w_1, \ldots, w_d$. Each vertex of $\Delta_{n-1} \times \Delta_{d-1}$ corresponds to an edge of $K_{n,d}$. The vertices of each simplex in $\Delta_{n-1} \times \Delta_{d-1}$ determine a subgraph of $K_{n,d}$. Each triangulation of $\Delta_{n-1} \times \Delta_{d-1}$ is then encoded by a collection of subgraphs of $K_{n,d}$. Figure 6 shows the three trees that encode the triangulation of Figure 5.

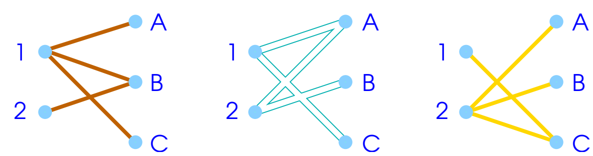


Figure 6. The trees corresponding to the triangulation of Figure 5.

Our next result is a combinatorial characterization of the triangulations of $\Delta_{n-1} \times \Delta_{d-1}$.

Proposition 6.2. *A collection of subgraphs $t_1, \ldots, t_k$ of $K_{n,d}$ encodes a triangulation of $\Delta_{n-1} \times \Delta_{d-1}$ if and only if:*

(1) *Each t_i is a spanning tree.*
(2) *For each t_i and each internal[4] edge e of t_i, there exists an edge f and a tree t_j with $t_j = t_i - e \cup f$.*
(3) *There do not exist two trees t_i and t_j, and a circuit C of $K_{n,d}$ which alternates between edges of t_i and edges of t_j.*

Proof. Omitted. $\square$

In light of Proposition 6.2, we will call a collection of spanning trees satisfying the above properties a triangulation of $\Delta_{n-1} \times \Delta_{d-1}$.

7. Subdivisions of $n\Delta_{d-1}$ and the matroid $\mathcal{T}_{n,d}$.

Having given a combinatorial characterization of the triangulations of the polytope $\Delta_{n-1} \times \Delta_{d-1}$ in Proposition 6.2, we are now in a position to prove the forward direction of Conjecture 6.1, which relates these triangulations to the matroid $\mathcal{T}_{n,d}$. The following combinatorial lemma will play an important role in our proof.

Proposition 7.1. *Let n, d, and $a_1, \ldots, a_d$ be non-negative integers such that $a_1 + \cdots + a_d \leq n - 1$. Suppose we have a coloring of the $n(n-1)$ edges of the directed complete graph K_n with d colors, such that each color defines a poset on $[n]$; in other words,*

(a) *the edges $u \to v$ and $v \to u$ have different colors, and*
(b) *if $u \to v$ and $v \to w$ have the same color, then $u \to w$ has that same color.*

[4] An edge of a tree is *internal* if it is not a leaf.

Call a vertex v outgoing *if, for every i, there exist at least a_i vertices w such that $v \to w$ has color i. Then the number of outgoing vertices is at most $n - a_1 - \cdots - a_d$.*

PROOF. Omitted. The intuition is the following. We have d poset structures on the set $[n]$, and this statement essentially says that we cannot have "too many" elements which are "very large" in all the posets. $\square$

We have now laid down the necessary groundwork to prove one direction of Conjecture 6.1.

PROPOSITION 7.2. *In any fine mixed subdivision of $n\Delta_{d-1}$,*
(a) *there are exactly n tiles which are simplices, and*
(b) *the locations of the n simplices give a basis of the matroid $\mathcal{T}_{n,d}$.*

PROOF OF PROPOSITION 7.2. Let us look back at the way we defined the correspondence between a triangulation T of $\Delta_{n-1} \times \Delta_{d-1}$ and a fine mixed subdivision $f(T)$ of $n\Delta_{d-1}$. It is clear that the simplices $f(t)$ of $f(T)$ arise from those simplices t of T whose vertices are $v_i \times w_1, \ldots, v_i \times w_d$ (for some i), and one $v_j \times w_{g(j)}$ for each $j \neq i$. Furthermore, the location of $f(t)$ in $n\Delta_{d-1}$ is given by the sum of the $w_{g(j)}$s.

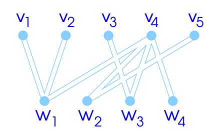


FIGURE 7. A spanning tree of $K_{5,4}$.

For instance the spanning tree of $K_{5,4}$ shown in Figure 7 gives rise to a simplex in a fine mixed subdivision of $5\Delta_3 = 5w_1w_2w_3w_4$ given by the Minkowski sum $w_1 + w_1 + w_3 + w_1w_2w_3w_4 + w_2$. The location of this simplex in $5\Delta_3$ corresponds to the point $(2, 1, 1, 0)$ of $T_{5,4}$, because the Minkowski sum above contains two w_1 summands, one w_2, and one w_3.

The simplices of the fine mixed subdivision of $n\Delta_{d-1}$ come from spanning trees t of $K_{n,d}$ for which one vertex v_i has degree d and the other v_js have degree 1. The coordinates of the location of $f(t)$ in $n\Delta_{d-1}$ are simply $(\deg_t w_1 - 1, \ldots, \deg_t w_d - 1)$. Call such a simplex, and the corresponding tree, *i-pure*. For instance, in Figure 5, there is a 1-pure tree and a 2-pure tree, which give simplices in locations $(0, 1, 0)$ and $(0, 0, 1)$ of $2\Delta_2$, respectively.
Proof of (a). We prove that in a triangulation T of $\Delta_{n-1} \times \Delta_{d-1}$ there is exactly one i-pure simplex for each i with $1 \leq i \leq n$. The details are omitted.
Proof of (b). The idea is to construct a coloring of the directed complete graph K_n which economically stores a description of the n pure trees, and invoke Proposition 7.1. Again, we omit the details. $\square$

For the converse of Conjecture 6.1, we would need to show that every basis of $\mathcal{T}_{n,d}$ arises from a fine mixed subdivision of $n\Delta_{d-1}$. We conjecture a stronger result.

CONJECTURE 7.3. *For any basis B of $\mathcal{T}_{n,d}$, there is a* **coherent** *fine mixed subdivision of $n\Delta_{d-1}$ whose n simplices are located at B.*

Given the correspondence between coherent fine mixed subdivisions of $n\Delta_{d-1}$ and the combinatorial types of arrangements of d generic tropical hyperplanes in tropical $(n-1)$-space [**4, 15**], Conjecture 7.3 is an invitation to study more closely those combinatorial types. This can naturally be thought of as the study of *tropical oriented matroids.*

8. Applications to Schubert calculus.

In this section, we show some of the implications of our work in the Schubert calculus of the flag variety. Throughout this section, we will assume some familiarity with the Schubert calculus, though we will recall some of the definitions and conventions that we will use; for more information, see for example [**6, 11**]. We will also need some of the results of Eriksson and Linusson [**5**] and Billey and Vakil [**2**] on Schubert varieties and permutation arrays.

Eriksson and Linusson [**5**] introduced certain higher-dimensional analogs of permutation matrices, called *permutation arrays.* A permutation array is an array of dots in the cells of a d-dimensional $n \times n \times \cdots \times n$

box, satisfying some quite restrictive properties. From a permutation array P, via a simple combinatorial rule, one can construct a *rank array* of integers, also of shape $[n]^d$. We denote it $\operatorname{rk} P$.

This definition is motivated by the observation that the relative position of d flags $E^1_\bullet, \ldots, E^d_\bullet$ in $\mathcal{F}\ell_n$ is described by a unique *permutation array* P, via the equations

$$\dim\bigl(E^1_{x_1} \cap \cdots \cap E^d_{x_d}\bigr) = \operatorname{rk} P[x_1, \ldots, x_d] \qquad \text{for all } \ 1 \leq x_1, \ldots, x_d \leq n.$$

This result initiated the study of *permutation array schemes*, which generalize Schubert varieties in the flag variety $\mathcal{F}\ell_n$.

The relative position of d generic flags is described by the *transversal permutation array*

$$\Bigl\{(x_1, \ldots, x_d) \in [n]^d \mid \sum_{i=1}^{d} x_i = (d-1)n + 1\Bigr\}.$$

The dot at position $(x_1, \ldots, x_d)$ represents a one-dimensional intersection $E^1_{x_1} \cap \cdots \cap E^d_{x_d}$. Naturally, we identify the dots in the transversal permutation array with the corresponding element of the matroid $\mathcal{T}_{n,d}$.

Given a fixed flag $E_\bullet$ in $\mathbb{C}^n$ and a permutation w in S_n, denote the *Schubert cell* and *Schubert variety* by

$$\begin{aligned} X^\circ_w(E_\bullet) &= \{F_\bullet \mid E_\bullet \text{ and } F_\bullet \text{ have relative position } w\} \\ &= \{F_\bullet \mid \dim(E_i \cap F_j) = \operatorname{rk} w[i,j] \ \text{ for all } \ 1 \leq i, j \leq n.\}, \text{ and} \\ X_w(E_\bullet) &= \{F_\bullet \mid \dim(E_i \cap F_j) \geq \operatorname{rk} w[i,j] \ \text{ for all } \ 1 \leq i, j \leq n.\}, \end{aligned}$$

respectively.

A *Schubert problem* asks for the number of flags $F_\bullet$ whose relative positions with respect to d given fixed flags $E^1_\bullet, \ldots, E^d_\bullet$ are given by the permutations $w^1, \ldots, w^d$. This question only makes sense when

$$X = X_{w^1}(E^1_\bullet) \cap \cdots \cap X_{w^d}(E^d_\bullet)$$

is 0-dimensional; that is, when $l(w^1) + \cdots + l(w^d) = \binom{n}{2}$. If $E^1_\bullet, \ldots, E^d_\bullet$ are sufficiently generic, the intersection X has a fixed number of points $c_{w^1 \ldots w^d}$ which only depends on the permutations $w^1, \ldots, w^d$.

This question is a fundamental one for several reasons. The numbers $c_{w^1 \ldots w^d}$ which answer this question appear in several different contexts. For instance, the cycles $[X_w]$ corresponding to the Schubert varieties form a $\mathbb{Z}$-basis for the cohomology ring of the flag variety $\mathcal{F}\ell_n$, and the numbers c_{uvw} are the multiplicative structure constants. (For this reason, if we know the answer to all Schubert problems with $d = 3$, we can easily obtain them for higher d.) The analogous structure constants in the Grassmannian are the Littlewood-Richardson coefficients, which are much better understood. For instance, even though the c_{uvw}s are known to be positive integers, it is a long standing open problem to find a combinatorial interpretation of them.

Billey and Vakil [**2**] showed that the permutation arrays of Eriksson and Linusson can be used to explicitly intersect Schubert varieties, and compute the numbers $c_{w^1 \ldots w^d}$.

THEOREM 8.1. *(Billey-Vakil, [**2**]) Suppose that*

$$X = X_{w^1}(E^1_\bullet) \cap \cdots \cap X_{w^d}(E^d_\bullet)$$

is a 0-dimensional and nonempty intersection, with $E^1_\bullet, \ldots, E^d_\bullet$ generic.

(1) *There exists a unique permutation array $P \subset [n]^{d+1}$, easily constructed from $w^1, \ldots, w^d$, such that*

$$\dim\bigl(E^1_{x_1} \cap \cdots \cap E^d_{x_d} \cap F_{x_{d+1}}\bigr) = rk\, P[x_1, \ldots, x_d, x_{d+1}],$$

for all $F_\bullet \in X$ and all $1 \leq x_1, \ldots, x_{d+1} \leq n$.

(2) *Given the permutation array P, and a vector $v_{a_1,\ldots,a_d}$ in each one-dimensional intersection $E_{a_1,\ldots,a_d} = E^1_{a_1} \cap \cdots \cap E^d_{a_d}$, we can write down an explicit set of polynomial equations defining X.*

Theorem 8.1 highlights the importance of studying the line arrangements $\mathbf{E}_{n,d}$ determined by intersecting d generic complete flags in $\mathbb{C}^n$. In principle, if we are able to construct such a line arrangement, we can compute the structure constants c_{uvw} for any $u, v, w \in S_n$. (In practice, we still have to solve the system of polynomial equations, which is not easy for large n.) Let us make two observations in this direction.

8.1. Matroid genericity versus Schubert genericity. We have been talking about the line arrangement $\mathbf{E}_{n,d}$ determined by a generic flag arrangement $E^1_\bullet, \ldots, E^d_\bullet$ in $\mathbb{C}^n$. We need to be careful, because we have given two different meanings to the word *generic*.

In Sections 3 and 4, we have shown that, if $E^1_\bullet, \ldots, E^d_\bullet$ are sufficiently generic, then the linear dependence relations in the line arrangement $\mathbf{E}_{n,d}$ are described by a fixed matroid $\mathcal{T}_{n,d}$. Let us say that the flags are *matroid-generic* if this is the case.

Recall that in the Schubert problem described by permutations $w^1, \ldots, w^d$ with $\sum l(w^i) = \binom{n}{2}$, the 0-dimensional intersection

$$X = X_{w^1}(E^1_\bullet) \cap \cdots \cap X_{w^d}(E^d_\bullet)$$

contains a fixed number of points $c_{w^1 \ldots w^d}$, provided that $E^1_\bullet, \ldots, E^d_\bullet$ are sufficiently generic. Let us say that the flags are *Schubert-generic* if they are sufficiently generic for any Schubert problem.

These notions depend only on the line arrangement $\mathbf{E}_{n,d}$. The line arrangement $\mathbf{E}_{n,d}$ is matroid-generic if its matroid is $\mathcal{T}_{n,d}$, and it is Schubert-generic if the equations of Theorem 8.1 give the correct number of solutions to every Schubert problem.

Our characterization of matroid-generic line arrangements (*i.e.*, our description of the matroid $\mathcal{T}_{n,d}$) does not tell us how to construct a Schubert-generic line arrangement. However, when $d = 3$ (which is the interesting case in the Schubert calculus),the cotransversality of the matroid $\mathcal{T}_{n,3}$ allows us to present such a line arrangement explicitly.

PROPOSITION 8.2. *The $\binom{n}{2}$ path vectors of Theorem 5.5 are Schubert-generic.*

PROOF. Omitted. □

Proposition 8.2 shows that when we plug the path vectors into the polynomial equations of Theorem 8.1, and compute the intersection X, we will have $|X| = c_{uvw}$. The advantage of this point of view is that the equations are now written in terms of combinatorial objects, without any reference to an initial choice of flags.

PROBLEM 8.3. *Interpret combinatorially the c_{uvw} solutions of the above system of equations, thereby obtaining a combinatorial interpretation for the structure constants c_{uvw}.*

8.2. A criterion for vanishing Schubert structure constants. Consider the Schubert problem

$$X = X_{w^1}(E^1_\bullet) \cap \cdots \cap X_{w^d}(E^d_\bullet).$$

Let $P \in [n]^{d+1}$ be the permutation array which describes the dimensions $\dim(E^1_{x_1} \cap \cdots \cap E^d_{x_d} \cap F_{x_{d+1}})$ for any flag $F_\bullet \in X$. Let $P_1, \ldots, P_n$ be the n "floors" of P, corresponding to $F_1, \ldots, F_n$, respectively. Each one of them is itself a permutation array of shape $[n]^d$.

Billey and Vakil proposed a simple criterion which is very efficient in detecting that many Schubert structure constants are equal to zero.

PROPOSITION 8.4. *(Billey-Vakil, [**2**]) If P_n is not the transversal permutation array, then $X = \emptyset$ and $c_{w^1 \ldots w^d} = 0$.*

Knowing the structure of the matroid $\mathcal{T}_{n,d}$, we can strengthen this criterion as follows.

PROPOSITION 8.5. *Suppose P_n is the transversal permutation array, and identify it with the set $T_{n,d}$. If, for some k, the rank of $P_k \cap P_n$ in $\mathcal{T}_{n,d}$ is greater than k, then $X = \emptyset$ and $c_{w^1 \ldots w^d} = 0$.*

PROOF. Each dot in P_n corresponds to a one-dimensional intersection of the form $E^1_{x_1} \cap \cdots \cap E^d_{x_d}$. Therefore, each dot in $P_k \cap P_n$ corresponds to a line that F_k is supposed to contain, if $F_\bullet$ is a solution to the Schubert problem. The rank of $P_k \cap P_n$ is the dimension of the subspace spanned by those lines; if $F_\bullet$ exists, that dimension must be at most k. □

Let us see how to apply Proposition 8.5 in an example. Following the algorithm of [**2**], the permutations $u = v = w = 213$ in S_3 give rise to the four-dimensional permutation array consisting of the dots $(3,3,1,1)$, $(1,3,3,2)$, $(3,1,3,2)$, $(3,3,1,2)$, $(1,3,3,3)$, $(2,2,3,3)$, $(2,3,2,3)$, $(3,1,3,3)$, $(3,2,2,3)$, and $(3,3,1,3)$. We follow [**5, 18**] in representing it as follows:

		1

		3
3		1

		3
	3	2
3	2	1

The three boards shown represent the three-dimensional floors P_1, P_2, and P_3 of P, form left to right. In each one of them, a dot in cell (i, j, k) is represented in two dimensions by a number k in cell (i, j).

It takes some practice to interpret these tables; but once one is used to them, it is very easy to proceed. Simply notice that $P_2 \cap P_3$ is a set of rank 3 in the matroid $\mathcal{T}_{3,3}$, and we are done! We conclude that $c_{213,213,213} = 0$. For $n = 3$, this is the only vanishing c_{uvw} which is not explained by Proposition 8.4.

We remark that there are other methods for detecting the vanishing of Schubert structure constants, due to Knutson, Lascoux and Schutzenberger, and Purbhoo. In comparing these methods for small values of n, we have found Proposition 8.5 to be quicker and simpler, but less complete than some of these methods.

However, Proposition 8.5 is only the very first observation that we can make from our understanding of the structure of $\mathcal{T}_{n,d}$. Our argument can be easily fine-tuned to explain all vanishing Schubert structure constants with $n \leq 5$. A systematic way of doing this in general would be very desirable.

9. Acknowledgments

We would like to thank Laci Lovasz, Jim Propp, David Wilson, and Andrei Zelevinsky for very helpful discussions.

References

[1] E. Babson and L. Billera. The geometry of products of minors. *Discrete Comput. Geom.* **20** (1998), 231-249.

[2] S. Billey and R. Vakil. Intersections of Schubert varieties and other permutation array schemes. Preprint, `arXiv:math.AG/0502468`, 2005.

[3] J. A. de Loera, J. Rambau, and F. Santos. *Triangulations of polyhedra and point sets*, monograph in preparation, 2005.

[4] M. Develin and B. Sturmfels. Tropical convexity. *Documenta Math.* **9** (2004) 1-27.

[5] K. Eriksson and S. Linusson. A decomposition of $\mathcal{F}\ell(n)^d$ indexed by permutation arrays. *Adv. in Appl. Math.* **25** (2000) 212-227.

[6] W. Fulton. *Young tableux, with applications to representation theory and geometry.* London Math. Soc. Student Texts 35, Cambridge, 1997.

[7] I. Gessel and X. Viennot. Binomial determinants, paths and hook formulae. *Adv. Math* **58** (1985) 300-321.

[8] S. Karlin and G. MacGregor. Coincidence probabilities. *Pacific J. Math.* **9** (1959) 1141-1164.

[9] A. Knutson, T. Tao, and C. Woodward. The honeycomb model of GL_n tensor products II: Facets of the Littlewood-Richardson cone. *J. Amer. Math. Soc.* **17** (2004) 19-48.

[10] B. Lindström. On the vector representations of induced matroids. *Bull. London Math. Soc.* **5** (1973) 85-90.

[11] L. Manivel. *Symmetric functions, Schubert polynomials and degeneracy loci,* transl. by J. Swallow. SMF/AMS Texts and Monographs, vol. 6, AMS, Providence, 2001.

[12] J. Mason On a class of matroids arising from paths in graphs. *Proc. London Math. Soc.* (3) **25** (1972) 55-74.

[13] J. G. Oxley. *Matroid theory.* Oxford University Press. New York, 1992.

[14] J. Propp. Enumeration of matchings: problems and progress, in *New perspectives in algebraic combinatorics*, Math. Sci. Res. Inst. Publ., vol. 38, Cambridge, 1999.

[15] F. Santos. The Cayley trick and triangulations of simplices, in *Integer Points in Polyhedra – Geometry, Number Theory, Algebra, Optimization*, edited by A. Barvinok, M. Beck, C. Haase, B. Reznick, and V. Welker, American Mathematical Society, 2005.

[16] F. Santos. Triangulations of oriented matroids. *Memoirs of the American Mathematical Society* **156** (2002), Number 741.

[17] B. Sturmfels. *Solving systems of polynomial equations.* CBMS Regional Conference Series in Mathematics, vol. 97. American Mathematical Society, Providence, 2002.

[18] R. Vakil. A geometric Littlewood-Richardson rule. Preprint, `arXiv:math.AG/0302294`, 2003, to appear in *Annals of Math.*

Department of Mathematics, University of Washington, Seattle, WA, USA.
E-mail address: `fardila@alum.mit.edu`

Department of Mathematics, University of Washington, Seattle, WA, USA.
E-mail address: `billey@math.washington.edu`

Finite generation of symmetric ideals

Matthias Aschenbrenner and Christopher J. Hillar

In memoriam Karin Gatermann (1965–2005).

Abstract. Let A be a commutative Noetherian ring, and let $R = A[X]$ be the polynomial ring in an infinite collection X of indeterminates over A. Let $\mathfrak{S}_X$ be the symmetric group of X. The group $\mathfrak{S}_X$ acts on R in a natural way, and this in turn gives R the structure of a left module over the group ring $R[\mathfrak{S}_X]$. We prove that all ideals of R invariant under the action of $\mathfrak{S}_X$ are finitely generated as $R[\mathfrak{S}_X]$-modules. The proof involves introducing a certain partial order on monomials and showing that it is a well-quasi-ordering. We also consider the concept of an invariant chain of ideals for finite-dimensional polynomial rings and relate it to the finite generation result mentioned above. Finally, a motivating question from chemistry is presented, with the above framework providing a suitable context in which to study it.

Résumé. Soit A un anneau Noetherien commutatif, et $R = A[X]$ l'anneau des polynomes en une infinité d'indéterminées X sur A. Soit $\mathfrak{S}_X$ le groupe symétrique de X. Le groupe $\mathfrak{S}_X$ agit sur R de manière naturelle, ce qui donne à R la structure d'un module gauche sur l'anneau $R[\mathfrak{S}_X]$. Nous prouvons que tous les idéaux de R invariants sous l'action de $\mathfrak{S}_X$ sont finitement engendrés comme $R[\mathfrak{S}_X]$-modules. La démonstration utilise le fait qu'un certain ordre partiel sur les monomes est un quasi-ordre. Nous utilisons aussi le concept de chaˆ***i***ne invariante des idéaux pour les anneaux de polynômes de dimension finie, que nous relions au résultat de génération finie mentionné plus haut. Finalement, nous présentons une motivation pour notre travail issue de la chimie.

1. Introduction

A pervasive theme in invariant theory is that of finite generation. A fundamental example is a theorem of Hilbert stating that the invariant subrings of finite-dimensional polynomial algebras over finite groups are finitely generated [**5**, Corollary 1.5]. In this article, we study invariant ideals of infinite-dimensional polynomial rings. Of course, when the number of indeterminates is finite, Hilbert's basis theorem tells us that any ideal (invariant or not) is finitely generated.

Our setup is as follows. Let X be an infinite collection of indeterminates, and let $\mathfrak{S}_X$ be the group of permutations of X. Fix a commutative Noetherian ring A and let $R = A[X]$ be the polynomial ring in the indeterminates X. The group $\mathfrak{S}_X$ acts naturally on R: if $\sigma \in \mathfrak{S}_X$ and $f \in A[x_1, \dots, x_n]$ where $x_i \in X$, then

$$\sigma f(x_1, x_2, \dots, x_n) = f(\sigma x_1, \sigma x_2, \dots, \sigma x_n) \in R.$$

This in turn gives R the structure of a left module over the (non-commutative) group ring $R[\mathfrak{S}_X]$. An ideal $I \subseteq R$ is called *invariant under* $\mathfrak{S}_X$ (or simply *invariant*) if

$$\mathfrak{S}_X I := \{\sigma f : \sigma \in \mathfrak{S}_X, f \in I\} \subseteq I.$$

Notice that invariant ideals are simply the $R[\mathfrak{S}_X]$-submodules of R. We may now state our main result.

1991 *Mathematics Subject Classification.* 13E05, 13E15, 20B30, 06A07.

Key words and phrases. Invariant ideal, well-quasi-ordering, symmetric group.

The first author is partially supported by the National Science Foundation Grant DMS 03-03618. The work of the second author is supported under a National Science Foundation Graduate Research Fellowship.

Theorem 1.1. *Every ideal of $R = A[X]$ invariant under $\mathfrak{S}_X$ is finitely generated as an $R[\mathfrak{S}_X]$-module. (Stated more succinctly, R is a Noetherian $R[\mathfrak{S}_X]$-module.)*

For the purposes of this work, we will use the following notation. Let B be a ring and let G be a subset of a B-module M. Then $\langle f : f \in G\rangle_B$ will denote the B-submodule of M generated by elements of G.

Example 1.2. Suppose that $X = \{x_1, x_2, \ldots\}$. The invariant ideal $I = \langle x_1, x_2, \ldots\rangle_R$ is clearly not finitely generated over R, however, it does have the compact representation $I = \langle x_1\rangle_{R[\mathfrak{S}_X]}$.

The outline of this paper is as follows. In Section 2, we define a partial order on monomials and show that it can be used to obtain a well-quasi-ordering of the monomials in R. Section 3 then goes on to detail our proof of Theorem 1.1, using the main result of Section 2 in a fundamental way. In the penultimate section, we discuss a relationship between invariant ideals of R and chains of increasing ideals in finite-dimensional polynomial rings. The notions introduced there provide a suitable framework for studying a problem arising from chemistry, the subject of the final section of this article.

2. The Symmetric Cancellation Ordering

We begin this section by briefly recalling some basic order-theoretic notions. We also discuss some fundamental results due to Higman and Nash-Williams and some of their consequences. We define the ordering mentioned in the section heading, and give a sufficient condition for it to be a well-quasi-ordering; this is needed in the proof of Theorem 1.1.

2.1. Preliminaries. A *quasi-ordering* on a set S is a binary relation $\le$ on S which is reflexive and transitive. A *quasi-ordered set* is a pair $(S, \le)$ consisting of a set S and a quasi-ordering $\le$ on S. When there is no confusion, we will omit $\le$ from the notation, and simply call S a quasi-ordered set. If in addition the relation $\le$ is *anti-symmetric* ($s \le t \wedge t \le s \Rightarrow s = t$, for all $s, t \in S$), then $\le$ is called an *ordering* (sometimes also called a *partial ordering*) on the set S. The *trivial* ordering on S is given by $s \le t \iff s = t$ for all $s, t \in S$. A quasi-ordering $\le$ on a set S induces an ordering on the set $S/{\sim} = \{s/{\sim} : s \in S\}$ of equivalence classes of the equivalence relation $s \sim t \iff s \le t \ \wedge\ t \le s$ on S. If s and t are elements of a quasi-ordered set, we write as usual $s \le t$ also as $t \ge s$, and we write $s < t$ if $s \le t$ and $t \not\le s$.

A map $\varphi\colon S \to T$ between quasi-ordered sets S and T is called *increasing* if $s \le t \Rightarrow \varphi(s) \le \varphi(t)$ for all $s, t \in S$, and *strictly increasing* if $s < t \Rightarrow \varphi(s) < \varphi(t)$ for all $s, t \in S$. We also say that $\varphi\colon S \to T$ is a *quasi-embedding* if $\varphi(s) \le \varphi(t) \Rightarrow s \le t$ for all $s, t \in S$.

An *antichain* of S is a subset $A \subseteq S$ such that $s \not\le t$ and $t \not\le s$ for all $s \not\sim t$ in A. A *final segment* of a quasi-ordered set $(S, \le)$ is a subset $F \subseteq S$ which is closed upwards: $s \le t \ \wedge\ s \in F \Rightarrow t \in F$, for all $s, t \in S$. We can view the set $\mathcal{F}(S)$ of final segments of S as an ordered set, with the ordering given by reverse inclusion. Given a subset M of S, the set $\{t \in S : \exists s \in M \text{ with } s \le t\}$ is a final segment of S, the final segment *generated by M*. An *initial segment* of S is a subset of S whose complement is a final segment. An initial segment I of S is *proper* if $I \ne S$. For $a \in S$ we denote by $S^{\le a}$ the initial segment consisting of all $s \in S$ with $s \le a$.

A quasi-ordered set S is said to be *well-founded* if there is no infinite strictly decreasing sequence $s_1 > s_2 > \cdots$ in S, and *well-quasi-ordered* if in addition every antichain of S is finite. The following characterization of well-quasi-orderings is classical (see, for example, [**8**]). An infinite sequence $s_1, s_2, \ldots$ in S is called *good* if $s_i \le s_j$ for some indices $i < j$, and *bad* otherwise.

Proposition 2.1. *The following are equivalent, for a quasi-ordered set S:*

(1) *S is well-quasi-ordered.*
(2) *Every infinite sequence in S is good.*
(3) *Every infinite sequence in S contains an infinite increasing subsequence.*
(4) *Any final segment of S is finitely generated.*
(5) *$(\mathcal{F}(S), \supseteq)$ is well-founded (i.e., the ascending chain condition holds for final segments of S).* □

Let $(S, \le_S)$ and $(T, \le_T)$ be quasi-ordered sets. If there exists an increasing surjection $S \to T$ and S is well-quasi-ordered, then T is well-quasi-ordered, and if there exists a quasi-embedding $S \to T$ and T is well-quasi-ordered, then so is S. Moreover, the cartesian product $S \times T$ can be turned into a quasi-orderd set by using the cartesian product of $\le_S$ and $\le_T$:

$$(s, t) \le (s', t') \quad :\Longleftrightarrow \quad s \le_S s' \wedge t \le_T t', \qquad \text{for } s, s' \in S,\ t, t' \in T.$$

Using Proposition 2.1 we see that the cartesian product of two well-quasi-ordered sets is again well-quasi-ordered.

Of course, a total ordering $\leq$ is well-quasi-ordered if and only if it is well-founded; in this case $\leq$ is called a *well-ordering.* Every well-ordered set is isomorphic to a unique ordinal number, called its *order type.* The order type of $\mathbb{N} = \{0, 1, 2, \ldots\}$ with its usual ordering is ω.

2.2. A lemma of Higman. Given a set X, we let X^* denote the set of all finite sequences of elements of X (including the empty sequence). We may think of the elements of X^* as *non-commutative words* $x_1 \cdots x_m$ with letters $x_1, \ldots, x_m$ coming from the alphabet X. With the concatenation of such words as operation, X^* is the free monoid generated by X. A quasi-ordering $\leq$ on X yields a quasi-ordering $\leq_{\mathrm{H}}$ (the *Higman quasi-ordering*) on X^* as follows:

$$x_1 \cdots x_m \leq_{\mathrm{H}} y_1 \cdots y_n \quad :\Longleftrightarrow \quad \begin{cases} \text{there exists a strictly increasing function} \\ \varphi\colon \{1, \ldots, m\} \to \{1, \ldots, n\} \text{ such that} \\ x_i \leq y_{\varphi(i)} \text{ for all } 1 \leq i \leq m. \end{cases}$$

If $\leq$ is an ordering on X, then $\leq_{\mathrm{H}}$ is an ordering on X^*. The following fact was shown by Higman [**6**] (with an ingenious proof due to Nash-Williams [**12**]):

Lemma 2.2. *If $\leq$ is a well-quasi-ordering on X, then $\leq_{\mathrm{H}}$ is a well-quasi-ordering on X^*.* □

It follows that if $\leq$ is a well-quasi-ordering on X, then the quasi-ordering $\leq^*$ on X^* defined by

$$x_1 \cdots x_m \leq^* y_1 \cdots y_n \quad :\Longleftrightarrow \quad \begin{cases} \text{there exists an injective function} \\ \varphi\colon \{1, \ldots, m\} \to \{1, \ldots, n\} \text{ such} \\ \text{that } x_i \leq y_{\varphi(i)} \text{ for all } 1 \leq i \leq m \end{cases}$$

is also a well-quasi-ordering. (Since $\leq^*$ extends $\leq_{\mathrm{H}}$.)

We also let $X^\diamond$ be the set of *commutative words* in the alphabet X, that is, the free commutative monoid generated by X (with identity element denoted by 1). We sometimes also refer to the elements of $X^\diamond$ as *monomials* (in the set of indeterminates X). We have a natural surjective monoid homomorphism $\pi\colon X^* \to X^\diamond$ given by simply "making the indeterminates commute" (i.e., interpreting a non-commutative word from X^* as a commutative word in $X^\diamond$). Unlike $\leq_{\mathrm{H}}$, the quasi-ordering $\leq^*$ is compatible with π in the sense that $v \leq^* w \Rightarrow v' \leq^* w'$ for all $v, v', w, w' \in X^*$ with $\pi(v) = \pi(v')$ and $\pi(w) = \pi(w')$. Hence $\pi(v) \leq^\diamond \pi(w) :\Longleftrightarrow v \leq^* w$ defines a quasi-ordering $\leq^\diamond$ on $X^\diamond = \pi(X^*)$ making π an increasing map. The quasi-ordering $\leq^\diamond$ extends the divisibility relation in the monoid $X^\diamond$:

$$v|w \quad :\Longleftrightarrow \quad uv = w \text{ for some } u \in X^\diamond.$$

If we take for $\leq$ the trivial ordering on X, then $\leq^\diamond$ corresponds exactly to divisibility in $X^\diamond$, and this ordering is a well-quasi-ordering if and only if X is finite. In general we have, as an immediate consequence of Higman's lemma (since π is a surjection):

Corollary 2.3. *If $\leq$ is a well-quasi-ordering on the set X, then $\leq^\diamond$ is a well-quasi-ordering on $X^\diamond$.* □

2.3. A theorem of Nash-Williams. Given a totally ordered set S and a quasi-ordered set X, we denote by $\mathrm{Fin}(S, X)$ the set of all functions $f\colon I \to X$, where I is a proper initial segment of S, whose range $f(I)$ is *finite.* We define a quasi-ordering $\leq_{\mathrm{H}}$ on $\mathrm{Fin}(S, X)$ as follows: for $f\colon I \to X$ and $g\colon J \to X$ from $\mathrm{Fin}(S, X)$ put

$$f \leq_{\mathrm{H}} g \quad :\Longleftrightarrow \quad \begin{cases} \text{there exists a strictly increasing function } \varphi\colon I \to J \\ \text{such that } f(i) \leq g(\varphi(i)) \text{ for all } i \in I. \end{cases}$$

We may think of an element of $\mathrm{Fin}(S, X)$ as a sequence of elements of X indexed by indices in some proper intial segment of S. So for $S = \mathbb{N}$ with its usual ordering, we can identify elements of $\mathrm{Fin}(\mathbb{N}, X)$ with words in X^*, and then $\leq_{\mathrm{H}}$ for $\mathrm{Fin}(\mathbb{N}, X)$ agrees with $\leq_{\mathrm{H}}$ on X^* as defined above. We will have occasion to use a far-reaching generalization of Lemma 2.2:

Theorem 2.4. *If X is well-quasi-ordered and S is well-ordered, then* $\mathrm{Fin}(S, X)$ *is well-quasi-ordered.* □

This theorem was proved by Nash-Williams [**13**]; special cases were shown earlier in [**4, 11, 14**].

2.4. Term orderings. A *term ordering* of $X^\diamond$ is a well-ordering $\leq$ of $X^\diamond$ such that

(1) $1 \leq x$ for all $x \in X$, and
(2) $v \leq w \Rightarrow xv \leq xw$ for all $v, w \in X^\diamond$ and $x \in X$.

Every ordering $\leq$ of $X^\diamond$ satisfying (1) and (2) extends the ordering $\leq^\diamond$ obtained from the restriction of $\leq$ to X. In particular, $\leq$ extends the divisibility ordering on $X^\diamond$. By the corollary above, a total ordering $\leq$ of $X^\diamond$ which satisfies (1) and (2) is a term ordering if and only if its restriction to X is a well-ordering.

Example 2.5. Let $\leq$ be a total ordering of X. We define the induced *lexicographic ordering* $\leq_{\mathrm{lex}}$ of monomials as follows: given $v, w \in X^\diamond$ we can write $v = x_1^{a_1} \cdots x_n^{a_n}$ and $w = x_1^{b_1} \cdots x_n^{b_n}$ with $x_1 < \cdots < x_n$ in X and all $a_i, b_i \in \mathbb{N}$; then

$$v \leq_{\mathrm{lex}} w \quad :\Longleftrightarrow \quad (a_n, \ldots, a_1) \leq (b_n, \ldots, b_1) \text{ lexicographically (from the left).}$$

The ordering $\leq_{\mathrm{lex}}$ is total and satisfies (1), (2); hence if the ordering $\leq$ of X is a well-ordering, then $\leq_{\mathrm{lex}}$ is a term ordering of $X^\diamond$.

Remark 2.6. Let $\leq$ be a total ordering of X. For $w \in X^\diamond$, $w \neq 1$, we let

$$|w| := \max\{x \in X : x|w\} \quad \text{(with respect to } \leq\text{)}.$$

We also put $|1| := -\infty$ where we set $-\infty < x$ for all $x \in X$. One of the perks of using the lexicographic ordering as a term ordering on $X^\diamond$ is that if v and w are monomials with $v \leq_{\mathrm{lex}} w$, then $|v| \leq |w|$. Below, we often use this observation.

The previous example shows that for every set X there exists a term ordering of $X^\diamond$, since every set can be well-ordered by the Axiom of Choice. In fact, every set X can be equipped with a well-ordering every proper initial segment of which has strictly smaller cardinality than X; in other words, the order type of this ordering (a certain ordinal number) is a cardinal number. We shall call such an ordering of X a *cardinal well-ordering* of X.

Lemma 2.7. *Let X be a set equipped with a cardinal well-ordering, and let I be a proper initial segment of X. Then every injective function $I \to X$ can be extended to a permutation of X.*

Proof. Since this is clear if X is finite, suppose that X is infinite. Let $\varphi \colon I \to X$ be injective. Since I has cardinality $|I| < |X|$ and X is infinite, we have $|X| = \max\{|X \setminus I|, |I|\} = |X \setminus I|$. Similarly, since $|\varphi(I)| = |I| < |X|$, we also have $|X \setminus \varphi(I)| = |X|$. Hence there exists a bijection $\psi \colon X \setminus I \to X \setminus \varphi(I)$. Combining φ and ψ yields a permutation of X as desired. □

2.5. A new ordering of monomials. Let G be a permutation group on a set X, that is, a group G together with a faithful action $(\sigma, x) \mapsto \sigma x \colon G \times X \to X$ of G on X. The action of G on X extends in a natural way to a faithful action of G on $X^\diamond$: $\sigma w = \sigma x_1 \cdots \sigma x_n$ for $\sigma \in G$, $w = x_1 \cdots x_n \in X^\diamond$. Given a term ordering $\leq$ of $X^\diamond$, we define a new relation on $X^\diamond$ as follows:

Definition 2.8. (The symmetric cancellation ordering corresponding to G and $\leq$.)

$$v \preceq w \quad :\Longleftrightarrow \quad \begin{cases} v \leq w \text{ and there exist } \sigma \in G \text{ and a monomial} \\ u \in X^\diamond \text{ such that } w = u\sigma v \text{ and for all } v' \leq v, \\ \text{we have } u\sigma v' \leq w. \end{cases}$$

Remark 2.9. Every term ordering $\leq$ is *linear*: $v \leq w \Longleftrightarrow uv \leq uw$ for all monomials u, v, w. Hence the condition above may be rewritten as: $v \leq w$ and there exists $\sigma \in G$ such that $\sigma v | w$ and $\sigma v' \leq \sigma v$ for all $v' \leq v$. (We say that "σ witnesses $v \preceq w$.")

Example 2.10. Let $X = \{x_1, x_2, \ldots\}$ be a countably infinite set of indeterminates, ordered such that $x_1 < x_2 < \cdots$, and let $\leq\, = \,\leq_{\mathrm{lex}}$ be the corresponding lexicographic ordering of $X^\diamond$. Let also G be the group of permutations of $\{1, 2, 3, \ldots\}$, acting on X via $\sigma x_i = x_{\sigma(i)}$. As an example of the relation $\preceq$, consider the following chain:

$$x_1^2 \preceq x_1 x_2^2 \preceq x_1^3 x_2 x_3^2.$$

To verify the first inequality, notice that $x_1 x_2^2 = x_1 \sigma(x_1^2)$, in which σ is the transposition $(1\,2)$. If $v' = x_1^{a_1} \cdots x_n^{a_n} \leq x_1^2$ with $a_1, \ldots, a_n \in \mathbb{N}$, $a_n > 0$, then it follows that $n = 1$ and $a_1 \leq 2$. In particular, $x_1 \sigma v' = x_1 x_2^{a_1} \leq x_1 x_2^2$. For the second relationship, we have that $x_1^3 x_2 x_3^2 = x_1^3 \tau(x_1 x_2^2)$, in which τ is the cycle $(1\,2\,3)$. Additionally, if $v' = x_1^{a_1} \cdots x_n^{a_n} \leq x_1 x_2^2$ with $a_1, \ldots, a_n \in \mathbb{N}$, $a_n > 0$, then $n \leq 2$, and if $n = 2$, then either $a_2 = 1$ or $a_2 = 2$, $a_1 \leq 1$. In each case we get $x_1^3 \tau v' = x_1^3 x_2^{a_1} x_3^{a_2} \leq x_1^3 x_2 x_3^2$.

Although Definition 2.8 appears technical, we will soon present a nice interpretation of it that involves leading term cancellation of polynomials. First we verify that it is indeed an ordering.

LEMMA 2.11. *The relation $\preceq$ is an ordering on monomials.*

PROOF. First notice that $w \preceq w$ since we may take $u = 1$ and $\sigma =$ the identity permutation. Next, suppose that $u \preceq v \preceq w$. Then there exist permutations σ, τ in G and monomials u_1, u_2 in $X^\diamond$ such that $v = u_1\sigma u$, $w = u_2\tau v$. In particular, $w = u_2(\tau u_1)(\tau\sigma u)$. Additionally, if $v' \leq u$, then $u_1\sigma v' \leq v$, so that $u_2\tau(u_1\sigma v') \leq w$. It follows that $u_2(\tau u_1)(\tau\sigma v') \leq w$. This shows transitivity; anti-symmetry of $\preceq$ follows from anti-symmetry of $\leq$. $\square$

We offer a useful interpretation of this ordering (which motivates its name). We fix a commutative ring A and let $R = A[X]$ be the ring of polynomials with coefficients from A in the collection of commuting indeterminates X. Its elements may be written uniquely in the form

$$f = \sum_{w \in X^\diamond} a_w w$$

where $a_w \in A$ for all $w \in X^\diamond$, and all but finitely many a_w are zero. We say that a monomial w *occurs* in f if $a_w \neq 0$. Given a non-zero $f \in R$ we define $\operatorname{lm}(f)$, the *leading monomial* of f (with respect to our choice of term ordering $\leq$) to be the largest monomial w (with respect to $\leq$) which occurs in f. If $w = \operatorname{lm}(f)$, then a_w is the *leading coefficient* of f, denoted by $\operatorname{lc}(f)$, and $a_w w$ is the *leading term* of f, denoted by $\operatorname{lt}(f)$. By convention, we set $\operatorname{lm}(0) = \operatorname{lc}(0) = \operatorname{lt}(0) = 0$. We let $R[G]$ be the group ring of G over R (with multiplication given by $f\sigma \cdot g\tau = fg(\sigma\tau)$ for $f, g \in R$, $\sigma, \tau \in G$), and we view R as a left $R[G]$-module in the natural way.

LEMMA 2.12. *Let $f \in R$, $f \neq 0$, and $u, w \in X^\diamond$. Suppose that $\sigma \in G$ witnesses $\operatorname{lm}(f) \preceq w$, and let $u \in X^\diamond$ with $u\sigma \operatorname{lm}(f) = w$. Then $\operatorname{lm}(u\sigma f) = u\sigma \operatorname{lm}(f)$.*

PROOF. Put $v = \operatorname{lm}(f)$. Every monomial occurring in $u\sigma f$ has the form $u\sigma v'$, where v' occurs in f. Hence $v' \leq v$, and since σ witnesses $v \preceq w$, this yields $u\sigma v' \leq w$. $\square$

Suppose that A is a field, let $v \preceq w$ be in $X^\diamond$ and let f, g be two polynomials in R with leading monomials v, w, respectively. Then, from the definition and the lemma above, there exists a $\sigma \in G$ and a term cu ($c \in A \setminus \{0\}$, $u \in X^\diamond$) such that all monomials occurring in

$$h = g - cu\sigma f$$

are strictly smaller (with respect to $\leq$) than w. For readers familiar with the theory of Gröbner bases, the polynomial h can be viewed as a kind of symmetric version of the S-polynomial (see, for instance, [**5**, Chapter 15]).

EXAMPLE 2.13. In the situation of Example 2.10 above, let $f = x_1x_2^2 + x_2 + x_1^2$ and $g = x_1^3x_2x_3^2 + x_3^2 + x_1^4x_3$. Set $\sigma = (1\,2\,3)$, and observe that

$$g - x_1^3\sigma f = x_1^4x_3 + x_3^2 - x_1^3x_3 - x_1^3x_2^2$$

has a smaller leading monomial than g.

We are mostly interested in the case where our term ordering on $X^\diamond$ is $\leq_{\text{lex}}$, and $G = \mathfrak{S}_X$. Under these assumptions we have:

LEMMA 2.14. *Let $v, w \in X^\diamond$ with $v \preceq w$. Then for every $\sigma \in \mathfrak{S}_X$ witnessing $v \preceq w$ we have $\sigma(X^{\leq|v|}) \subseteq X^{\leq|w|}$. Moreover, if the order type of $(X, \leq)$ is $\leq \omega$, then we can choose such σ with the additional property that $\sigma(x) = x$ for all $x > |w|$.*

PROOF. To see the first claim, suppose for a contradiction that $\sigma x > |w|$ for some $x \in X$, $x \leq |v|$. We have $\sigma v | w$, so if $x|v$, then $\sigma x | w$, contradicting $\sigma x > |w|$. In particular $x < |v|$, which yields $x <_{\text{lex}} v$ and thus $\sigma x \leq_{\text{lex}} \sigma v \leq_{\text{lex}} w$, again contradicting $\sigma x > |w|$. Now suppose that the order type of X is $\leq \omega$, and let σ witness $v \preceq w$. Then $|v| \leq |w|$, and $\sigma X^{\leq|v|}$ can be extended to a permutation σ' of the finite set $X^{\leq|w|}$. We further extend σ' to a permutation of X by setting $\sigma'(x) = x$ for all $x > |w|$. One checks easily that σ' still witnesses $v \preceq w$. $\square$

2.6. Lovely orderings. We say that a term ordering $\leq$ of $X^\diamond$ is *lovely* for G if the corresponding symmetric cancellation ordering $\preceq$ on $X^\diamond$ is a well-quasi-ordering. If $\leq$ is lovely for a subgroup of G, then $\leq$ is lovely for G.

Example 2.15. The symmetric cancellation ordering corresponding to $G = \{1\}$ and a given term ordering $\leq$ of $X^\diamond$ is just

$$v \preceq w \quad\Longleftrightarrow\quad v \leq w \ \wedge\ v|w.$$

Hence a term ordering of $X^\diamond$ is lovely for $G = \{1\}$ if and only if divisibility in $X^\diamond$ has no infinite antichains; that is, exactly if X is finite.

This terminology is inspired by the following definition from [**3**] (which in turn goes back to an idea in [**2**]):

Definition 2.16. Given an ordering $\leq$ of X, consider the following ordering of X:

$$x \sqsubseteq y \quad :\Longleftrightarrow\quad \begin{cases} x \leq y \text{ and there exists } \sigma \in G \text{ such that } \sigma x = y \\ \text{and for all } x' \leq x, \text{ we have } \sigma x' \leq y. \end{cases}$$

A well-ordering $\leq$ of X is called *nice* (for G) if $\sqsubseteq$ is a well-quasi-ordering.

In [**2**] one finds various examples of nice orderings, and in [**3**] it is shown that if X admits a nice ordering with respect to G, then for every field F, the free F-module FX with basis X is Noetherian as a module over $F[G]$. It is clear that the restriction to X of a lovely ordering of $X^\diamond$ is nice. However, there do exist permutation groups (G, X) for which X admits a nice ordering, but $X^\diamond$ does not admit a lovely ordering; see Example 3.4 and Proposition 5.2 below.

Example 2.17. Suppose that X is countable. Then every well-ordering of X of order type ω is nice for $\mathfrak{S}_X$. To see this, we may assume that $X = \mathbb{N}$ with its usual ordering. It is then easy to see that if $x \leq y$ in $\mathbb{N}$, then $x \sqsubseteq y$, witnessed by any extension σ of the strictly increasing map $n \mapsto n + y - x \colon \mathbb{N}^{\leq x} \to \mathbb{N}$ to a permutation of $\mathbb{N}$.

The following crucial fact (generalizing the last example) is needed for our proof of Theorem 1.1:

Theorem 2.18. *The lexicographic ordering of $X^\diamond$ corresponding to a cardinal well-ordering of a set X is lovely for the full symmetric group $\mathfrak{S}_X$ of X.*

For the proof, let as above $\operatorname{Fin}(X, \mathbb{N})$ be the set of all sequences in $\mathbb{N}$ indexed by elements in some proper initial segment of X which have finite range, quasi-ordered by $\leq_{\mathrm{H}}$. For a monomial $w \neq 1$ we define $w^* \colon X^{\leq |w|} \to \mathbb{N}$ by

$$w^*(x) := \max\{a \in \mathbb{N} : x^a | w\}.$$

Then clearly $w^* \in \operatorname{Fin}(X, \mathbb{N})$, in fact, $w^*(x) = 0$ for all but finitely many $x \in X^{\leq |w|}$. We also let $1^* :=$ the empty sequence $\emptyset \to \mathbb{N}$ (the unique smallest element of $\operatorname{Fin}(X, \mathbb{N})$). We now quasi-order $X^\diamond \times \operatorname{Fin}(X, \mathbb{N})$ by the cartesian product of the ordering $\leq_{\mathrm{lex}}$ on $X^\diamond$ and the quasi-ordering $\leq_{\mathrm{H}}$ on $\operatorname{Fin}(X, \mathbb{N})$. By Corollary 2.3, Theorem 2.4, and the remark following Proposition 2.1, $X^\diamond \times \operatorname{Fin}(X, \mathbb{N})$ is well-quasi-ordered. Therefore, in order to finish the proof of Theorem 2.18, it suffices to show:

Lemma 2.19. *The map*

$$w \mapsto (w, w^*) \colon X^\diamond \to X^\diamond \times \operatorname{Fin}(X, \mathbb{N})$$

is a quasi-embedding with respect to the symmetric cancellation ordering on $X^\diamond$ and the quasi-ordering on $X^\diamond \times \operatorname{Fin}(X, \mathbb{N})$.

Proof. Suppose that v, w are monomials with $v \leq_{\mathrm{lex}} w$ and $v^* \leq_{\mathrm{H}} w^*$; we need to show that $v \preceq w$. For this we may assume that $v, w \neq 1$. So there exists a strictly increasing function $\varphi \colon X^{\leq |v|} \to X^{\leq |w|}$ such that

$$v^*(x) \leq w^*(\varphi(x)) \qquad \text{for all } x \in X \text{ with } x \leq |v|. \tag{2.1}$$

By Lemma 2.7 there exists $\sigma \in \mathfrak{S}_X$ such that $\sigma X^{\leq |v|} = \varphi X^{\leq |v|}$. Then clearly $\sigma v | w$ by (2.1). Now let $v' \leq_{\mathrm{lex}} v$; we claim that $\sigma v' \leq_{\mathrm{lex}} \sigma v$. Again we may assume $v' \neq 1$. Then $|v'| \leq |v|$, hence we may write

$$v' = x_1^{a_1} \cdots x_n^{a_n}, \quad v = x_1^{b_1} \cdots x_n^{b_n}$$

with $x_1 < \cdots < x_n \leq |v|$ in X and $a_i, b_j \in \mathbb{N}$. Put $y_1 := \varphi(x_1), \ldots, y_n := \varphi(x_n)$. Then $y_1 < \cdots < y_n$ and

$$\sigma v' = y_1^{a_1} \cdots y_n^{a_n}, \quad \sigma v = y_1^{b_1} \cdots y_n^{b_n},$$

and therefore $\sigma v' \leq_{\text{lex}} \sigma v$ as required. $\square$

2.7. The case of countable X. In Section 4 we will apply Theorem 2.18 in the case where X is countable. Then the order type of X is at most ω, and in the proof of the theorem given above we only need to appeal to a special instance (Higman's Lemma) of Theorem 2.4. We finish this section by giving a self-contained proof of this important special case of Theorem 2.18, avoiding Theorem 2.4. Let $\mathfrak{S}_{(X)}$ denote the subgroup of $\mathfrak{S}_X$ consisting of all $\sigma \in \mathfrak{S}_X$ with the property that $\sigma(x) = x$ for all but finitely many letters $x \in X$.

Theorem 2.20. *The lexicographic ordering of $X^\diamond$ corresponding to a cardinal well-ordering of a countable set X is lovely for $\mathfrak{S}_{(X)}$.*

Let X be countable and let $\leq$ be a cardinal well-ordering of X. Enumerate the elements of X as $x_1 < x_2 < \cdots$. We assume that X is infinite; this is not a restriction, since by Lemma 2.14 we have:

Lemma 2.21. *If the lexicographic ordering of $X^\diamond$ is lovely for $\mathfrak{S}_{(X)}$, then for any n and $X_n := \{x_1, \ldots, x_n\}$, the lexicographic ordering of $(X_n)^\diamond$ is lovely for $\mathfrak{S}_{X_n}$.* $\square$

We begin with some preliminary lemmas. Here, $\preceq$ is the symmetric cancellation ordering corresponding to $\mathfrak{S}_{(X)}$ and $\leq_{\text{lex}}$. We identfy $\mathfrak{S}_{(X)}$ and $\mathfrak{S}_\infty := \mathfrak{S}_{(\mathbb{N})}$ in the natural way, and for every n we regard $\mathfrak{S}_n$, the group of permutations of $\{1, 2, \ldots, n\}$, as a subgroup of $\mathfrak{S}_\infty$; then $\mathfrak{S}_n \leq \mathfrak{S}_{n+1}$ for each n, and $\mathfrak{S}_\infty = \bigcup_n \mathfrak{S}_n$.

Lemma 2.22. *Suppose that $x_1^{a_1} \cdots x_n^{a_n} \preceq x_1^{b_1} \cdots x_n^{b_n}$ where $a_i, b_j \in \mathbb{N}$, $b_n > 0$. Then for any $c \in \mathbb{N}$ we have $x_1^{a_1} \cdots x_n^{a_n} \preceq x_1^c x_2^{b_1} \cdots x_{n+1}^{b_n}$.*

Proof. Let $v := x_1^{a_1} \cdots x_n^{a_n}$, $w := x_1^{b_1} \cdots x_n^{b_n}$. We may assume $v \neq 1$. Clearly $v \leq_{\text{lex}} w$ and $b_n > 0$ yield $x_1^{a_1} \cdots x_n^{a_n} \leq_{\text{lex}} x_1^c x_2^{b_1} \cdots x_{n+1}^{b_n}$. Let now $\sigma \in \mathfrak{S}_\infty$ witness $v \preceq w$. Let τ be the cyclic permutation $\tau = (1\,2\,3 \cdots (n+1))$ and set $\widehat{\sigma} := \tau\sigma$. Then $\sigma v | w$ yields $\widehat{\sigma} v | \tau w$, hence $\widehat{\sigma} v | x_1^c \tau w = x_1^c x_2^{b_1} \cdots x_{n+1}^{b_n}$. Next, suppose that $v' \leq_{\text{lex}} v$; then $\sigma v' \leq_{\text{lex}} \sigma v$. By Lemma 2.14 and the nature of τ, the map $\tau\sigma(\{1, \ldots, |v|\})$ is strictly increasing, which gives $\widehat{\sigma} v' = \tau\sigma v' \leq_{\text{lex}} \tau\sigma v = \widehat{\sigma} v$. Hence $\widehat{\sigma}$ witnesses $x_1^{a_1} \cdots x_n^{a_n} \preceq x_1^c x_2^{b_1} \cdots x_{n+1}^{b_n}$. $\square$

Lemma 2.23. *If $x_1^{a_1} \cdots x_n^{a_n} \preceq x_1^{b_1} \cdots x_n^{b_n}$, where $a_i, b_j \in \mathbb{N}$, $b_n > 0$, and $a, b \in \mathbb{N}$ are such that $a \leq b$, then $x_1^a x_2^{a_1} \cdots x_{n+1}^{a_n} \preceq x_1^b x_2^{b_1} \cdots x_{n+1}^{b_{n+1}}$.*

Proof. As before let $v := x_1^{a_1} \cdots x_n^{a_n}$, $w := x_1^{b_1} \cdots x_n^{b_n}$. Once again, we may assume $v \neq 1$, and it is clear that $x_1^a x_2^{a_1} \cdots x_{n+1}^{a_n} \leq_{\text{lex}} x_1^b x_2^{b_1} \cdots x_{n+1}^{b_{n+1}}$. Let $\sigma \in \mathfrak{S}_\infty$ witness $v \preceq w$. By Lemma 2.14 we may assume that $\sigma(x_i) = x_i$ for all $i > n$. Let τ be the cyclic permutation $\tau = (1\,2 \cdots (n+1))$. Setting $\widehat{\sigma} = \tau\sigma\tau^{-1}$, we have $\widehat{\sigma} x_1 = x_1$, hence

$$\widehat{\sigma}(x_1^a x_2^{a_1} \cdots x_{n+1}^{a_n}) = \widehat{\sigma}(x_1^a)\widehat{\sigma}(x_2^{a_1} \cdots x_{n+1}^{a_n}) = x_1^a \tau\sigma v. \tag{2.2}$$

Since $\sigma v | w$, this last expression divides $x_1^b \tau w = x_1^b x_2^{b_1} \cdots x_{n+1}^{b_n}$. Suppose that $v' = x_1^{c_1} \cdots x_{n+1}^{c_{n+1}} \leq_{\text{lex}} x_1^a x_2^{a_1} \cdots x_{n+1}^{a_n}$, where $c_i \in \mathbb{N}$. Then, since we are using a lexicographic order, we have

$$x_2^{c_2} \cdots x_{n+1}^{c_{n+1}} \leq_{\text{lex}} x_2^{a_1} \cdots x_{n+1}^{a_n}$$

and therefore

$$\tau^{-1}(x_2^{c_2} \cdots x_{n+1}^{c_{n+1}}) = x_1^{c_2} \cdots x_n^{c_{n+1}} \leq_{\text{lex}} \tau^{-1}(x_2^{a_1} \cdots x_{n+1}^{a_n}) = v.$$

By assumption, this implies that $\sigma\tau^{-1}(x_2^{c_2} \cdots x_{n+1}^{c_{n+1}}) \leq_{\text{lex}} \sigma v$ and thus by (2.2)

$$\widehat{\sigma}(x_2^{c_2} \cdots x_{n+1}^{c_{n+1}}) \leq_{\text{lex}} \tau\sigma v = \widehat{\sigma}(x_2^{a_1} \cdots x_{n+1}^{a_n}).$$

If this inequality is strict, then since $1 \notin \widehat{\sigma}\big(\{2, \ldots, n+1\}\big)$, clearly

$$\widehat{\sigma} v' = x_1^{c_1} \widehat{\sigma}(x_2^{c_2} \cdots x_{n+1}^{c_{n+1}}) <_{\text{lex}} x_1^a \tau\sigma v = \widehat{\sigma}(x_1^a x_2^{a_1} \cdots x_{n+1}^{a_n}).$$

Otherwise $x_2^{c_2} \cdots x_{n+1}^{c_{n+1}} = x_2^{a_1} \cdots x_{n+1}^{a_n}$, hence $c_1 \leq a$, in which case we still have $\widehat{\sigma} v' \leq_{\text{lex}} \widehat{\sigma}(x_1^a x_2^{a_1} \cdots x_{n+1}^{a_n})$. Therefore $\widehat{\sigma}$ witnesses $x_1^a x_2^{a_1} \cdots x_{n+1}^{a_n} \preceq x_1^b x_2^{b_1} \cdots x_{n+1}^{b_{n+1}}$. This completes the proof. $\square$

We now have enough to show Theorem 2.20. The proof uses the basic idea from Nash-Williams' proof [**13**] of Higman's lemma. Assume for the sake of contradiction that there exists a bad sequence

$$w^{(1)}, w^{(2)}, \dots, w^{(n)}, \dots \qquad \text{in } X^{\diamond}.$$

For $w \in X^{\diamond} \setminus \{1\}$ let $j(w)$ be the index $j \geq 1$ with $|w| = x_j$, and put $j(1) := 0$. We may assume that the bad sequence is chosen in such a way that for every n, $j(w^{(n)})$ is *minimal* among the $j(w)$, where w ranges over all elements of $X^{\diamond}$ with the property that $w^{(1)}, w^{(2)}, \dots, w^{(n-1)}, w$ can be continued to a bad sequence in $X^{\diamond}$. Because $1 \leq_{\text{lex}} w$ for all $w \in X^{\diamond}$, we have $j(w^{(n)}) > 0$ for all n. For every $n > 0$, write $w^{(n)} = x_1^{a^{(n)}} v^{(n)}$ with $a^{(n)} \in \mathbb{N}$ and $v^{(n)} \in X^{\diamond}$ not divisible by x_1. Since $\mathbb{N}$ is well-ordered, there is an infinite sequence $1 \leq i_1 < i_2 < \cdots$ of indices such that $a^{(i_1)} \leq a^{(i_2)} \leq \cdots$. Consider the monoid homomorphism $\alpha \colon X^{\diamond} \to X^{\diamond}$ given by $\alpha(x_{i+1}) = x_i$ for all $i > 1$. Then $j(\alpha(w)) = j(w) - 1$ if $w \neq 1$. Hence by minimality of $w^{(1)}, w^{(2)}, \dots$, the sequence

$$w^{(1)}, w^{(2)}, \dots, w^{(i_1 - 1)}, \alpha(v^{(i_1)}), \alpha(v^{(i_2)}), \dots, \alpha(v^{(i_n)}), \dots$$

is good; that is, there exist $j < i_1$ and k with $w^{(j)} \preceq \alpha(v^{(i_k)})$, or there exist $k < l$ with $\alpha(v^{(i_k)}) \preceq \alpha(v^{(i_l)})$. In the first case we have $w^{(j)} \preceq w^{(i_k)}$ by Lemma 2.22; and in the second case, $w^{(i_k)} \preceq w^{(i_l)}$ by Lemma 2.23. This contradicts the badness of our sequence $w^{(1)}, w^{(2)}, \dots$, finishing the proof.

Question. Careful inspection of the proof of Theorem 2.18 (in particular Lemma 2.7) shows that in the statement of the theorem, we can replace $\mathfrak{S}_X$ by its subgroup consisting of all σ with the property that the set of $x \in X$ with $\sigma(x) \neq x$ has cardinality $< |X|$. In Theorem 2.18, can one always replace $\mathfrak{S}_X$ by $\mathfrak{S}_{(X)}$?

3. Proof of the Finiteness Theorem

We now come to the proof our main result. Throughout this section we let A be a commutative Noetherian ring, X an arbitrary set, $R = A[X]$, and we let G be a permutation group on X. An $R[G]$-submodule of R will be called a *G-invariant ideal* of R, or simply an *invariant ideal*, if G is understood. We will show:

Theorem 3.1. *If $X^{\diamond}$ admits a lovely term ordering for G, then R is Noetherian as an $R[G]$-module.*

For $G = \{1\}$ and X finite, this theorem reduces to Hilbert's basis theorem, by Example 2.15. We also obtain Theorem 1.1:

Corollary 3.2. *The $R[\mathfrak{S}_X]$-module R is Noetherian.*

Proof. Choose a cardinal well-ordering of X. Then the corresponding lexicographic ordering of $X^{\diamond}$ is lovely for $\mathfrak{S}_X$, by Theorem 2.18. Apply Theorem 3.1. □

Remark 3.3. It is possible to replace the use of Theorem 2.18 in the proof of the corollary above by the more elementary Theorem 2.20. This is because if the $R[\mathfrak{S}_X]$-module R was not Noetherian, then one could find a *countably generated* $R[\mathfrak{S}_X]$-submodule of R which is not finitely generated, and hence a countable subset X' of X such that $R' = A[X']$ is not a Noetherian $R'[\mathfrak{S}_{X'}]$-module.

The following example shows how the conclusion of Theorem 3.1 may fail:

Example 3.4. Suppose that G has a cyclic subgroup H which acts freely and transitively on X. Then X has a nice ordering (see [**2**]), but $R = \mathbb{Q}[X^{\diamond}]$ is not Noetherian. To see this let σ be a generator for H, and let $x \in X$ be arbitrary. Then the $R[G]$-submodule of $R = \mathbb{Q}[X^{\diamond}]$ generated by the elements $\sigma^n x \sigma^{-n} x$ ($n \in \mathbb{N}$) is not finitely generated. So by Theorem 3.1, $X^{\diamond}$ does not admit a lovely term ordering for G.

For the proof of Theorem 3.1 we develop a bit of Gröbner basis theory for the $R[G]$-module R. For the time being, we fix an arbitrary term ordering $\leq$ (not necessarily lovely for G) of $X^{\diamond}$.

3.1. Reduction of polynomials. Let $f \in R$, $f \neq 0$, and let B be a set of non-zero polynomials in R. We say that f is *reducible by B* if there exist pairwise distinct $g_1, \dots, g_m \in B$, $m \geq 1$, such that for each i we have $\operatorname{lm}(g_i) \preceq \operatorname{lm}(f)$, witnessed by some $\sigma_i \in G$, and

$$\operatorname{lt}(f) = a_1 w_1 \sigma_1 \operatorname{lt}(g_1) + \cdots + a_m w_m \sigma_m \operatorname{lt}(g_m)$$

for non-zero $a_i \in A$ and monomials $w_i \in X^\diamond$ such that $w_i\sigma_i \operatorname{lm}(g_i) = \operatorname{lm}(f)$. In this case we write $f \xrightarrow[B]{} h$, where

$$h = f - \big(a_1 w_1 \sigma_1 g_1 + \cdots + a_m w_m \sigma_m g_m\big),$$

and we say that f *reduces to* h by B. We say that f is *reduced* with respect to B if f is not reducible by B. By convention, the zero polynomial is reduced with respect to B. Trivially, every element of B reduces to 0.

EXAMPLE 3.5. Suppose that A is a field. Then f is reducible by B if and only if there exists some $g \in B$ such that $\operatorname{lm}(g) \preceq \operatorname{lm}(f)$.

EXAMPLE 3.6. Suppose that f is reducible by B as defined (for finite X) in, say, [**1**, Chapter 4], that is: there exist $g_1, \dots, g_m \in B$ and $a_1, \dots, a_m \in A$ ($m \geq 1$) such that $\operatorname{lm}(g_i) | \operatorname{lm}(f)$ for all i and

$$\operatorname{lc}(f) = a_1 \operatorname{lc}(g_1) + \cdots + a_m \operatorname{lc}(g_m).$$

Then f is reducible by B in the sense defined above. (Taking $\sigma_i = 1$ for all i.)

REMARK 3.7. Suppose that $G = \mathfrak{S}_X$, the term ordering $\leq$ of $X^\diamond$ is $\leq_{\text{lex}}$, and the order type of $(X, \leq)$ is $\leq \omega$. Then in the definition of reducibility by B above, we may require that the σ_i satisfy $\sigma_i(x) = x$ for all $1 \leq i \leq m$ and $x > |\operatorname{lm}(f)|$. (By Lemma 2.14.)

The smallest quasi-ordering on R extending the relation $\xrightarrow[B]{}$ is denoted by $\xrightarrow[B]{*}$. If $f, h \neq 0$ and $f \xrightarrow[B]{} h$, then $\operatorname{lm}(h) < \operatorname{lm}(f)$, by Lemma 2.12. In particular, every chain

$$h_0 \xrightarrow[B]{} h_1 \xrightarrow[B]{} h_2 \xrightarrow[B]{} \cdots$$

with all $h_i \in R \setminus \{0\}$ is finite. (Since the term ordering $\leq$ is well-founded.) Hence there exists $r \in R$ such that $f \xrightarrow[B]{*} r$ and r is reduced with respect to B; we call such an r a *normal form* of f with respect to B.

LEMMA 3.8. *Suppose that $f \xrightarrow[B]{*} r$. Then there exist $g_1, \dots, g_n \in B$, $\sigma_1, \dots, \sigma_n \in G$ and $h_1, \dots, h_n \in R$ such that*

$$f = r + \sum_{i=1}^{n} h_i \sigma_i g_i \quad \textit{and} \quad \operatorname{lm}(f) \geq \max_{1 \leq i \leq n} \operatorname{lm}(h_i \sigma_i g_i).$$

(In particular, $f - r \in \langle B \rangle_{R[G]}$.)

PROOF. This is clear if $f = r$. Otherwise we have $f \xrightarrow[B]{} h \xrightarrow[B]{*} r$ for some $h \in R$. Inductively we may assume that there exist $g_1, \dots, g_n \in B$, $\sigma_1, \dots, \sigma_n \in G$ and $h_1, \dots, h_n \in R$ such that

$$h = r + \sum_{i=1}^{n} h_i \sigma_i g_i \quad \text{and} \quad \operatorname{lm}(h) \geq \max_{1 \leq i \leq n} \operatorname{lm}(h_i \sigma_i g_i).$$

There are also $g_{n+1}, \dots, g_{n+m} \in B$, $\sigma_{n+1}, \dots, \sigma_{n+m} \in G$, $a_{n+1}, \dots, a_{n+m} \in A$ and $w_{n+1}, \dots, w_{n+m} \in X^\diamond$ such that $\operatorname{lm}(w_{n+i}\sigma_{n+i} g_{n+i}) = \operatorname{lm}(f)$ for all i and

$$\operatorname{lt}(f) = \sum_{i=1}^{m} a_{n+i} w_{n+i} \sigma_{n+i} \operatorname{lt}(g_{n+i}), \qquad f = h + \sum_{i=1}^{m} a_{n+i} w_{n+i} \sigma_{n+i} g_{n+i}.$$

Hence putting $h_{n+i} := a_{n+i} w_{n+i}$ for $i = 1, \dots, m$ we have $f = r + \sum_{j=1}^{n+m} h_j \sigma_j g_j$ and $\operatorname{lm}(f) > \operatorname{lm}(h) \geq \operatorname{lm}(h_j \sigma_j g_j)$ if $1 \leq j \leq n$, $\operatorname{lm}(f) = \operatorname{lm}(h_j \sigma_j g_j)$ if $n < j \leq n + m$. $\square$

REMARK 3.9. Suppose that $G = \mathfrak{S}_X$, $\leq \, = \, \leq_{\text{lex}}$, and X has order type $\leq \omega$. Then in the previous lemma we can choose the σ_i such that in addition $\sigma_i(x) = x$ for all i and all $x > |\operatorname{lm}(f)|$. (By Remark 3.7.)

3.2. Gröbner bases. Let B be a subset of R. We let

$$\mathrm{lt}(B) := \big\langle \mathrm{lc}(g)w : 0 \neq g \in B,\ \mathrm{lm}(g) \preceq w \big\rangle_A$$

be the A-submodule of R generated by all elements of the form $\mathrm{lc}(g)w$, where $g \in B$ is non-zero and w is a monomial with $\mathrm{lm}(g) \preceq w$. Clearly for non-zero $f \in R$ we have: $\mathrm{lt}(f) \in \mathrm{lt}(B)$ if and only if f is reducible by B. In particular, $\mathrm{lt}(B)$ contains $\big\{ \mathrm{lt}(g) : g \in B \big\}$, and for an ideal I of R which is G-invariant, we simply have

$$\mathrm{lt}(I) = \big\{ \mathrm{lt}(f) : f \in I \big\}.$$

(Use Lemma 2.12.) We say that a subset B of an invariant ideal I of R is a *Gröbner basis* for I (with respect to our choice of term ordering $\leq$) if $\mathrm{lt}(I) = \mathrm{lt}(B)$.

Lemma 3.10. *Let I be an invariant ideal of R and B be a set of non-zero elements of I. The following are equivalent:*

(1) *B is a Gröbner basis for I.*
(2) *Every non-zero $f \in I$ is reducible by B.*
(3) *Every $f \in I$ has normal form 0. (In particular, $I = \langle B \rangle_{R[G]}$.)*
(4) *Every $f \in I$ has unique normal form 0.*

Proof. The implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) are either obvious or follow from the remarks preceding the lemma. Suppose that (4) holds. Every $f \in I \setminus \{0\}$ with $\mathrm{lt}(f) \notin \mathrm{lt}(B)$ is reduced with respect to B, hence has two distinct normal forms (0 and f), a contradiction. Thus $\mathrm{lt}(I) = \mathrm{lt}(B)$. $\square$

Suppose that B is a Gröbner basis for an ideal I of the polynomial ring $R = A[X^\diamond]$, in the usual sense of the word (as defined, for finite X, in [**1**, Chapter 4]); if I is invariant, then B is a Gröbner basis for I as defined above (by Example 3.6). Moreover, for $G = \{1\}$, the previous lemma reduces to a familiar characterization of Gröbner bases in the usual case of polynomial rings. It is probably possible to also introduce a notion of S-polynomial and to prove a Buchberger-style criterion for Gröbner bases in our setting, leading to a completion procedure for the construction of Gröbner bases. At this point, we will not pursue these issues further, and rather show:

Proposition 3.11. *Suppose that the term ordering $\leq$ of $X^\diamond$ is lovely for G. Then every invariant ideal of R has a finite Gröbner basis.*

For a subset B of R let $\mathrm{lm}(B)$ denote the final segment of $X^\diamond$ with respect to $\preceq$ generated by the $\mathrm{lm}(g)$, $g \in B$. If A is a field, then a subset B of an invariant ideal I of R is a Gröbner basis for I if and only if $\mathrm{lm}(B) = \mathrm{lm}(I)$. Hence in this case, the proposition follows immediately from the equivalence of (1) and (4) in Proposition 2.1. For the general case we use the following observation:

Lemma 3.12. *Let S be a well-quasi-ordered set and T be a well-founded ordered set, and let $\varphi\colon S \to T$ be decreasing: $s \leq t \Rightarrow \varphi(s) \geq \varphi(t)$, for all $s, t \in S$. Then the quasi-ordering $\leq_\varphi$ on S defined by*

$$s \leq_\varphi t \quad :\Longleftrightarrow \quad s \leq t \ \wedge\ \varphi(s) = \varphi(t)$$

is a well-quasi-ordering. $\square$

Proof of Proposition 3.11. Suppose now that our term ordering of $X^\diamond$ is lovely for G, and let I be an invariant ideal of R. For $w \in X^\diamond$ consider

$$\mathrm{lc}(I, w) := \big\{ \mathrm{lc}(f) : f \in I, \text{ and } f = 0 \text{ or } \mathrm{lm}(f) = w \big\},$$

an ideal of A. Note that if $v \preceq w$, then $\mathrm{lc}(I, v) \subseteq \mathrm{lc}(I, w)$. We apply the lemma to $S = X^\diamond$, quasi-ordered by $\preceq$, $T =$ the collection of all ideals of A, ordered by reverse inclusion, and φ given by $w \mapsto \mathrm{lc}(I, w)$. Thus by (4) in Proposition 2.1, applied to the final segment $X^\diamond$ of the well-quasi-ordering $\leq_\varphi$, we obtain finitely many $w_1, \dots, w_m \in X^\diamond$ with the following property: for every $w \in X^\diamond$ there exists some $i \in \{1, \dots, m\}$ such that $w_i \preceq w$ and $\mathrm{lc}(I, w_i) = \mathrm{lc}(I, w)$. Using Noetherianity of A, for every i we now choose finitely many non-zero elements $g_{i1}, \dots, g_{in_i}$ of I ($n_i \in \mathbb{N}$), each with leading monomial w_i, whose leading coefficients generate the ideal $\mathrm{lc}(I, w_i)$ of A. We claim that

$$B := \{g_{ij} : 1 \leq i \leq m,\ 1 \leq j \leq n_i\}$$

is a Gröbner basis for I. To see this, let $0 \neq f \in I$, and put $w := \operatorname{lm}(f)$. Then there is some i with $w_i \preceq w$ and $\operatorname{lc}(I, w_i) = \operatorname{lc}(I, w)$. This shows that f is reducible by $\{g_{i1}, \dots, g_{i,n_i}\}$, and hence by B. By Lemma 3.10, B is a Gröbner basis for I. □

From Proposition 3.11 and the implication (1) ⇒ (3) in Lemma 3.10 we obtain Theorem 3.1.

3.3. A partial converse of Theorem 3.1. Consider now the quasi-ordering $|_G$ of $X^\diamond$ defined by

$$v|_G w \quad :\Longleftrightarrow \quad \exists \sigma \in G : \sigma v | w,$$

which extends every symmetric cancellation ordering corresponding to a term ordering of $X^\diamond$. If M is a set of monomials from $X^\diamond$ and F the final segment of $(X^\diamond, |_G)$ generated by M, then the invariant ideal $\langle M \rangle_{R[G]}$ of R is finitely generated as an $R[G]$-module if and only if F is generated by a finite subset of M. Hence by the implication (4) ⇒ (1) in Proposition 2.1 we get:

Lemma 3.13. *If R is Noetherian as an $R[G]$-module, then $|_G$ is a well-quasi-ordering.* □

This will be used in Section 5 below.

3.4. Connection to a concept due to Michler. Let $\leq$ be a term ordering of $X^\diamond$. For each $\sigma \in G$ we define a term ordering $\leq_\sigma$ on $X^\diamond$ by

$$v \leq_\sigma w \quad \Longleftrightarrow \quad \sigma v \leq \sigma w.$$

We denote the leading monomial of $f \in R$ with respect to $\leq_\sigma$ by $\operatorname{lm}_\sigma(f)$. Clearly we have

$$\sigma \operatorname{lm}(f) = \operatorname{lm}_{\sigma^{-1}}(\sigma f) \qquad \text{for all } \sigma \in G \text{ and } f \in R. \tag{3.1}$$

Let I be an invariant ideal of R. Generalizing terminology introduced in [**10**], let us call a set B of non-zero elements of I a *universal G-Gröbner basis* for I (with respect to $\leq$) if B contains, for every $\sigma \in G$, a Gröbner basis (in the usual sense of the word) for the ideal I with respect to the term ordering $\leq_\sigma$. If the set X of indeterminates is finite, then every invariant ideal of R has a finite universal G-Gröbner basis. By the remark following Lemma 3.10, every universal G-Gröbner basis for an invariant ideal I of R is a Gröbner basis for I. We finish this section by observing:

Lemma 3.14. *Suppose that A is field. If B is a Gröbner basis for the invariant ideal I of R, then*

$$GB = \{\sigma g : \sigma \in G,\ g \in B\}$$

is a universal G-Gröbner basis for I.

Proof. Let $\sigma \in G$ and $f \in I$, $f \neq 0$. Then $\sigma f \in I$, hence there exists $\tau \in G$ and $g \in B$ such that $w \leq \operatorname{lm}(g) \Rightarrow w \leq_\tau \operatorname{lm}(g)$ for all $w \in X^\diamond$, and $\tau \operatorname{lm}(g) | \operatorname{lm}(\sigma f)$. The first condition implies in particular that $\tau \operatorname{lm}(g) = \operatorname{lm}(\tau g)$, hence $\sigma^{-1}\tau \operatorname{lm}(g) = \operatorname{lm}_\sigma(\sigma^{-1}\tau g)$ and $\sigma^{-1} \operatorname{lm}(\sigma f) = \operatorname{lm}_\sigma(f)$ by (3.1). Put $h := \sigma^{-1}\tau g \in GB$. Then $\operatorname{lm}_\sigma(h) | \operatorname{lm}_\sigma(f)$ by the second condition. This shows that GB contains a Gröbner basis for I with respect to $\leq_\sigma$, as required. □

Example 3.15. Suppose that $G = \mathfrak{S}_n$, the group of permutations of $\{1, 2, \dots, n\}$, acting on $X = \{x_1, \dots, x_n\}$ via $\sigma x_i = x_{\sigma(i)}$. The invariant ideal $I = \langle x_1, \dots, x_n \rangle_R$ has Gröbner basis $\{x_1\}$ with respect to the lexicographic ordering; a corresponding (minimal) universal $\mathfrak{S}_n$-Gröbner basis for I is $\{x_1, \dots, x_n\}$.

4. Invariant Chains of Ideals

In this section we describe a relationship between certain chains of increasing ideals in finite-dimensional polynomials rings and invariant ideals of infinite-dimensional polynomial rings. We begin with an abstract setting that is suitable for placing the motivating problem (described in the next section) in a proper context. Throughout this section, m and n range over the set of positive integers. For each n, let R_n be a commutative ring, and assume that R_n is a subring of R_{n+1}, for each n. Suppose that the symmetric group on n letters $\mathfrak{S}_n$ gives an action (not necessarily faithful) on R_n such that $f \mapsto \sigma f \colon R_n \to R_n$ is a ring homomorphism, for each $\sigma \in \mathfrak{S}_n$. Furthermore, suppose that the natural embedding of $\mathfrak{S}_n$ into $\mathfrak{S}_m$ for $n \leq m$ is compatible with the embedding of rings $R_n \subseteq R_m$; that is, if $\sigma \in \mathfrak{S}_n$ and $\widehat{\sigma}$ is the corresponding element in $\mathfrak{S}_m$, then $\widehat{\sigma} R_n = \sigma$. Note that there exists a unique action of $\mathfrak{S}_\infty$ on the ring $R := \bigcup_{n \geq 1} R_n$ which extends the action of each $\mathfrak{S}_n$ on R_n. An ideal of R is *invariant* if $\sigma f \in I$ for all $\sigma \in \mathfrak{S}_\infty$, $f \in I$.

We will need a method for lifting ideals of smaller rings into larger ones, and one such technique is as follows.

Definition 4.1. For $m \geq n$, the *m-symmetrization* $L_m(B)$ of a set B of elements of R_n is the $\mathfrak{S}_m$-invariant ideal of R_m given by

$$L_m(B) = \langle g : g \in B \rangle_{R_m[\mathfrak{S}_m]}$$

In order for us to apply this definition sensibly, we must make sure that the m-symmetrization of an ideal can be defined in terms of generators.

Lemma 4.2. *If B is a set of generators for the ideal $I_B = \langle B \rangle_{R_n}$ of R_n, then $L_m(I_B) = L_m(B)$.*

Proof. Suppose that B generates the ideal $I_B \subseteq R_n$. Clearly, $L_m(B) \subseteq L_m(I_B)$. Therefore, it is enough to show the inclusion $L_m(I_B) \subseteq L_m(B)$. Suppose that $h \in L_m(I_B)$ so that $h = \sum_{j=1}^{s} f_j \cdot \sigma_j h_j$ for elements $f_j \in R_m$, $h_j \in I_B$ and $\sigma_j \in \mathfrak{S}_m$. Next express each $h_j = \sum_{i=1}^{r_j} p_{ij} g_{ij}$ for $p_{ij} \in R_n$ and $g_{ij} \in B$. Substitution into the expression above for h gives us

$$h = \sum_{j=1}^{s} \sum_{i=1}^{r_j} f_j \cdot \sigma_j p_{ij} \cdot \sigma_j g_{ij}.$$

This is easily seen to be an element of $L_m(B)$, completing the proof. □

Example 4.3. Let $S = \mathbb{Q}[t_1, t_2]$, $R_n = \mathbb{Q}[x_1, \ldots, x_n]$, and consider the natural action of $\mathfrak{S}_n$ on R_n. Let Q be the kernel of the homomorphism induced by the map $\phi\colon R_3 \to S$ given by $\phi(x_1) = t_1^2$, $\phi(x_2) = t_2^2$, and $\phi(x_3) = t_1 t_2$. Then, $Q = \langle x_1 x_2 - x_3^2 \rangle$, and $L_4(Q) \subseteq R_4$ is generated by the following 12 polynomials:

$$\begin{gathered} x_1x_2 - x_3^2,\ x_1x_2 - x_4^2,\ x_1x_3 - x_2^2,\ x_1x_3 - x_4^2, \\ x_1x_4 - x_3^2,\ x_1x_4 - x_2^2,\ x_2x_3 - x_1^2,\ x_2x_3 - x_4^2, \\ x_2x_4 - x_1^2,\ x_2x_4 - x_3^2,\ x_3x_4 - x_1^2,\ x_3x_4 - x_2^2. \end{gathered}$$

We would also like a way to project a set of elements in R_m down to a smaller ring R_n ($n \leq m$).

Definition 4.4. Let $B \subseteq R_m$ and $n \leq m$. The *n-projection* $P_n(B)$ of B is the $\mathfrak{S}_n$-invariant ideal of R_n given by

$$P_n(B) = \langle g : g \in B \rangle_{R_m[\mathfrak{S}_m]} \cap R_n.$$

We now consider increasing chains $I_\circ$ of ideals $I_n \subseteq R_n$:

$$I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots,$$

simply called *chains* below. Of course, such chains will usually fail to stabilize since they are ideals in larger and larger rings. However, it is possible for these ideals to stabilize "up to the action of the symmetric group," a concept we make clear below. For the purposes of this work, we will only consider a special class of chains; namely, a *symmetrization invariant chain* (resp. *projection invariant chain*) is one for which $L_m(I_n) \subseteq I_m$ (resp. $P_n(I_m) \subseteq I_n$) for all $n \leq m$. If $I_\circ$ is both a symmetrization and a projection invariant chain, then it will be simply called an *invariant chain*. We will encounter some concrete invariant chains in the next section. The stabilization definition alluded to above is as follows.

Definition 4.5. A symmetrization invariant chain of ideals $I_\circ$ as above *stabilizes modulo the symmetric group* (or simply *stabilizes*) if there exists a positive integer N such that

$$L_m(I_n) = I_m \qquad \text{for all } m \geq n > N.$$

To put it another way, accounting for the natural action of the symmetric group, the ideals I_n are the same for large enough n. Let us remark that if for a symmetrization invariant chain $I_\circ$, there is some integer N such that $L_m(I_N) = I_m$ for all $m > N$, then $I_\circ$ stabilizes. This follows from the inclusions

$$I_m = L_m(I_N) \subseteq L_m(I_n) \subseteq I_m, \quad n > N.$$

Any chain $I_\circ$ naturally gives rise to an ideal $\mathcal{I}(I_\circ)$ of $R = \bigcup_{n \geq 1} R_n$ by way of

$$\mathcal{I}(I_\circ) := \bigcup_{n \geq 1} I_n.$$

Conversely, if I is an ideal of R, then

$$I_n = \mathcal{J}_n(I) := I \cap R_n$$

defines the components of a chain $\mathcal{J}(I) := I_\circ$. Clearly, for any ideal $I \subseteq R$, we have $\mathcal{I} \circ \mathcal{J}(I) = I$, but, as is easily seen, it is not true in general that $\mathcal{J} \circ \mathcal{I}(I_\circ) = I_\circ$. However, for invariant chains, this relationship does hold, as the following straightforward lemma describes.

Lemma 4.6. *There is a one-to-one, inclusion-preserving correspondence between invariant chains $I_\circ$ and invariant ideals I of R given by the maps $\mathcal{I}$ and $\mathcal{J}$.* □

For the remainder of this section we consider the case where, for a commutative Noetherian ring A, we have $R_n = A[x_1, \dots, x_n]$ for each n, endowed with the natural action of $\mathfrak{S}_n$ on the indeterminates $x_1, \dots, x_n$. Then $R = A[X^\diamond]$ where $X = \{x_1, x_2, \dots\}$. We use the results of the previous section to demonstrate the following.

Theorem 4.7. *Every symmetrization invariant chain stabilizes modulo the symmetric group.*

Proof. Given a symmetrization invariant chain, construct the invariant ideal $I = \mathcal{I}(I_\circ)$ of R. One would now like to apply Theorem 1.1, however, more care is needed to prove stabilization. Let $\leq$ be a well-ordering of X of order type ω, and let B be a finite Gröbner basis for I with respect to the corresponding term ordering $\leq_{\text{lex}}$ of $X^\diamond$. (Theorem 2.20 and Proposition 3.11.) Choose a positive integer N such that $B \subseteq I_N$; we claim that $I_m = L_m(I_N)$ for all $m \geq N$. Let $f \in I_m$, $f \neq 0$. By the equivalence of (1) and (3) in Lemma 3.10 we have $f \xrightarrow[B]{*} 0$. Hence by Lemma 3.8 there are $g_1, \dots, g_n \in B$, $h_1, \dots, h_n \in R$, as well as $\sigma_1, \dots, \sigma_n \in \mathfrak{S}_\infty$, such that

$$f = h_1\sigma_1 g_1 + \cdots + h_n\sigma_n g_n \quad \text{and} \quad \operatorname{lm}(f) = \max_i \operatorname{lm}(h_i\sigma_i g_i).$$

By Remark 3.9 we may assume that in fact $\sigma_i \in \mathfrak{S}_m$ for each i. Moreover $\operatorname{lm}(h_i) \leq_{\text{lex}} \operatorname{lm}(f)$, hence $|\operatorname{lm}(h_i)| \leq |\operatorname{lm}(f)| \leq m$, for each i. Therefore $h_i \in R_m$ for each i. This shows that $f \in L_m(B) \subseteq L_m(I_N)$ as desired. □

5. A Chemistry Motivation

We can now discuss the details of the basic problem that is of interest to us. It was brought to our attention by Bernd Sturmfels, who, in turn, learned about it from Andreas Dress.

Fix a natural number $k \geq 1$. Given a set S we denote by $\langle S\rangle^k$ the set of all ordered k-element subsets of S, that is, $\langle S\rangle^k$ is the set of all k-tuples $\boldsymbol{u} = (u_1, \dots, u_k) \in S^k$ with pairwise distinct $u_1, \dots, u_k$. We also just write $\langle n\rangle^k$ instead of $\langle\{1, \dots, n\}\rangle^k$. Let K be a field, and for $n \geq k$ consider the polynomial ring

$$R_n = K\big[\{x_{\boldsymbol{u}}\}_{\boldsymbol{u}\in\langle n\rangle^k}\big].$$

We let $\mathfrak{S}_n$ act on $\langle n\rangle^k$ by

$$\sigma(u_1, \dots, u_k) = \big(\sigma(u_1), \dots, \sigma(u_k)\big).$$

This induces an action $(\sigma, x_{\boldsymbol{u}}) \mapsto \sigma x_{\boldsymbol{u}} = x_{\sigma\boldsymbol{u}}$ of $\mathfrak{S}_n$ on the indeterminates $x_{\boldsymbol{u}}$, which we extend to an action of $\mathfrak{S}_n$ on R_n in the natural way. We also put $R = \bigcup_{n\geq k} R_n$. Note that

$$R = K\big[\{x_{\boldsymbol{u}}\}_{\boldsymbol{u}\in\langle\Omega\rangle^k}\big],$$

where $\Omega = \{1, 2, 3, \dots\}$ is the set of positive integers, and that the actions of $\mathfrak{S}_n$ on R_n combine uniquely to an action of $\mathfrak{S}_\infty$ on R. Let now $f(y_1, \dots, y_k) \in K[y_1, \dots, y_k]$, let $t_1, t_2, \dots$ be an infinite sequence of pairwise distinct indeterminates over K, and for $n \geq k$ consider the K-algebra homomorphism

$$\phi_n \colon R_n \to K[t_1, \dots, t_n], \qquad x_{(u_1,\dots,u_k)} \mapsto f(t_{u_1}, \dots, t_{u_k}).$$

The ideal

$$Q_n = \ker\, \phi_n$$

of R_n determined by such a map is the prime ideal of algebraic relations between the quantities $f(t_{u_1}, \dots, t_{u_k})$. Such ideals arise in chemistry [**9, 15, 16**]; of specific interest there is when f is a Vandermonde polynomial $\prod_{i<j}(y_i - y_j)$. In this case, the ideals Q_n correspond to relations among a series of experimental measurements. One would then like to understand the limiting behavior of such relations, and in particular, to see that they stabilize up to the action of the symmetric group.

EXAMPLE 5.1. The permutation $\sigma = (1\,2\,3) \in \mathfrak{S}_3$ acts on the elements

$$(1,2),\ (2,1),\ (1,3),\ (3,1),\ (2,3),\ (3,2)$$

of $\langle 3\rangle^2$ to give

$$(2,3),\ (3,2),\ (2,1),\ (1,2),\ (3,1),\ (1,3),$$

respectively. Let $f(t_1,t_2) = t_1^2 t_2$. Then the action of σ on the valid relation $x_{12}^2 x_{31} - x_{13}^2 x_{21} \in Q_3$ gives us another relation $x_{23}^2 x_{12} - x_{21}^2 x_{32} \in Q_3$.

It is easy to see that by construction, the chain $Q_\circ$ of ideals

$$Q_k \subseteq Q_{k+1} \subseteq \cdots \subseteq Q_n \subseteq \cdots$$

(which we call the chain of ideals *induced by the polynomial* f) is an invariant chain. As in the proof of Theorem 4.7, we would like to form the ideal $Q = \bigcup_{n\geq k} Q_n$ of the infinite-dimensional polynomial ring $R = \bigcup_{n\geq k} R_n$, and then apply a finiteness theorem to conclude that $Q_\circ$ stabilizes in the sense mentioned above (Definition 4.5). For $k = 1$, Theorem 4.7 indeed does the job. Unfortunately however, this simple-minded approach fails for $k \geq 2$:

PROPOSITION 5.2. *For $k \geq 2$, the $R[\mathfrak{S}_\infty]$-module R is not Noetherian.*

PROOF. Let us make the dependence on k explicit and denote R by $R^{(k)}$. Then

$$x_{(u_1,\dots,u_k,u_{k+1})} \mapsto x_{(u_1,\dots,u_k)}$$

defines a surjective K-algebra homomomorphism $\pi_k\colon R^{(k+1)} \to R^{(k)}$ with invariant kernel. Hence if $R^{(k+1)}$ is Noetherian as an $R[\mathfrak{S}_\infty]$-module, then so is $R^{(k)}$; thus it suffices to prove the proposition in the case $k = 2$. Suppose therefore that $k = 2$. By Lemma 3.13 it is enough to produce an infinite bad sequence for the quasi-ordering $|_{\mathfrak{S}_\infty}$ of $X^\diamond$, where $X = \{x_{\boldsymbol{i}} : \boldsymbol{i} \in \langle\Omega\rangle^2\}$. For this, consider the sequence of monomials

$$\begin{aligned}
s_3 &= x_{(1,2)}x_{(3,2)}x_{(3,4)} \\
s_4 &= x_{(1,2)}x_{(3,2)}x_{(4,3)}x_{(4,5)} \\
s_5 &= x_{(1,2)}x_{(3,2)}x_{(4,3)}x_{(5,4)}x_{(6,7)} \\
&\ \,\vdots \\
s_n &= x_{(1,2)}x_{(3,2)}x_{(4,3)}\cdots x_{(n,n-1)}x_{(n,n+1)} \qquad (n = 3,4,\dots) \\
&\ \,\vdots
\end{aligned}$$

Now for $n < m$ and any $\sigma \in \mathfrak{S}_\infty$, the monomial σs_n does not divide s_m. To see this, suppose otherwise. Note that $x_{(1,2)}$, $x_{(3,2)}$ is the only pair of indeterminates which divides s_n or s_m and has the form $x_{(i,j)}$, $x_{(l,j)}$ $(i,j,l \in \Omega)$. Therefore $\sigma(2) = 2$, and either $\sigma(1) = 1$, $\sigma(3) = 3$, or $\sigma(1) = 3$, $\sigma(3) = 1$. But since 1 does not appear as the second component j of a factor $x_{(i,j)}$ of s_m, we have $\sigma(1) = 1$, $\sigma(3) = 3$. Since $x_{(4,3)}$ is the only indeterminate dividing s_n or s_m of the form $x_{(i,3)}$ with $i \in \Omega$, we get $\sigma(4) = 4$; since $x_{(5,4)}$ is the only indeterminate dividing s_n or s_m of the form $x_{(i,4)}$ with $i \in \Omega$, we get $\sigma(5) = 5$; etc. Ultimately this yields $\sigma(i) = i$ for all $i = 1,\dots,n$. But the only indeterminate dividing s_m of the form $x_{(n,j)}$ with $j \in \Omega$ is $x_{(n,n-1)}$, hence the factor $\sigma x_{(n,n+1)} = x_{(n,\sigma(n+1))}$ of σs_n does not divide s_m. This shows that $s_3, s_4, \dots$ is a bad sequence for the quasi-ordering $|_{\mathfrak{S}_\infty}$, as claimed. $\square$

REMARK 5.3. The construction of the infinite bad sequence $s_3, s_4, \dots$ in the proof of the previous proposition was inspired by an example in [**7**].

5.1. A criterion for stabilization. Our next goal is to give a condition for the chain $Q_\circ$ to stabilize. Given $g \in R$, we define the *variable size* of g to be the number of distinct indeterminates $x_{\boldsymbol{u}}$ that appear in g. For example, $g = x_{12}^5 + x_{45}x_{23} + x_{45}$ has variable size 3.

LEMMA 5.4. *A chain of ideals $Q_\circ$ induced by a polynomial $f \in K[y_1,\dots,y_k]$ stabilizes modulo the symmetric group if and only if there exist integers M and N such that for all $n > N$, there are generators for Q_n with variable sizes at most M. Moreover, in this case a bound for stabilization is given by* $\max(N, kM)$.

Proof. Suppose M and N are integers with the stated property. To see that $Q_\circ$ stabilizes, since $Q_\circ$ is an invariant chain, we need only verify that $N' = \max(N, kM)$ is such that $Q_m \subseteq L_m(Q_n)$ for $m \geq n > N'$. For this inclusion, it suffices that each generator in a generating set for the ideal Q_m of R_m is in $L_m(Q_n)$. Since $m > N$, there are generators B for Q_m with variable sizes at most M. If $g \in B$, then there are at most kM different integers appearing as subscripts of indeterminates in g. We can form a permutation $\sigma \in \mathfrak{S}_m$ such that $\sigma g \in R_{N'}$ and thus in R_n. But then $\sigma g \in P_n(Q_m) \subseteq Q_n$ so that $g = \sigma^{-1}\sigma g \in L_m(Q_n)$ as desired.

Conversely, suppose that $Q_\circ$ stabilizes. Then there exists an N such that $Q_m = L_m(Q_N)$ for all $m > N$. Let B be any finite generating set for Q_N. Then for all $m > N$, $Q_m = L_m(B)$ is generated by elements of bounded variable size, by Lemma 4.2. $\square$

Although this condition is a very simple one, it will prove useful. Below we will apply it together with a preliminary reduction to the case that each indeterminate $y_1, \dots, y_k$ actually occurs in the polynomial f, which we explain next. For this we let $\pi_k \colon R^{(k+1)} \to R^{(k)}$ be the surjective K-algebra homomorphism defined in the proof of Proposition 5.2. We write $Q^{(k)}$ for Q, and considering $f \in K[y_1, \dots, y_k]$ as an element of $K[y_1, \dots, y_k, y_{k+1}]$, we also let $Q^{(k+1)}$ be the kernel of the K-algebra homomorphsm

$$R^{(k+1)} \to K[t_1, t_2, \dots], \qquad x_{(u_1,\dots,u_k,u_{k+1})} \mapsto f(t_{u_1}, \dots, t_{u_k}, t_{u_{k+1}}) \\ (= f(t_{u_1}, \dots, t_{u_k})).$$

Note that $\pi_k(Q^{(k+1)}) = Q^{(k)}$, and the ideal $\ker \pi_k$ of $R^{(k+1)}$ is generated by the elements

$$x_{(u_1,\dots,u_k,i)} - x_{(u_1,\dots,u_k,j)} \qquad (i, j \in \Omega),$$

in particular $\ker \pi_k \subseteq Q^{(k+1)}$. It is easy to see that as an $R^{(k+1)}[\mathfrak{S}_\infty]$-module, $\ker \pi_k$ is generated by the single element $x_{(1,\dots,k,k+1)} - x_{(1,\dots,k,k+2)}$. These observations now yield:

Lemma 5.5. *Suppose that the invariant ideal $Q^{(k)}$ of $R^{(k)}$ is finitely generated as an $R^{(k)}[\mathfrak{S}_\infty]$-module. Then the invariant ideal $Q^{(k+1)}$ of $R^{(k+1)}$ is finitely generated as an $R^{(k+1)}[\mathfrak{S}_\infty]$-module.* $\square$

We let $\mathfrak{S}_k$ act on $\langle\Omega\rangle^k$ by

$$\tau(u_1, \dots, u_k) = (u_{\tau(1)}, \dots, u_{\tau(k)}) \qquad \text{for } \tau \in \mathfrak{S}_k,\ (u_1, \dots, u_k) \in \langle\Omega\rangle^k.$$

This action gives rise to an action of $\mathfrak{S}_k$ on $\{x_{\boldsymbol{u}}\}_{\boldsymbol{u} \in \langle\Omega\rangle^k}$ by $\tau x_{\boldsymbol{u}} = x_{\tau\boldsymbol{u}}$, which we extend to an action of $\mathfrak{S}_k$ on R in the natural way. We also let $\mathfrak{S}_k$ act on $K[y_1, \dots, y_k]$ by $\tau f(y_1, \dots, y_k) = f(y_{\tau(1)}, \dots, y_{\tau(k)})$. Note that

$$\tau Q_k \subseteq \tau Q_{k+1} \subseteq \cdots \subseteq \tau Q_n \subseteq \cdots$$

is the chain induced by τf. Using the lemma above we obtain:

Corollary 5.6. *Let $f \in K[y_1, \dots, y_k]$. There are $i \in \{0, \dots, k\}$ and $\tau \in \mathfrak{S}_k$ such that $\tau f \in K[y_1, \dots, y_i]$ and each of the indeterminates $y_1, \dots, y_i$ occurs in τf. If the chain of ideals induced by the polynomial τf stabilizes, then so does the chain of ideals induced by f.* $\square$

5.2. Chains induced by monomials. If the given polynomial f is a monomial, then the homomorphism ϕ_n from above produces a (homogeneous) toric kernel Q_n. In particular, there is a finite set of binomials that generate Q_n (see [**17**]). Although a proof for the general toric case eludes us, we do have the following.

Theorem 5.7. *The sequence of kernels induced by a square-free monomial $f \in K[y_1, \dots, y_k]$ stabilizes modulo the symmetric group. Moreover, a bound for when stabilization occurs is $N = 4k$.*

To prepare for the proof of this result, we discuss in detail the toric encoding associated to our problem (see [**17**, Chapter 14] for more details). By Corollary 5.6, we may assume that $f = y_1 \cdots y_k$. Then $g - \tau g \in Q$ for all $g \in R$. We say that $\boldsymbol{u} = (u_1, \dots, u_k) \in \langle\Omega\rangle^k$ is *sorted* if $u_1 < \cdots < u_k$, and *unsorted* otherwise; similarly we say that $x_{\boldsymbol{u}}$ is sorted (unsorted) if $\boldsymbol{u}$ is sorted (unsorted, respectively). For example, x_{135} is a sorted indeterminate, whereas x_{315} is not. Consider the set of vectors

$$\mathcal{A}_n = \left\{(i_1, \dots, i_n) \in \mathbb{Z}^n : i_1 + \cdots + i_n = k,\ 0 \leq i_1, \dots, i_n \leq 1\right\}.$$

View $\mathcal{A}_n$ as an n-by-$\binom{n}{k}$ matrix entries with 0 and 1, whose with columns are indexed by sorted indeterminates $x_{\boldsymbol{u}}$ and whose rows are indexed by t_i ($i = 1, \dots, n$). (See Example 5.9 below.) Let $\operatorname{sort}(\cdot)$ denote the operator which takes any word in $\{1,\dots,n\}^*$ and sorts it in increasing order. By [**17**, Remark 14.1], the toric ideal $I_{\mathcal{A}_n}$ associated to $\mathcal{A}_n$ is generated (as K-vector space) by the binomials $x_{\boldsymbol{u}_1}\cdots x_{\boldsymbol{u}_r} - x_{\boldsymbol{v}_1}\cdots x_{\boldsymbol{v}_r}$, where $r\in\mathbb{N}$ and the $\boldsymbol{u}_i$, $\boldsymbol{v}_j$ are sorted elements of $\langle n\rangle^k$ such that $\operatorname{sort}(\boldsymbol{u}_1\cdots\boldsymbol{u}_r)=\operatorname{sort}(\boldsymbol{v}_1\cdots\boldsymbol{v}_r)$. In particular, we have $I_{\mathcal{A}_n}\subseteq Q_n$. Let B be any set of generators for the ideal $I_{\mathcal{A}_n}$.

LEMMA 5.8. *A generating set for the ideal Q_n of R_n is given by*

$$S = B\cup\{x_{\boldsymbol{u}} - x_{\tau\boldsymbol{u}}:\ \tau\in\mathfrak{S}_k,\ \boldsymbol{u} \text{ is sorted}\}.$$

PROOF. Elements of Q_n are of the form $g = x_{\boldsymbol{u}_1}\cdots x_{\boldsymbol{u}_r} - x_{\boldsymbol{v}_1}\cdots x_{\boldsymbol{v}_r}$, in which the $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ are ordered k-element subsets of $\{1,\dots,n\}$ such that $\operatorname{sort}(\boldsymbol{u}_1\cdots\boldsymbol{u}_r)=\operatorname{sort}(\boldsymbol{v}_1\cdots\boldsymbol{v}_r)$. We induct on the number t of $\boldsymbol{u}_i$ and $\boldsymbol{v}_j$ that are not sorted. If $t=0$, then $g\in I_{\mathcal{A}_n}$, and we are done. Suppose now that $t>0$ and assume without loss of generality that $\boldsymbol{u}_1$ is not sorted. Let $\tau\in\mathfrak{S}_k$ be such that $\tau\boldsymbol{u}_1$ is sorted, and consider the element $h = x_{\tau\boldsymbol{u}_1}x_{\boldsymbol{u}_2}\cdots x_{\boldsymbol{u}_r} - x_{\boldsymbol{v}_1}\cdots x_{\boldsymbol{v}_r}$ of Q_n. This binomial involves $t-1$ unsorted indeterminates, and therefore, inductively, can be expressed in terms of S. But then

$$g = h - (x_{\tau\boldsymbol{u}_1} - x_{\boldsymbol{u}_1})x_{\boldsymbol{u}_2}\cdots x_{\boldsymbol{u}_r}$$

can as well, completing the proof. ☐

EXAMPLE 5.9. Let $k=2$ and $n=4$. Then

	x_{12}	x_{13}	x_{14}	x_{23}	x_{24}	x_{34}
t_1	1	1	1	0	0	0
t_2	1	0	0	1	1	0
t_3	0	1	0	1	0	1
t_4	0	0	1	0	1	1

represents the matrix associated to $\mathcal{A}_4$. The ideal $I_{\mathcal{A}_4}$ is generated by the two binomials $x_{13}x_{24} - x_{12}x_{34}$ and $x_{14}x_{23} - x_{12}x_{34}$. Hence Q_4 is generated by these two elements along with

$$\{x_{12}-x_{21}, x_{13}-x_{31}, x_{14}-x_{41}, x_{23}-x_{32}, x_{24}-x_{42}, x_{34}-x_{43}\}.$$

We are now in a position to prove Theorem 5.7.

PROOF OF THEOREM 5.7. By Lemma 5.4, we need only show that there exist generators for Q_n which have bounded variable sizes. Using [**17**, Theorem 14.2], it follows that $I_{\mathcal{A}_n}$ has a quadratic (binomial) Gröbner basis for each n (with respect to some term ordering of R_n). By Lemma 5.8, there is a set of generators for Q_n with variable sizes at most 4. This proves the theorem. ☐

We close with a conjecture that generalizes Theorem 5.7.

CONJECTURE 5.10. *The sequence of kernels induced by a monomial f stabilizes modulo the symmetric group.*

6. Acknowledgment

We would like to thank Bernd Sturmfels for bringing the problem found in Section 5 (originating from Andreas Dress) to our attention and for making us aware of Theorem 14.2 in [**17**].

References

[1] W. Adams, P. Loustaunau, *An Introduction to Gröbner Bases,* Graduate Studies in Mathematics, vol. 3, American Mathematical Society, Providence, RI, 1994.

[2] G. Ahlbrandt, M. Ziegler, *Quasi-finitely axiomatizable totally categorical theories,* Stability in model theory (Trento, 1984), Ann. Pure Appl. Logic **30** (1986), no. 1, 63–82.

[3] A. R. Camina, D. M. Evans, *Some infinite permutation modules*, Quart. J. Math. Oxford Ser. (2) **42** (1991), no. 165, 15–26.

[4] P. Erdős, R. Rado, *A theorem on partial well-ordering of sets of vectors,* J. London Math. Soc. **34** (1959), 222–224.

[5] D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Graduate Texts in Mathematics, vol. 203, Springer-Verlag, New York, 1995.

[6] G. Higman, *Ordering by divisibility in abstract algebras*, Proc. London Math. Soc. (3) **2** (1952), 326–336.

[7] T. A. Jenkyns, C. St. J. A. Nash-Williams, *Counterexamples in the theory of well-quasi-ordered sets.* in: F. Harary (ed.), *Proof Techniques in Graph Theory* (*Proc. Second Ann Arbor Graph Theory Conf., Ann Arbor, Mich., 1968*), pp. 87–91, Academic Press, New York, 1969.

[8] J. B. Kruskal, *The theory of well-quasi-ordering: A frequently discovered concept*, J. Combinatorial Theory Ser. A **13** (1972), 297–305.
[9] A. Mead, E. Ruch, A. Schönhofer, *Theory of chirality functions, generalized for molecules with chiral ligands*. Theor. Chim. Acta **29** (1973), 269–304.
[10] R. Michler, *Gröbner bases of symmetric quotients and applications,* in: C. Christensen et al. (eds.), *Algebra, Arithmetic and Geometry with Applications* (*West Lafayette, IN, 2000*)*,* pp. 627–637, Springer-Verlag, Berlin, 2004.
[11] E. Milner, *Well-quasi-ordering of sequences of ordinal numbers,* J. London Math. Soc. **43** (1968), 291–296.
[12] C. St. J. A. Nash-Williams, *On well-quasi-ordering finite trees*, Proc. Cambridge Philos. Soc. **59** (1963), 833–835.
[13] ———, *On well-quasi-ordering transfinite sequences*, Proc. Cambridge Philos. Soc. **61** (1965), 33–39.
[14] R. Rado, *Partial well-ordering of sets of vectors,* Mathematika **1** (1954), 89–95.
[15] E. Ruch, A. Schönhofer, *Theorie der Chiralitätsfunktionen*, Theor. Chim. Acta **19** (1970), 225–287.
[16] E. Ruch, A. Schönhofer, I. Ugi, *Die Vandermondesche Determinante als Näherungsansatz für eine Chiralitätsbeobachtung, ihre Verwendung in der Stereochemie und zur Berechnung der optischen Aktivität*, Theor. Chim. Acta **7** (1967), 420–432.
[17] B. Sturmfels, *Gröbner Bases and Convex Polytopes*, AMS University Lecture Series, vol. 8, American Mathematical Society, Providence, RI, 1996.

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, Chicago, IL 60607.

E-mail address: maschenb@math.uic.edu

Department of Mathematics, University of California, Berkeley, CA 94720.

E-mail address: chillar@math.berkeley.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Multivariate Fuss-Catalan numbers and B-quasisymmetric functions

Jean-Christophe Aval

Abstract. We study the ideal generated by constant-term free B-quasisymmetric polynomials, and prove that the quotient of the polynomial ring by this ideal has dimension given by $\frac{1}{2n+1}\binom{3n}{n}$, the number of ternary trees, or Fuss-Catalan number of order 3. This leads us to introduce and study multivariate Fuss-Catalan numbers, whose combinatorial interpretation is given by some statistics on ternary trees and plane paths.

Résumé. Nous étudions l'idéal engendré par les polynômes B-quasisymétriques (sans terme constant), et prouvons que le quotient de l'anneau des polynômes par cet idéal est de dimension $\frac{1}{2n+1}\binom{3n}{n}$, le nombre d'arbres ternaires, ou nombre de Fuss-Catalan d'ordre 3. Nous en profitons pour introduire et étudier combinatoirement certains nombres de Fuss-Catalan multivariés, ce qui fait apparaître une bi-statistique sur les arbres ternaires et certains chemins du plan.

1. Introduction

To start with, we recall a small part of the story of the study of ideals and quotients related to symmetric or quasisymmetric polynomials. The root of this work is a result of Artin [**1**]. Let us consider the set of variables $X_n = x_1, x_2, \ldots, x_n$. The space of polynomials in the variables X_n with rational coefficients is denoted by $\mathbb{Q}[X_n]$. The subspace of symmetric polynomials is denoted by Sym_n. If $\mathbf{V}$ is a subset of the polynomial ring, we denote by $\langle \mathbf{V}^+ \rangle$ the ideal generated by elements of a $\mathbf{V}$ with no constant term. Artin's result is given by:

$$\dim \mathbb{Q}[X_n]/\langle Sym_n^+ \rangle = n!\,. \tag{1.1}$$

Another, more recent, part of the story deals with quasisymmetric polynomials. The space $QSym_n \subset \mathbb{Q}[X_n]$ of quasisymmetric polynomials was introduced by Gessel [**13**] as generating functions for Stanley's P-partitions [**21**]. This is the starting point of many recent works in several areas of combinatorics [**9, 12, 16, 17, 22**].

In [**4, 5**], Aval *et. al.* study the problem analogous to Artin's work in the case of quasisymmetric polynomials. Their main result is that the dimension of the quotient is given by Catalan numbers:

$$\dim \mathbb{Q}[X_n]/\langle QSym_n^+ \rangle = C_n = \frac{1}{n+1}\binom{2n}{n}. \tag{1.2}$$

An interesting axis of research is the extension of these results to 2 sets of variables. Let $\mathcal{A}_n$ denote the alphabet

$$\mathcal{A}_n = x_1, y_1, x_2, y_2, \ldots, x_n, y_n.$$

Since symmetric (resp. quasisymmetric) polynomials may be seen as S_n-invariants under the action that permutes variables (resp. under Hivert's action [**16**]), one can define diagonal analogues by letting S_n

---

2000 *Mathematics Subject Classification.* Primary 05E15; Secondary 05A15.

Key words and phrases. algebraic combinatorics, quasisymmetric functions, Gröbner bases, enumeration of paths and trees.

This research has been supported by EC's IHRP Programme, within the Research Training Network "Algebraic Combinatorics in Europe," grant HPRN-CT-2001-00272.

act simulaneously on the x's and y's. We then obtain the space $DSym_n$ (resp. $DQSym_n$) of diagonally symmetric (resp. quasisymmetric) functions.

The diagonal coinvariant space $\mathbb{Q}[\mathcal{A}_n]/\langle DSym_n^+\rangle$ has been studied extensively in the last 15 years by several authors [**8, 11, 14**]. A great achievment in this area is Haiman's proof of the following equality (*cf.* [**14**]):

$$\dim \mathbb{Q}[\mathcal{A}_n]/\langle DSym_n^+\rangle = (n+1)^{n-1}\,.$$

The space $DQSym_n$ was introduced in [**19**], then recently studied in [**7**], [**18**], and [**6**], where the coinvariant space is investigated, and a conjectural basis is presented.

To end this presentation, we introduce the space $QSym_n(B)$ of B-quasisymmetric polynomials, which is the focus of this article. This space, whose definition appears implicitly in [**19**], is studied with more details in [**7**]. A precise definition will be given in the next section, and we only mention here that $QSym_n(B)$ is a subspace (and in fact a subalgebra, *cf.* [**7**]) of $DQSym_n$.

We now state the main result of this work, which appears as a generalization of equation (1.2).

Theorem 1.1.

$$\dim \mathbb{Q}[\mathcal{A}_n]/\langle QSym_n(B)^+\rangle = \frac{1}{2n+1}\binom{3n}{n}. \tag{1.3}$$

Observe that in Equations (1.2) and (1.3), the dimensions $\frac{1}{n+1}\binom{2n}{n}$ and $\frac{1}{2n+1}\binom{3n}{n}$ are respectively the numbers of binary and ternary trees (*cf.* [**20**]). Since we deal with polynomials in two alphabets (and since the ideal $\langle QSym_n(B)^+\rangle$ is homogeneous), we can study the bigraded version of Equation (1.3). More precisely, we look at the subspace of $\mathbb{Q}[\mathcal{A}_n]/\langle QSym_n(B)^+\rangle$ of polynomials of degree k in $x_1,\dots,x_n$ and degree l in $y_1,\dots,y_n$, and consider its dimension, which we denote by $B(n,k,l)$. It appears that these numbers present their own interest, which led us to study them.

Let us now give the plan of this article. We have decided to deal first with the combinatorial part, *i.e.* the study of the numbers $B(n,k,l)$, which is the subject of the next section, and the algebraic part is developped in the last section of this paper.

Remark. This paper is the extended abstract of our work. More details and the complete proofs (here are ony given the sketches of some proofs) can be found in [**2, 3**].

2. Multivariate Fuss-Catalan numbers

2.1. Catalan triangle, binary trees, and Dyck paths.

The *Catalan numbers*

$$C(n) = \frac{1}{n+1}\binom{2n}{n}$$

are integers that appear in many combinatorial problems. These numbers first arose in the work of Catalan as the number of triangulations of a polygon by mean of non-intersecting diagonals. Stanley [**21, 23**] maintains a dynamic list of exercises related to Catalan numbers, including (at this date) 127 combinatorial interpretations.

Closely related to Catalan numbers are *ballot numbers*. To serve our purpose, we shall neither state the so-called *ballot problem*, nor give an explicit formula, but we introduce integers $B(n,k)$ for $(n,k)\in\mathbb{N}^*\times\mathbb{N}$ defined by the following recurrence:

- $B(1,0)=1$
- $\forall n>1$ and $0\le k<n$, $B(n,k)=\sum_{i=0}^{k} B(n-1,i)$
- $\forall k\ge n$, $B(n,k)=0$.

Observe that the recursive formula in the second condition is equivalent to:

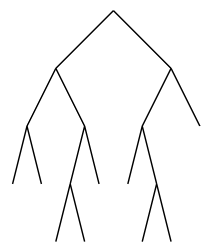
$$B(n,k) = B(n-1,k) + B(n,k-1). \tag{2.1}$$

We shall present the $B(n,k)$'s by the following triangular representation (zero entries are omitted) where moving down increases n and moving right increases k.

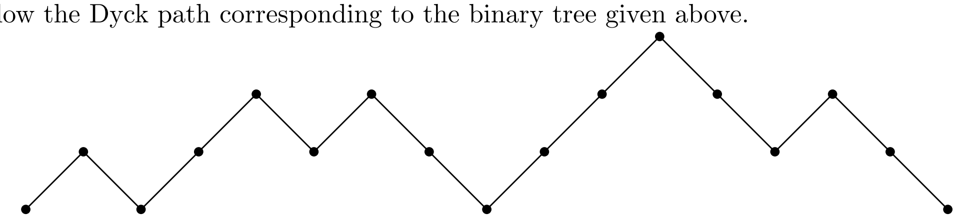
1					
1	1				
1	2	2			
1	3	5	5		
1	4	9	14	14	
1	5	14	28	42	42

The crucial observation is that computing the horizontal sums of these integers give : 1, 2, 5, 14, 42, 132. We recognize the first terms of the Catalan series, and this fact will be proven in Proposition 2.1, after introducing combinatorial objects.

A *binary tree* is a tree in which every internal node has exactly 2 sons. The number of binary trees with n internal nodes is given by the n-th Catalan number.



A *Dyck path* is a path consisting of steps $(1,1)$ and $(1,-1)$, starting from $(0,0)$, ending at $(2n,0)$, and remaining above the line $y=0$. The number of Dyck paths of length $2n$ is also given by the n-th Catalan number. More precisely, the depth-first search of the tree gives a bijection between binary trees and Dyck paths: we associate to each external node (except the left-most one) a $(1,1)$ step and to each internal node a $(1,-1)$ step by searching recursively the left son, then the right son, then the current node. As an example, we show below the Dyck path corresponding to the binary tree given above.



An inportant parameter in our study will be the length of the right-most sequence of $(1,-1)$ of the path. This parameter equals 2 in our example. Observe that under the correspondence between paths and trees, this parameter corresponds to the length of the right-most string of right sons in the tree. We shall use the expressions *last down sequence* and *last right string*, for these parts of the path and of the tree.

Now we come to the announced result. It is well-known and simple, but is the starting point of our work.

Proposition 2.1. *We have the following equality:*

$$\sum_{k=0}^{n-1} B(n,k) = C(n) = \frac{1}{n+1}\binom{2n}{n}.$$

Proof. Let us denote by $\mathcal{C}_{n,k}$ the set of Dyck paths of length $2n$ with a last down sequence of length equal to $n-k$.

We shall prove that $B(n,k)$ is the cardinality of $\mathcal{C}_{n,k}$.

The proof is done recursively on n. If $n=0$, this is trivial. If $n>0$, let us suppose that $B(n-1,k)$ is the cardinality of $\mathcal{C}_{n-1,k}$ for $0 \le k < n-1$. Let us consider an element of $\mathcal{C}_{n,k}$. If we erase the last step $(1,1)$ and the following step $(1,-1)$, we obtain a Dyck path of length $2(n-1)$ and with a last decreasing sequence of length $n-l \ge n-k$. If we keep track of the integer k, we obtain a bijection between $\mathcal{C}_{n,k}$ and $\cup_{l\le k}\mathcal{C}_{n-1,l}$. □

2.2. Fuss-Catalan tetrahedron and ternary trees. This subsection, which is the heart of this part of the work, is the study of a 3-dimensional analogue of the Catalan triangle presented in the previous section. We consider exactly the same recurrence, and let the array grow, not in 2, but in 3 dimensions.

More precisely, we introduce the sequence $B_3(n,k,l)$ indexed by integers n, k and l, and defined recursively by:

- $B_3(1,0,0)=1$
- $\forall n>1,\ k+l<n,\ B_3(n,k,l)=\sum_{0\le i\le k, 0\le j\le l} B_3(n-1,i,j)$
- $\forall k+l\ge n,\ B_3(n,k,l)=0.$

Observe that the recursive formula in the second condition is equivalent to:

$$B_3(n,k,l)=B_3(n-1,k,l)+B_3(n,k-1,l)+B_3(n,k,l-1)-B_3(n,k-1,l-1) \tag{2.2}$$

and this expression can be used to make some computations lighter, but the presentation above explains more about the generalization of the definition of the ballot numbers $B(n,k)$.

Because of the planar structure of the sheet of paper, we are led to present the tetrahedron of $B_3(n,k,l)$'s by its sections with a given n.

$$n=1\longrightarrow\begin{bmatrix}1\end{bmatrix}$$

$$n=2\longrightarrow\begin{bmatrix}1&1\\1&\end{bmatrix}$$

$$n=3\longrightarrow\begin{bmatrix}1&2&2\\2&3&\\2&&\end{bmatrix}$$

$$n=4\longrightarrow\begin{bmatrix}1&3&5&5\\3&8&10&\\5&10&&\\5&&&\end{bmatrix}$$

$$n=5\longrightarrow\begin{bmatrix}1&4&9&14&14\\4&15&30&35&\\9&30&45&&\\14&35&&&\\14&&&&\end{bmatrix}$$

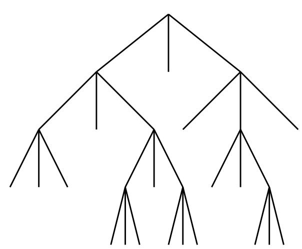
It is clear that $B_3(n,k,0)=B_3(n,0,k)=B(n,k)$. The reader may easily check that when we compute $\sum_{k,l}B_3(n,k,l)$, we obtain: 1, 3, 12, 55, 273. These integers are the first terms of the following sequence (*cf.* [**20**]):

$$C_3(n)=\frac{1}{2n+1}\binom{3n}{n}.$$

2.3. Combinatorial interpretation. Fuss[1]-Catalan numbers (*cf.* [**15**]) are given by the formula

$$C_p(n)=\frac{1}{(p-1)n+1}\binom{pn}{n}, \tag{2.3}$$

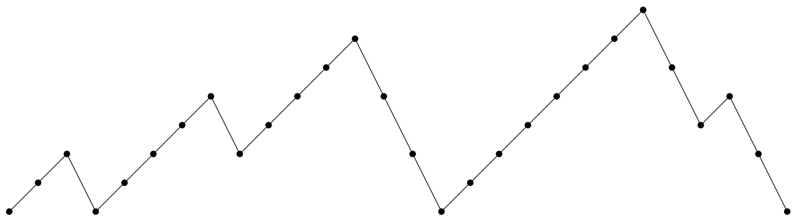
and $C_3(n)$ appear as order-3 Fuss-Catalan numbers. The integers $C_3(n)$ are known [**20**] to count *ternary trees*, *i.e.* trees in which every internal node has exactly 3 sons.



Ternary trees are in bijection with *2-Dyck paths*, which are defined as paths from $(0,0)$ to $(3n,0)$ with steps $(1,1)$ and $(1,-2)$, and remaining above the line $y=0$. The bijection between these objects is the same as in the case of binary trees, *i.e.* a depth-first search, with the difference that here an internal node

[1]Nikolai Fuss (Basel, 1755 – St Petersburg, 1826) helped Euler prepare over 250 articles for publication over a period on about seven years in which he acted as Euler's assistant, and was from 1800 to 1826 permanent secretary to the St Petersburg Academy.

is translated into a $(1,-2)$ step. To illustrate this bijection, we give the path corresponding to the previous example of ternary tree:



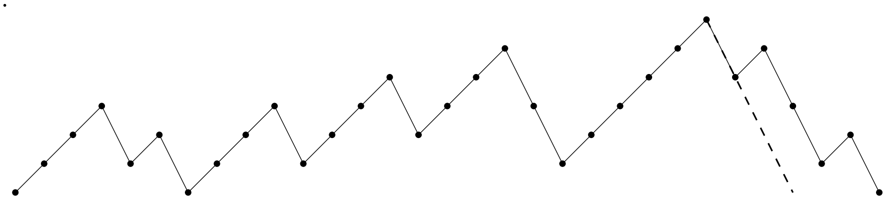
We shall consider these paths with respect to the position of their down steps. Let $\mathcal{D}_{n,k,l}$ denote the set of 2-Dyck paths of length $3n$, with k down steps at even height and l down steps at odd height. By convention, the last sequence of down steps is not considered (the number of these steps is by definition equal to $n-k-l$). In the previous example, $n=9$, $k=5$ and $l=2$.

Proposition 2.2. *We have*

$$\sum_{k,l} B_3(n,k,l) = C_3(n) = \frac{1}{2n+1}\binom{3n}{n}.$$

Moreover, $B_3(n,k,l)$ is the cardinality of $\mathcal{D}_{n,k,l}$.

Proof. Let k and l be fixed. Let us consider an element of $\mathcal{D}_{n,k,l}$. If we cut this path after its $(2n-2)$-th up step, and complete with down steps, we obtain a 2-Dyck path of length $3(n-1)$ (see figure below). It is clear that this path is an element of $\mathcal{D}_{n,i,j}$ for some $i \leq k$ and $j \leq l$. We can furthermore reconstruct the original path from the truncated one, if we know k and l. We only have to delete the last sequence of down steps (here the dashed line), to draw $k-i$ down steps, one up step, $l-j$ down steps, one up step, and to complete with down steps. This gives a bijection from $\mathcal{D}_{n,k,l}$ to $\cup_{0\leq i\leq k, 0\leq j\leq l}\mathcal{D}_{n-1,i,j}$, which implies Proposition 2.2.



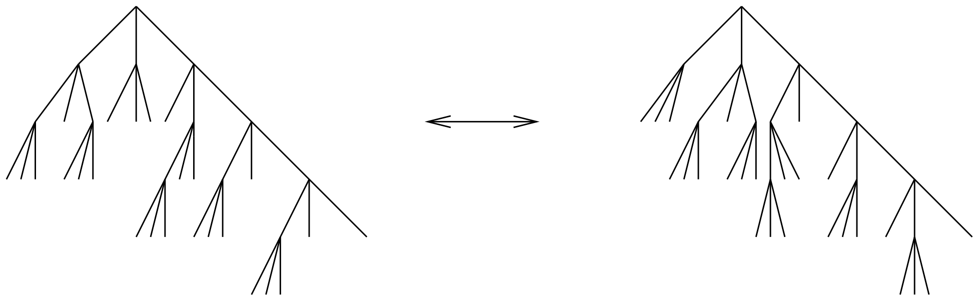
$\square$

Remark 2.1. It is interesting to translate the bi-statistics introduced on 2-Dyck paths to the case of ternary trees. As previously, we consider the depth-first search of the tree, and shall not consider the last right string. We define $\mathcal{T}_{n,k,l}$ as the set of ternary trees with n internal nodes, k of them being encountered in the search after an even number of leaves and l after and odd number of leaves. By the bijection between trees and paths, and Proposition 2.2, we have that the cardinality of $T_{n,k,l}$ is $B_3(n,k,l)$.

Remark 2.2. It is clear from the definition that:

$$B_3(n,k,l) = B_3(n,l,k).$$

But this fact is not obvious when considering trees or paths, since the statistics defined are not clearly symmetric. To explain this, we can introduce an involution on the set of ternary trees which sends an element of $\mathcal{T}_{n,k,l}$ to $\mathcal{T}_{n,l,k}$. To do this, we can exchange for each node of the last right string its left and its middle son, as in the following picture. Since the number of leaves of a ternary tree is odd, every "even" node becomes an "odd" one, and conversely.



2.4. Explicit formula. Now a natural question is to obtain explicit formulas for the $B_3(n,k,l)$. The answer is given by the following proposition.

Proposition 2.3. *The intergers* $B_3(n,k,l)$ *are given by*

$$B_3(n,k,l) = \binom{n+k-1}{k}\binom{n+l-1}{l}\frac{n-k-l}{n} \tag{2.4}$$

Proof. [sketch] The proof is a variation of the cycle lemma [**10**], used to enumerate $\mathcal{D}_{n,k,l}$. It is also possible, once we have the formula (2.4), to check the recurrence (2.2). □

3. Ideals of B-quasisymmetric functions

3.1. Definitions, notations and results. For these definitions, we follow [**7**], with some minor differences, for the sake of simplicity of the computations we will have to make.

Let $\mathbb{N}$ and $\bar{\mathbb{N}}$ denote two occurrrences of the set of nonnegative integers. We shall write $\bar{\mathbb{N}} = \{\bar{0}, \bar{1}, \bar{2}, \ldots\}$ and make no difference between the elements of $\mathbb{N}$ and $\bar{\mathbb{N}}$ in any arithmetical expression. We distinguish $\mathbb{N}$ and $\bar{\mathbb{N}}$ for the ease of reading.

A *bivector* is a vector $v = (v_1, v_2, \ldots, v_{2k-1}, v_{2k})$ such that the odd entries $\{c_{2i-1},\ i = 1..k\}$ are in $\mathbb{N}$, and the even entries $\{c_{2i},\ i = 1..k\}$ are in $\bar{\mathbb{N}}$.

A *bicomposition* is a bivector in which there is no consecutive zeros, *i.e.* no pattern $0\bar{0}$ or $\bar{0}0$.

The integer k is called the *size* of v. The *weight* of the vector v is by definition the couple $(|v|_{\mathbb{N}}, |v|_{\bar{\mathbb{N}}})$, where $|v|_{\mathbb{N}} = \sum_{i=1}^k v_{2i-1}$ and $|v|_{\bar{\mathbb{N}}} = \sum_{i=1}^k v_{2i}$. We also set $|v| = |v|_{\mathbb{N}} + |v|_{\bar{\mathbb{N}}}$.

For example $(1, \bar{0}, 2, \bar{1}, 0, \bar{2}, 3, \bar{0})$ is a bicomposition of size 4, and of weight $(6, 3)$.

To make notations lighter, we shall sometimes write bivectors or bicomposition as words, for example $1\bar{0}2\bar{1}0\bar{2}3\bar{0}$.

The *fundamental B-quasisymmetric functions*, indexed by bicompositions, are defined as follows

$$F_{c_1c_2\ldots c_{2k-1}c_{2k}}(\mathcal{A}_n) = \sum x_{i_1}\cdots x_{i_{|c|_{\mathbb{N}}}}\, y_{j_1}\cdots y_{j_{|c|_{\bar{\mathbb{N}}}}} \in \mathbb{Q}[\mathcal{A}_n]$$

where the sum is taken over indices i's and j's such that

$$i_1 \le \cdots i_{c_1} \le j_1 \le \cdots j_{c_2} < i_{c_1+1} \le \cdots i_{c_1+c_3} \le j_{c_2+1} \le \cdots \le j_{c_2+c_4} < i_{c_1+c_3+1} \le \cdots$$

We give some examples:
$F_{1\bar{2}} = \sum_{i\le j\le k} x_i y_j y_k$,
$F_{0\bar{2}1\bar{0}} = \sum_{i\le j<k} y_i y_j x_k$.

It is clear from the definition that the bidegree (*i.e.* the couple (degree in x, degree in y)) of F_c in $\mathbb{Q}[\mathcal{A}_n]$ is the weight of c. If the size of c is greater than n, we shall set $F_c(\mathcal{A}_n) = 0$.

The space of B-quasisymmetric functions, denoted by $QSym_n(B)$ is the vector subspace of $\mathbb{Q}[\mathcal{A}_n]$ generated by the $F_c(\mathcal{A}_n)$, for all bicompositions c.

Let us denote by $\mathcal{I}_n^2$ the ideal $\langle QSym_n(B)^+\rangle$ generated by B-quasisymmetric functions with zero constant term.

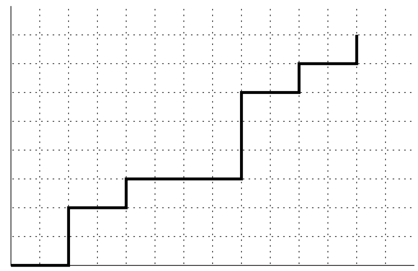
With these notations, our goal is to prove

$$\dim \mathbb{Q}[\mathcal{A}_n]/\langle QSym_n(B)^+\rangle = \frac{1}{2n+1}\binom{3n}{n}. \tag{3.1}$$

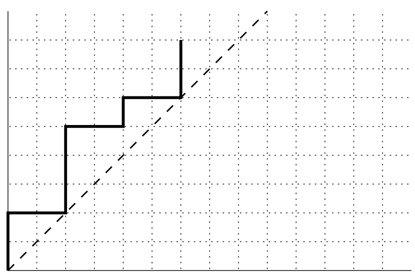
3.2. Paths and $\mathcal{G}$-set. The aim of this subsection is to construct a set $\mathcal{G}$ of polynomials, which will be proved in the next section to be a Gröbner basis of $\mathcal{I}_n^2$. This part of the work is greatly inspired from [**4, 5**].

Let $v = (v_1, v_2, \dots, v_{2k-1}, v_{2k})$ be a bivector of size n. We associate to v a path $\pi(v)$ in the plane $\mathbb{N} \times \mathbb{N}$, with steps (0,1) or (2,0). We start from (0,0) and add for each entry v_i (read from left to right): v_i steps (2,0), followed by one step (0,1).

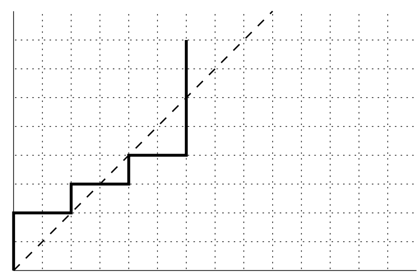
As an example, the path associated to $(1, \bar{0}, 1, \bar{2}, 0, \bar{0}, 1, \bar{1})$ is



We have two kinds of path, regarding their position to the diagonal $x = y$. If a path always remains above this line, we call it a *2-Dyck path*, and say that the corresponding vector is *2-Dyck*. Conversely, if the path enters the region $x < y$, we call both the path and the vector *transdiagonal*. For example, $v = (0, \bar{0}, 1, \bar{0}, 0, \bar{1}, 1, \bar{0})$ is 2-Dyck, whereas $w = (0, \bar{0}, 1, \bar{1}, 1, \bar{0}, 0, \bar{0})$ is transdiagonal.



π (v)



π (w)

A simple but important observation is that a vector $v = (v_1, v_2, \dots, v_{2k-1}, v_{2k})$ is transdiagonal if and only if there exists $1 \leq l \leq k$ such that

$$v_1 + v_2 + \cdots + v_{2l-1} + v_{2l} \geq l. \tag{3.2}$$

Our next task is to construct a set $\mathcal{G}$ of polynomials, mentionned above. From now on, unless otherwise indicated, vectors are of size n. For w a vector of size $k < n$, $w0^*$ denotes the vector (of size n) obtained by adding the desired number of $0\bar{0}$ patterns. We shall define the *length* $\ell(v)$ of a vector v as the integer k such that $v = v_1\, v_2\, \dots\, v_{2k-1}\, v_{2k}\, 0^*$ with $v_{2k-1}\, v_{2k} \neq 0\bar{0}$. In the case of bicompositions, the notions of size and length coincide.

For v a vector (of size n), we denote by $\mathcal{A}_n^v$ the monomial

$$\mathcal{A}_n^v = x_1^{v_1} y_1^{v_2} \cdots x_n^{v_{2n-1}} y_n^{v_{2n}}.$$

To deal with leading terms of polynomials, we will use the lexicographic order induced by the ordering of the variables:

$$x_1 > y_1 > x_2 > y_2 > \cdots > x_n > y_n.$$

The lexicographic order is defined on monomials as follows: $\mathcal{A}_n^v >_{\text{lex}} \mathcal{A}_n^w$ if and only if the first non-zero entry of $v - w$ (componentwise) is positive.

The set

$$\mathcal{G} = \{G_v\} \subset \mathcal{I}_n^2$$

is indexed by transdiagonal vectors. Let v be a transdiagonal vector.

For $v = c0^*$ with c a non-zero bicomposition of length $\geq n$ (which implies that v is transdiagonal), we define

$$G_v = F_c.$$

If v cannot be written as $c0^*$, the polynomial G_v is defined recursively. We look at the rightmost occurrence of two consecutive zeros (on the left of a non-zero entry: we do not consider the subword 0^*). Two cases are to be distinguished according to the parity of the position of this pattern:

- if $v = w0\bar{0}\alpha\beta c0^*$, with w a vector of size $k-1$, $\alpha \in \mathbb{N}$ (by definition non-zero), $\beta \in \bar{\mathbb{N}}$, c a bicomposition, we define

$$G_{w0\bar{0}\alpha\beta c0^*} = G_{w\alpha\beta c0^*} - x_k\, G_{w(\alpha-1)\beta c0^*}; \tag{3.3}$$

- if $v = w\alpha\bar{0}0\beta c0^*$, with w a vector of size $k-1$, $\alpha \in \mathbb{N}$, $\beta \in \bar{\mathbb{N}}$ (by definition non-zero), c a bicomposition, we define

$$G_{w\alpha\bar{0}0\beta c0^*} = G_{w\alpha\beta c0^*} - y_k\, G_{w\alpha(\beta-1)c0^*}. \tag{3.4}$$

We easily check that both terms on the right of (3.3) and (3.4) are indexed by vectors that are transdiagonal as soon as v is transdiagonal. We do it for (3.3) : let us denote $v' = w\alpha\beta c0^*$ and $v'' = w(\alpha-1)\beta c0^*$. Let l be the smallest integer such that (3.2) holds for v. If $l \geq k-1$ then w is transdiagonal thus so are v' and v'', and if not:

$$v'_1 + v'_2 + \cdots + v'_{2l-3} + v'_{2l-2} \geq l \quad \text{and} \quad v''_1 + v''_2 + \cdots + v''_{2l-3} + v''_{2l-2} \geq l-1.$$

Since v' and v'' are of length equal to $\ell(v)-1$, this defines any G_v for v transdiagonal by induction on $\ell(v)$.

It is interesting to develop an example, where we take $n = 3$.

$$\begin{array}{rcl}
G_{0\bar{0}1\bar{0}0\bar{2}} & = & G_{0\bar{0}1\bar{2}0\bar{0}} - y_2\, G_{0\bar{0}1\bar{1}0\bar{0}} \\
& = & (G_{1\bar{2}0\bar{0}0\bar{0}} - x_1\, G_{0\bar{2}0\bar{0}0\bar{0}}) - y_2\,(G_{1\bar{1}0\bar{0}0\bar{0}} - x_1\, G_{0\bar{1}0\bar{0}0\bar{0}}) \\
& = & (F_{1\bar{2}} - x_1\, F_{0\bar{2}}) - y_2(F_{1\bar{1}} - x_1\, F_{0\bar{1}}) \\
& = & (x_1y_1^2 + x_1y_1y_2 + x_1y_1y_3 + x_1y_2^2 + x_1y_2y_3 + x_1y_3^2 + x_2y_2^2 + x_2y_2y_3 \\
& & \quad +x_2y_3^2 + x_3y_3^2 - x_1(y_1^2 + y_1y_2 + y_1y_3 + y_2^2 + y_2y_3 + y_3^2)) \\
& & -y_2(x_1y_1 + x_1y_2 + x_1y_3 + x_2y_2 + x_2y_3 + x_3y_3 - x_1(y_1 + y_2 + y_3)) \\
& = & x_2y_3^2 - y_2x_3y_3 + x_3y_3^2
\end{array}$$

The monomials of the result are ordered with respect to the lexicographic order and we observe that the leading monomial (denoted LM) of $G_{0\bar{0}1\bar{0}0\bar{2}}$ is $\mathcal{A}_3^{0\bar{0}1\bar{0}0\bar{2}}$. The following proposition shows that this fact holds in general for the family $\mathcal{G}$.

Proposition 3.1. *Let v be a transdiagonal vector. The leading monomial of G_v is*

$$LM(G_v) = \mathcal{A}_n^v. \tag{3.5}$$

Proof. [sketch] It is done by induction on the length of v. □

3.3. Proof of the main theorem. The aim of this subsection is to prove Theorem 1.1, by showing that the set $\mathcal{G}$ constructed in the previous section is a Gröbner basis for $\mathcal{I}_n^2$. This will be achieved in several steps.

We introduce the notation $\mathcal{Q}_n = \mathbb{Q}[\mathcal{A}_n]/\mathcal{I}_n^2$ and define

$$\mathcal{B}_n = \{\mathcal{A}_n^v \;/\; \pi(v) \textit{ is a } 2\textit{−Dyck path}\}.$$

Lemma 3.1. *Any polynomial $P \in \mathbb{Q}[\mathcal{A}_n]$ is in the span of $\mathcal{B}_n$ modulo $\mathcal{I}_n^2$. That is*

$$P(\mathcal{A}_n) \equiv \sum_{\mathcal{A}_n^v \in \mathcal{B}_n} c_v \mathcal{A}_n^v. \tag{3.6}$$

Proof. It clearly suffices to show that (3.6) holds for any monomial $\mathcal{A}_n^v$, with v transdiagonal. We assume that there exists $\mathcal{A}_n^v$ not reducible of the form (3.6) and we choose $\mathcal{A}_n^w$ to be the smallest amongst them with respect to the lexicographic order. Let us write

$$\begin{array}{rcl}
\mathcal{A}_n^w & = & LM(G_w) \\
& = & (\mathcal{A}_n^w - G_w) + G_w \\
& \equiv & \mathcal{A}_n^w - G_w \pmod{\mathcal{I}_n^2}.
\end{array}$$

All monomials in $(\mathcal{A}_n^w - G_w)$ are lexicographically smaller than $\mathcal{A}_n^w$, thus they are reducible. This contradicts our assuption and completes the proof. □

This lemma implies that $\mathcal{B}_n$ spans the quotient $\mathcal{Q}_n$. We will now prove its linear independence. The next lemma is a crucial step.

Lemma 3.2. *If we denote by $\mathcal{L}[S]$ the linear span of a set S, then*

$$\mathbb{Q}[\mathcal{A}_n] = \mathcal{L}[\mathcal{A}_n^v F_c \ / \ \mathcal{A}_n^v \in \mathcal{B}_n,\ |c| \geq 0]. \tag{3.7}$$

Proof. We have the following reduction for any monomial $\mathcal{A}_n^w$ in $\mathbb{Q}[\mathcal{A}_n]$:

$$\mathcal{A}_n^w = \sum_{\mathcal{A}_n^v \in \mathcal{B}_n} c_v \mathcal{A}_n^v + \sum_{|c|>0} Q_c F_c, \quad Q_c \in \mathbb{Q}[\mathcal{A}_n]. \tag{3.8}$$

We then apply the reduction (3.6) to each monomial of the Q_c's. Now we use the algebra structure of QSym(B) (*cf.* Proposition 37 of [**7**]) to reduce products of fundamental B-quasisymmetric functions as linear combinations of F_c's. We obtain (3.7) in a finite number of operations since degrees strictly decrease at each operation, because $|c| > 0$ implies $\deg Q_c < |w|$. □

Now we come to the final step in the proof. Before stating this lemma, we introduce some notation, and make an observation. For $v = (v_1, v_2, v \ldots, v_{2k-1}, v_{2k})$ a bivector, let $r(v)$ denote the reverse bivector: $r(v) = (v_{2k}, v_{2k-1}, \ldots, v_2, v_1)$. In the same way, let $R(\mathcal{A}_n)$ denote the reverse alphabet of $\mathcal{A}_n$: $R(\mathcal{A}_n) = y_n, x_n, \ldots, y_1, x_1$. Then one has for any bicomposition c:

$$F_c(R(\mathcal{A}_n)) = F_{r(c)}(\mathcal{A}_n). \tag{3.9}$$

Lemma 3.3. *The set $\mathcal{G}$ is a linear basis of $\mathcal{I}_n^2$, i.e.*

$$\mathcal{I}_n^2 = \mathcal{L}[G_w \ / \ w \ \textit{transdiagonal}]. \tag{3.10}$$

Proof. [sketch] We use Lemma 3.2, observation (3.9), and the algebra structure of $QSym_n(B)$ to write:

$$\begin{aligned} \mathcal{I}_n^2 &= \langle F_c,\ |c| > 0 \rangle_{\mathbb{Q}[\mathcal{A}_n]} = \mathcal{L}[\mathcal{A}_n^v \, F_c \, F_{c'} \ / \ R(\mathcal{A}_n)^v \in \mathcal{B}_n,\ |c| > 0,\ |c'| \geq 0] \\ &= \mathcal{L}[\mathcal{A}_n^v \, F_{c''} / \ R(\mathcal{A}_n)^v \in \mathcal{B}_n,\ |c''| > 0]. \end{aligned}$$

Then we prove that we can reduce any term $\mathcal{A}_n^v \, F_{c''}$ using the G polynomials, and we illustrate this on an example, where $n = 5$:

$$\begin{aligned} x_1 \, y_2 \, F_{1\bar{0}0\bar{1}} &= y_2(x_1 \, F_{1\bar{0}0\bar{1}}) \\ &= y_2(G_{2\bar{0}0\bar{1}0\bar{0}0\bar{0}0\bar{0}} - G_{0\bar{0}2\bar{0}0\bar{1}0\bar{0}0\bar{0}}) \\ &= y_2 \, G_{2\bar{0}0\bar{1}0\bar{0}0\bar{0}0\bar{0}} - y_2 \, G_{0\bar{0}2\bar{0}0\bar{1}0\bar{0}0\bar{0}} \\ &= G_{2\bar{0}0\bar{2}0\bar{0}0\bar{0}0\bar{0}} - G_{2\bar{0}0\bar{0}0\bar{2}} - G_{0\bar{0}2\bar{1}0\bar{1}0\bar{0}0\bar{0}} + G_{0\bar{0}2\bar{0}0\bar{1}0\bar{1}0\bar{0}}. \end{aligned}$$

□

Now we are able to complete the proof of Theorem 1.1. We can even state a more precise result.

Theorem 3.4. *A basis of the quotient $\mathcal{Q}_n$ is given by the set*

$$\mathcal{B}_n = \{\mathcal{A}_n^v \ / \ \pi(v) \ \textit{is a } 2\textit{−Dyck path}\},$$

which implies

$$\dim \mathcal{Q}_n = \frac{1}{2n+1}\binom{3n}{n}. \tag{3.11}$$

Since $\mathcal{I}_n^2$ is bihomogeneous, the quotient $\mathcal{Q}_n$ is bigraded and we can consider $\mathbf{H}_{k,l}(\mathcal{Q}_n)$ the subspace of $\mathcal{Q}_n$ consisting of polynomials of bidegree (k, l), then

$$\dim \mathbf{H}_{k,l}(\mathcal{Q}_n) = \binom{n+k-1}{k}\binom{n+l-1}{l}\frac{n-k-l}{n}. \tag{3.12}$$

PROOF. By Lemma 3.1, the set $\mathcal{B}_n$ spans $\mathcal{Q}_n$. Assume we have a linear dependence:

$$P = \sum_{\mathcal{A}_n^v \in \mathcal{B}_n} a_v \, \mathcal{A}_n^v \;\in\; \mathcal{I}_n^2.$$

By Lemma 3.3, the set $\mathcal{G}$ spans $\mathcal{I}_n^2$, thus

$$P = \sum_{u \text{ transdiagonal}} b_u \, G_u.$$

This implies $LM(P) = \mathcal{A}_n^u$, with u transdiagonal, which is absurd. Hence $\mathcal{B}_n$ is a basis of the quotient $\mathcal{Q}_n$.

The expressions (3.11) and (3.12) are consequences of Section 2's results.

□

Remark. This work admits direct generalization. We can define quasisymmetric polynomials in p sets of variables. In this case, the quotient of the polynomial ring by the ideal generated by p-quasisymmetric polynomials (without constant term) has dimension given by $\frac{1}{pn+1}\binom{(p+1)n}{n}$. These numbers are Fuss-Catalan numbers, which enumerate $(p+1)$-ary trees. The combinatorial part corresponds to let the "Catalan recurrence" grow in $(p+1)$ dimensions, and we obtain multivariate Fuss-Catalan numbers of order $(p+1)$. All details can be found in [**2, 3**].

Acknowledgement. The author thanks C. Hohlweg for introducing him to B-quasisymmetric polynomials, and for valuable comments and explanations.

References

[1] E. Artin, Galois Theory, Notre Dame Mathematical Lecture **2** (1944), Notre Dame, IN.

[2] J.-C. Aval, Multivariate Fuss-Catalan numbers, preprint.

[3] J.-C. Aval, Ideals and quotients of B-quasisymmetric functions, preprint.

[4] J.-C. Aval and N. Bergeron, Catalan Paths and Quasi-Symmetric Functions, Proc. Amer. Math. Soc. **131** (2003), 1053-1062.

[5] J.-C. Aval, F. Bergeron and N. Bergeron, Ideals of Quasi-Symmetric Functions and Super-Coinvariant Polynomials for S_n, Adv. in Math., **181** (2004), no. 2, 353–367..

[6] J.-C. Aval, F. Bergeron and N. Bergeron, Diagonal Temperley-Lieb invariants and harmonics, preprint.

[7] P. Baumann and C. Hohlweg, A Solomon theory for the wreath product $G \wr \mathfrak{S}_n$, preprint.

[8] F. Bergeron, N. Bergeron, A. Garsia, M. Haiman and G. Tesler, Lattice Diagram Polynomials and Extended Pieri Rules, Adv. Math., **142** (1999), 244-334.

[9] N. Bergeron, S. Mykytiuk, F. Sottile, and S. van Willigenburg, Pieri Operations on Posets, J. of Comb. Theory, Series A, **91** (2000), 84–110.

[10] N. Dershowitz and S. Zaks, The Cycle Lemma and Some Applications, Europ. J. Combinatorics, **11** (1990), 35–40.

[11] A. M. Garsia and M. Haiman, A graded representation model for Macdonald's polynomials, Proc. Nat. Acad. Sci. U.S.A. **90** (1993), no. 8, 3607–3610.

[12] I. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. Retakh, and J.-Y. Thibon, Noncommutative symmetric functions, Adv. in Math., **112** (1995), 218–348.

[13] I. Gessel, Multipartite P-partitions and products of skew Schur functions, in Combinatorics and Algebra (Boulder, Colo., 1983), C. Greene, ed., vol. 34 of Contemp. Math., AMS, 1984, pp. 289–317.

[14] M. Haiman, Vanishing theorems and character formulas for the Hilbert scheme of points in the plane, Invent. Math. **149**, no. 2 (2002), 371–407.

[15] P. Hilton and J. Pedersen, Catalan Numbers, their Generalizations, and their Uses, Math. Intelligencer, **13** (1001), 64–75.

[16] F. Hivert, Hecke algebras, difference operators, and quasi-symmetric functions, Adv. in Math., **155** (2000), 181–238.

[17] C. Malvenuto and C. Reutenauer, Duality between quasi-symmetric functions and the Solomon descent algebra, Journal of Algebra, **177** (1995), 967–982.

[18] J.-C. Novelli and J.-Y. Thibon, Free quasi-symmetric functions of arbitrary level, preprint.

[19] S. Poirier, Cycle type and descent set in wreath products, Proceedings of the 7th Conference on Formal Power Series and Algebraic Combinatorics (Noisy-le-Grand, 1995). Discrete Math. **180** (1998), no. 1-3, 315–343.

[20] N. J. A. Sloane, A Handbook of Integer Sequences, New York and London: Academic Press (1973).

[21] R. Stanley, Enumerative Combinatorics, Vol. 1, Wadsworth and Brooks/Cole, 1986.

[22] R. Stanley, Enumerative Combinatorics Vol. 2, no. 62 in Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1999. Appendix 1 by Sergey Fomin.

[23] R. Stanley, Catalan Addendum, (web) `http://www-math.mit.edu/ rstan/ec/catadd.pdf`.

LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence cedex, FRANCE
E-mail address: `aval@labri.fr`
URL: `http://www.labri.fr/perso/aval`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Bijective counting of Kreweras walks and loopless triangulations

Olivier Bernardi

Abstract. We consider lattice walks in the plane starting at the origin, remaining in the first quadrant $i, j \geq 0$ and made of West, South and North-East steps. There are nice formulas for the number of such walks. But, although several proofs of these formulas have been proposed over the years, none of them provides a combinatorial explanation. We give such an explanation. Beside these walks, we enumerate loopless triangulations of the sphere bijectively. Our proofs rely on bijections between walks and triangulations with a distinguished spanning tree. As a by-product, we also enumerate an important class of spanning trees on cubic maps.

Résumé. On considère les chemins planaires partant de l'origine, restant dans le quart de plan $i, j \geq 0$ et faits de pas Ouest, Sud et Nord-Est. Il existe de jolies formules énumératives pour ces chemins. Mais, alors que plusieurs démonstrations ont été proposées pour ces formules par le passé, aucune ne fournit d'explication combinatoire. Nous donnons une telle explication. En sus de ces chemins, nous énumérons bijectivement les triangulations sans boucle de la sphère. Nos preuves reposent sur des bijections entre des chemins et des triangulations munies d'un arbre couvrant. Dans le même temps, nous énumérons une famille importante d'arbres couvrants sur les cartes cubiques.

1. Introduction

We consider lattice walks in the plane starting from the origin (0,0), remaining in the first quadrant $i, j \geq 0$ and made of three kind of steps: West, South and North-East. These walks were first studied by Germain Kreweras [**4**] and inherited his name. A Kreweras walk ending at the origin is represented in Figure 1.

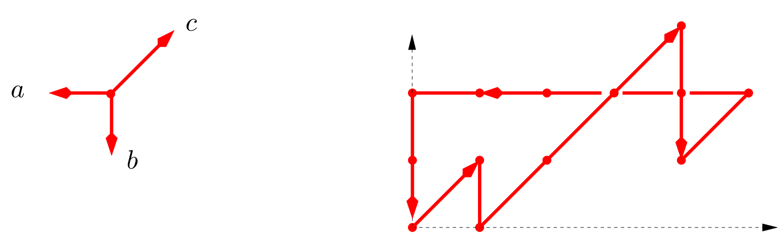


FIGURE 1. The Kreweras walk *cbcccbbcaaaaabb*.

These walks have remarkable enumerative properties. Kreweras proved in 1965 that the number of walks of length $3n$ ending at the origin is:

$$k_n = \frac{4^n}{(n+1)(2n+1)}\binom{3n}{n}. \tag{1.1}$$

The original proof of this result is complicated and somewhat unsatisfactory. It was performed by *guessing* the number of walks of size n ending at point (i, j). The conjectured formulas were then checked using the

2000 *Mathematics Subject Classification.* Primary 05A15.

Key words and phrases. planar walk, Kreweras, planar map, triangulation, bijection, counting.

recurrence relations between these numbers. The *checking part* involved several hypergeometric identities which were later simplified by Niederhausen [**6**]. In 1986, Gessel gave a different proof in which the *guessing part* was reduced [**3**]. More recently, Bousquet-Mélou proposed a constructive proof (that is, without guessing) of these results and some extensions [**1**]. Still, the simple looking formula (1.1) remained without a direct combinatorial explanation. The problem of finding a combinatorial explanation was raised by Stanley in [**9**]. One of our goals in this paper is to provide such an explanation.

Formula (1.1) for the number of Kreweras walks is to be compared to another formula proved the same year. In 1965, Mullin, following the seminal steps of Tutte, proved via a generating function approach [**5**] that the number of loopless triangulations of size n (see below for precise definitions) is

$$t_n \quad = \quad \frac{2^n}{(n+1)(2n+1)}\binom{3n}{n}. \tag{1.2}$$

A bijective proof of Formula (1.2) was outlined by Schaeffer in his Ph.D thesis [**7**]. See also [**8**] for a more general construction concerning non-separable triangulations of a k-gon. We will give an alternative bijective proof for the number of loopless triangulations. Technically speaking, we will work instead on *cubic maps without isthmus* which are the dual of loopless triangulations.

2. How the proofs work

We begin with an account of this paper's content in order to underline the (slightly unusual) logic structure of our proofs.

• In Section 3, we recall some definitions on rooted planar maps. Then, we define a special class of spanning trees called *depth trees*. Depth trees are closely related to the trees that can be obtained by a *depth first search* algorithm.

• In Section 4, we describe a bijection Φ between Kreweras walks ending at the origin and cubic maps without isthmus covered by a depth tree. As an immediate enumerative corollary, we obtain the relation

$$k_n = d_n,$$

between the number k_n of Kreweras walks of size n ending at the origin and the number d_n of cubic maps without isthmus of size n covered by a depth tree.

• In Section 5, we extend the mapping Φ to a larger class of walks called *extended Kreweras walks*. These walks (made of West, South and North-East steps) start from the origin (0,0) and remain in the half-plane $i + j \geq 0$. An extended Kreweras walk ending on the second diagonal (i.e. the line $i + j = 0$) is represented in Figure 2.

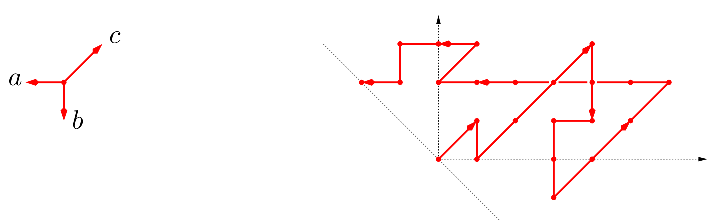


FIGURE 2. An extended Kreweras walk ending on the second diagonal

Unlike the Kreweras walks, the extended Kreweras walks are easy to count. A simple application of the cycle lemma (see Section 5.3 of [**10**]) allows one to prove that the number of extended Kreweras walks of length $3n$ ending on the second diagonal is

$$e_n = \frac{4^n}{2n+1}\binom{3n}{n}.$$

Then, we prove that the mapping Φ can be generalized into a bijection between extended Kreweras walks ending on the second diagonal and cubic maps without isthmus of size n covered by a depth tree with a marked edge not in the tree. Since any cubic map of size n has exactly $n+1$ edges not in the spanning tree, we obtain

$$e_n = (n+1)d_n.$$

As a result, we get

$$d_n \; = \; k_n \; = \; \frac{4^n}{(n+1)(2n+1)}\binom{3n}{n},$$

and recover Equation (1.1).

• In Section 6, we enumerate depth trees on cubic maps. We prove that the number of such trees for a cubic map of size n is 2^n. This result implies that the number of cubic maps of size n is

$$c_n \; = \; \frac{d_n}{2^n} \; = \; \frac{2^n}{(n+1)(2n+1)}\binom{3n}{n}.$$

Thus, we obtain a combinatorial proof of Formula (1.2).

• In Section 7, we extend the mapping Φ to Kreweras walks ending at $(i,0)$ and discuss some open problems.

3. Definitions and notations

3.1. Kreweras walks. In the following, Kreweras walks are considered as words on the alphabet $\{a,b,c\}$. The letter a (resp. b, c) corresponds to a West (resp. South, North-East) step. For instance, the walk in Figure 1 is *cbcccbbcaaaaabb*. The length of a word w is denoted by $|w|$ and the number of occurrences of a given letter α is denoted by $|w|_\alpha$. Kreweras walks are the words w on the alphabet $\{a,b,c\}$ such that any prefix w' of w satisfies

$$|w'|_a \leq |w'|_c \quad \text{and} \quad |w'|_b \leq |w'|_c \; . \tag{3.1}$$

Kreweras walks ending at the origin satisfy the additional constraint

$$|w|_a \; = \; |w|_b \; = \; |w|_c. \tag{3.2}$$

These conditions can be interpreted as a ballot problem with three candidates. This is why Kreweras walks sometimes appear under this formulation in the literature [**6**].

Similarly, the extended Kreweras walks (i.e. the walks remaining in the half-plane $i+j \geq 0$) are the words w on $\{a,b,c\}$ such that any prefix w' of w satisfies

$$|w'|_a \; + \; |w'|_b \; \leq \; 2|w'|_c \; , \tag{3.3}$$

and walks ending on the second diagonal satisfy the additional constraint

$$|w|_a \; + \; |w|_b \; = \; 2|w|_c. \tag{3.4}$$

Note that the length of any walk ending on the second diagonal is a multiple of 3. The *size* of such a walk of length $3n$ is n. Note also that a walk ending at point $(i,0)$ has a length of the form $l = 3n+2i$ where n is a non-negative integer. A Kreweras walk of length $l = 3n+2i$ ending at $(i,0)$ has *size* n.

3.2. Cubic maps. We recall some definitions about planar maps. A *planar map*, or *map* for short, is an embedding of a connected planar graph in the sphere without intersecting edges, defined up to orientation preserving homeomorphisms of the sphere. Loops and multiple edges are allowed. The *faces* are the connected components of the complement of the graph. Each edge has two *half-edges*, each incident to one of the endpoints. A map is *rooted* if one of its half-edges is distinguished as the *root*. The endpoint of the root is the *root-vertex*. Graphically, the root is indicated by an arrow pointing on the root-vertex (see Figure 3). All the maps considered in this paper are rooted and we shall not further precise it.

Our constructions lead us to consider *legs*, that is, half-edges that are not part of a complete edge. A *growing map* is a map together with some legs, one of them being distinguished as the *head*. We require the

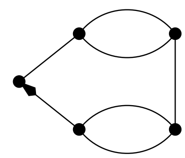


Figure 3. A rooted map.

legs to be all in the same face called *head-face*. The endpoint of the head is the *head-vertex*. Graphically, the head is indicated by an arrow pointing away from the *head-vertex*. The root of a growing map can be a leg or a regular half-edge. For instance, the growing map in Figure 4 has 2 legs beside the head, and its root is not a leg.

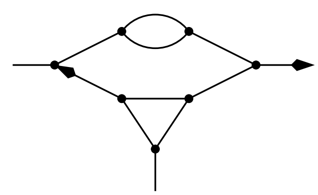


Figure 4. A growing map.

A map (or growing map) is *cubic* if any vertex has degree 3. It is k-near-cubic if the root-vertex has degree k and any other vertex has degree 3. For instance, the map in Figure 3 is 2-near-cubic and the growing map in Figure 4 is cubic. Observe that cubic maps are in bijection with 2-near-cubic maps not reduced to a loop by the mapping illustrated in Figure 5.

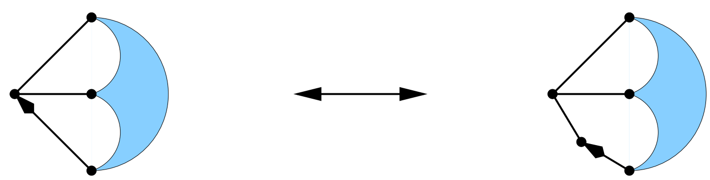


Figure 5. Bijection between cubic maps and 2-near-cubic maps.

We will be interested in *non-separable* k-near-cubic maps. A map is *separable* if the edge set can be partitioned into two non-empty parts such that exactly one vertex is incident to some edges in both parts. It is *non-separable* otherwise. In particular, a non-separable map has no loop nor isthmus (i.e. edge whose deletion disconnect the map). For cubic maps and 2-near-cubic maps it is equivalent to be non-separable or without isthmus.
The incidence relation between vertices and edges in cubic maps shows that the number of edges is always a multiple of 3. More generally, if M is a k-near-cubic map with e edges and v vertices, the incidence relation reads: $3(v-1)+k=2e$. Equivalently, $3(v-k+1)=2(e-2k+3)$. It can be shown that $v-k+1$ is non-negative. Hence, the number of edges has the form $e=3n+2k-3$ where n is a non-negative integer. We say that a k-near-cubic map has *size* n if it has $e=3n+2k-3$ edges (and $v=2n+k-1$ vertices). In particular, the mapping of Figure 5 is a bijection between cubic maps of size n ($3n+3$ edges) and 2-near-cubic maps of size $n+1$ ($3n+4$ edges).
The cubic maps without isthmus form an important class of maps because their *duals* are the loopless triangulations. Recall that the dual M^* of a map M is the map obtained by putting a vertex of M^* in each face of M and an edge of M^* across each edge of M.

3.3. Depth trees. A *tree* is a connected graph without cycle. A subgraph T of a connected graph G is a *spanning tree* if it is a tree containing every vertex of G. An edge of G is said to be *internal* if it is in the spanning tree T and *external* otherwise. For any pair of vertices u, v in G, there is a unique path between u and v in the spanning tree T. We call it the T*-path* between u and v.
A map (or growing map) M with a distinguished spanning tree T will be denoted by M_T. Graphically, we shall indicate the spanning tree by thick lines as in Figure 6. A vertex u of M_T is an *ancestor* of another vertex v if it is on the T-path between the root-vertex and v. In this case, v is a *descendant* of u. Two vertices are *comparable* if one is the ancestor of the other. For instance, in Figure 6, the vertices u_1 and v_1 are comparable whereas u_2 and v_2 are not. A *depth tree* is a spanning tree such that any external edge joins comparable vertices. Moreover, we require the edge containing the root to be external. In Figure 6, the tree on the left side is a depth tree but the tree on the right side is not a depth tree since the edge (u_2, v_2) breaks the rule. Finally, a *depth-map* is a map with a distinguished depth tree.

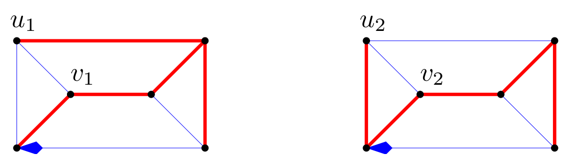


FIGURE 6. A depth tree (left) and a non-depth tree (right).

4. A bijection between Kreweras walks and cubic depth-maps

We define a bijection Φ between Kreweras walks ending at the origin and 2-near-cubic depth-maps (i.e. 2-near-cubic maps with a distinguished depth tree) without isthmus. The general principle of this bijection is to read the walk *from right to left* and interpret each letter as an operation for constructing the map and the tree. We illustrated this step-by-step construction in Figure 8. The intermediary steps are *tree-growing maps*, that is, growing maps together with a distinguished depth tree (indicated by thick lines).

• We start with the tree-growing map M_T^0 consisting of one vertex and two legs. One of the legs is the root, the other is the head (see Figure 7). The spanning tree is reduced to the vertex which is both the root-vertex and the head-vertex.

• We apply successively certain elementary mappings φ_a, φ_b, φ_c (Definition 4.1) corresponding to the letters a, b, c of the Kreweras walk read from right to left.

• When the whole walk is read, we *close* the tree-growing map, that is, we glue the head and the root together as was done in Figure 9.

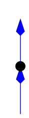


FIGURE 7. The tree-growing map M_T^0.

Let us enter in the details and define the bijection Φ. Consider a growing map M. We *make a tour* of the head-face if we follow its border in counterclockwise direction (i.e. the border of the head-face stays on our left-hand side) starting from the head (see Figure 10). This journey induces a linear order on the legs of M. We shall talk about the *first* and *last* legs of M. Moreover, if the root is a leg, we call *left* (resp. *right*) the legs encountered before (resp. after) the root during the tour of the head-face. For instance, the growing map of Figure 10 has one left leg and two right legs.

We define three mappings φ_a, φ_b, φ_c on tree-growing maps.

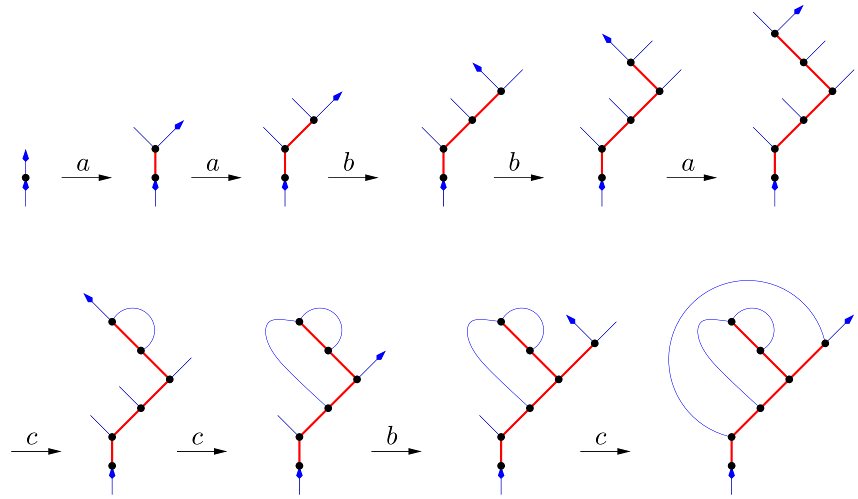


FIGURE 8. Successive applications of the mappings φ_a, φ_b, φ_c for the walk *cbccabbaa*.

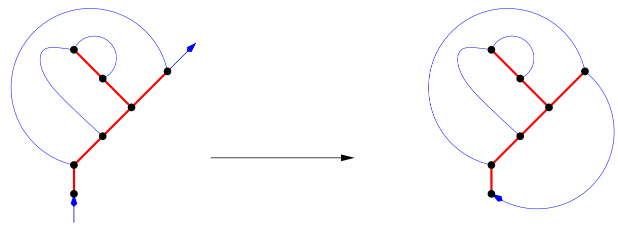


FIGURE 9. Closing the map.

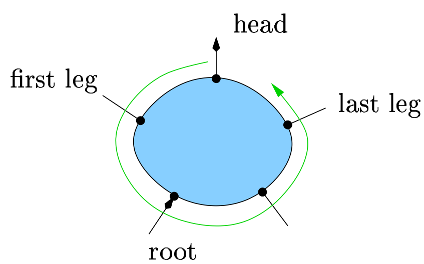


FIGURE 10. Making the tour of the head-face.

DEFINITION 4.1. Let M_T be a tree-growing map (the map is M and the distinguished tree is T).

• The mappings φ_a and φ_b are represented in Figure 11. The tree-growing map $M'_{T'} = \varphi_a(M_T)$ (resp. $\varphi_b(M_T)$) is obtained from M_T by replacing the head by an edge e together with a new vertex v incident with the new head and another leg at its left (resp. right). The tree T' is obtained from T by adding the edge e and the vertex v.

• The tree-growing map $\varphi_c(M_T)$ is only defined if the first and last legs exist (that is, if the head-face contains some legs beside the head) and have distinct and comparable endpoints. We call these legs s and t *with the convention that the endpoint of* s *is an ancestor of the endpoint of* t.
In this case, the tree-growing map $M'_T = \varphi_c(M_T)$ is obtained from M_T by gluing together the head and the leg s while the leg t becomes the new head (see Figure 12). The spanning tree T is unchanged.

• For a word $w = a_1 a_2 \dots a_n$ on the alphabet $\{a, b, c\}$, we denote by φ_w the mapping $\varphi_{a_1} \circ \varphi_{a_2} \circ \cdots \circ \varphi_{a_n}$.

FIGURE 11. The mappings φ_a and φ_b.

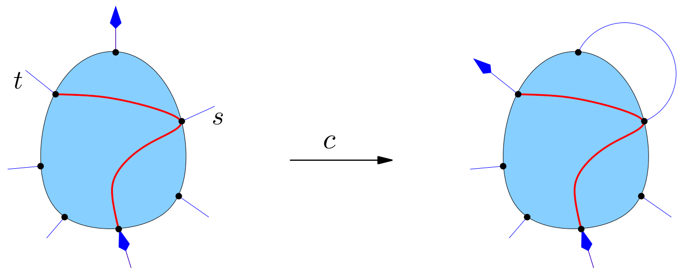


FIGURE 12. The mapping φ_c.

We are now ready to define the mapping Φ on Kreweras walks ending at the origin.

DEFINITION 4.2. Let w be a Kreweras walk ending at the origin. The image of w by the mapping Φ is the map with a distinguished spanning tree obtained by gluing together the root and the head of the tree-growing map $\varphi_w(M_T^0)$.

The mapping Φ has been applied to the walk *cbccabbaa* in Figure 8 and 9. One has to prove that this mapping is well defined. We omit the proof in this extended abstract. However, we highlight one of the key properties: for any suffix w' of w, the tree-growing map $\varphi_{w'}(M_T^0)$ has $|w'|_a - |w'|_c$ left legs and $|w'|_b - |w'|_c$ right legs. (These quantities are non-negative by Equations (3.1) and (3.2).)

We now state the main result of this section.

THEOREM 4.3. *The mapping Φ is a bijection between Kreweras walks of size n (length $3n$) ending at the origin and 2-near-cubic depth-maps without isthmus of size n ($3n+1$ edges).*

COROLLARY 4.4. *The number k_n of Kreweras walks of size n is equal to the number d_n of 2-near-cubic depth-maps without isthmus of size n.*

Observe that d_n is also the number of cubic depth-maps of size $n-1$ without isthmus since the bijection between cubic maps and 2-near-cubic maps represented in Figure 5 can be trivially turned into a bijection between cubic depth-maps and 2-near-cubic depth-maps.

We omit the proof of Theorem 4.3. The general idea is to define the inverse mapping Ψ. This mapping destructs the tree-growing map that Φ constructs and recover the walk. Looking at Figure 8 from bottom-to-top and right-to-left we see how Ψ works.

5. Enumeration of Kreweras walks

Recall that *extended Kreweras walks* are the walks starting from the origin and remaining in the half-plane $i + j \geq 0$. An extended Kreweras walk ending on the *second diagonal* (i.e. the line $i + j = 0$) is represented in Figure 2. The counting of extended Kreweras walks reduces to finding the number of 1-dimensional walks with steps $+2$, and -1 starting at 0 and remaining non-negative. This number is easily found by applying the cycle lemma (see Section 5.3 of [**10**]). We obtain the following result:

PROPOSITION 5.1. *There are*

$$e_n = \frac{4^n}{2n+1}\binom{3n}{n} \tag{5.1}$$

extended Kreweras walks of size n ending on the second diagonal.

We now extend the mapping Φ (Definition 4.2) into an injective mapping Φ' on extended Kreweras walks ending on the second diagonal. The mapping Φ' returns a map with a distinguished spanning tree and a marked external edge. In what follows, a map with a distinguished spanning tree is said *marked* if an external edge is marked.

DEFINITION 5.2. Let w be an extended Kreweras walk ending on the second diagonal. The image of w by the mapping Φ' is the map with a distinguished spanning tree obtained from the tree-growing map $\varphi_w(M_T^0)$ by gluing together the head and the unique remaining leg. The (external) edge obtained by gluing these legs is marked.

We applied the mapping Φ' to the extended Kreweras walks *cabccaaaa* in Figure 13. The marked edge is dashed.

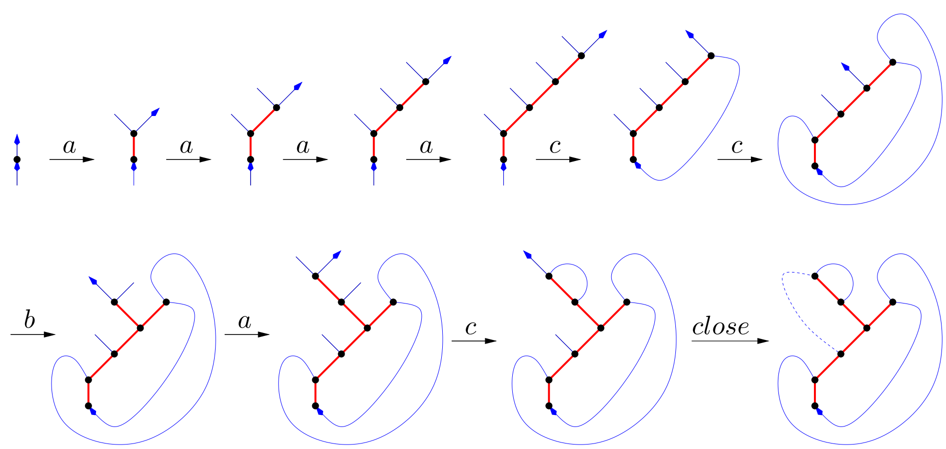


FIGURE 13. The bijection Φ' on the walk *cabccaaaa*.

Observe that the mappings Φ and Φ' coincide on Kreweras walks ending at the origin. In this case, the marked edge is the edge containing the root.

We now state the main result of this section.

THEOREM 5.3. *The mapping Φ' is a bijection between extended Kreweras walks of size n ending on the second diagonal and marked 2-near-cubic depth-maps of size n without isthmus.*

We will not prove Theorem 5.3 but we do explore its consequences. We know from Corollary 4.4 that the number d_n of 2-near-cubic depth-maps without isthmus of size n is equal to the number k_n of Kreweras walks of size n ending at the origin. Consider a 2-near-cubic map M of size n ($3n+1$ edges, $2n+1$ vertices). Since a spanning tree T has $2n$ edges, there are $n+1$ external edges. Therefore, there are $(n+1)d_n = (n+1)k_n$ marked 2-near-cubic depth-map without isthmus. By Theorem 5.3, this is also the number e_n of extended Kreweras walks of size n ending on the second diagonal. We know the number e_n explicitly by Proposition 5.1. Hence, we obtain $(n+1)k_n = e_n = \frac{4^n}{2n+1}\binom{3n}{n}$. This result deserves to be stated as a theorem.

THEOREM 5.4. *There are* $k_n = \dfrac{4^n}{(n+1)(2n+1)}\dbinom{3n}{n}$ *Kreweras walks of size n (length $3n$) ending at the origin.*

6. Enumerating depth trees and cubic maps

In the previous section, we exhibited a bijection Φ' between extended Kreweras walks ending on the second diagonal and marked 2-near-cubic depth-maps without isthmus. As a corollary we obtained the number of 2-near-cubic depth-maps without isthmus of size n: $d_n = \frac{4^n}{(n+1)(2n+1)}\binom{3n}{n}$. In this section, we prove that any 2-near-cubic map of size n has 2^n depth trees (Corollary 6.5). This implies that the number of 2-near-cubic maps of size n without isthmus is $c_n = \frac{d_n}{2^n} = \frac{2^n}{(n+1)(2n+1)}\binom{3n}{n}$. Given the bijection between 2-near-cubic maps and cubic maps (see Figure 5), we obtain the following theorem.

THEOREM 6.1. *There are* $c_n = \dfrac{2^n}{(n+1)(2n+1)}\dbinom{3n}{n}$ *cubic maps without isthmus having* $3n$ *edges.*

By duality, c_n is also the number of loopless triangulations with $3n$ edges. Hence, we recover Equation (1.2) announced in the introduction.

The rest of this section is devoted to the counting of depth trees on cubic maps and, more generally, on cubic (potentially non-planar) graphs. We first give an alternative characterization of depth trees. This characterization is based on the depth-first search (DFS) algorithm (see Section 23.3 of **[2]**). We consider the DFS algorithm as an algorithm for constructing a spanning tree T of a graph.

We consider a graph G with a distinguished vertex v_0. In the definition of the DFS algorithm (see below), the subgraph T remains a tree. The vertex v_0 is considered as the root-vertex of the tree. Hence, any vertex in T distinct from v_0 has a *father* in T.

DEFINITION 6.2. Depth-first search (DFS) algorithm.

Initialization: The current vertex is v_0 and the tree T is reduced to v_0.

Core: While the current vertex v is adjacent to a vertex not in T or is distinct from v_0 we do:
If there are some edges linking the current vertex v to a vertex not in T, we choose one of them e at random. We add e and its other endpoint v' to the tree T. The vertex v' becomes the current vertex.
Else, we backtrack, that is, we set the current vertex to be the father of v in T.

End: We return the tree T.

It is well known that the DFS algorithm returns a spanning tree. It is also known **[2]** that the two following properties are equivalent for a spanning tree T of a graph G having a distinguished vertex v_0:
(i) Any external edge joins comparable vertices.
(ii) The tree T can be obtained by a DFS algorithm on the graph G starting from v_0.

Before stating the main result of this section, we need an easy preliminary lemma.

LEMMA 6.3. *Let G be a connected graph with a distinguished vertex v_0 whose deletion does not disconnect the graph. Then, any spanning tree T satisfying conditions $(i)-(ii)$ has at exactly one edge incident to v_0.*

THEOREM 6.4. *Let G be a loopless connected graph with a distinguished vertex v_0 whose deletion does not disconnect the graph. Let e be an edge incident to v_0. If G is a k-near-cubic graph (v_0 has degree k and the other vertices have degree 3) of size n ($3n+2k-3$ edges), then the number of trees containing e and satisfying conditions $(i)-(ii)$ is 2^n.*

Given that the depth trees are the spanning trees satisfying conditions $(i)-(ii)$ and not containing the root, the following corollary is immediate.

COROLLARY 6.5. *For any 2-near-cubic map without isthmus of size n ($3n+1$ edges), there are 2^n depth trees.*

The proof of Theorem 6.4 relies on the intuition that exactly n *real* binary choices have to be made during the execution of a DFS algorithm on a k-near-cubic map of size n. However, making this intuition into a proof requires some work and we shall not do it here.

Remark: Theorem 6.4 shows that any k-near-cubic loopless graph of size n has $k2^n$ trees satisfying the conditions $(i)-(ii)$.

7. Extensions and open problems

7.1. Random generation of triangulations. We introduced the family of extended Kreweras walks ending on the second diagonal. The random generation of such walks of length $3n$ (with uniform distribution) reduces to the random generation of 1-dimensional walks of length $3n$ with steps $+2$, -1 starting and ending at 0 and remaining non-negative. The random generation of these walks is known to be feasible in linear time. (One just needs to generate (with uniform distribution) a word of length $3n+1$ containing n '+2' and $2n+1$ '-1' and to apply the cycle lemma.) Given an extended Kreweras walk w ending on the second diagonal, the construction of the 2-near-cubic depth-map $\Phi'(w)$ can be performed in linear time. Therefore, we have a linear time algorithm for the random generation (with uniform distribution) of 2-near-cubic depth-maps with a marked edge. For any 2-near-cubic map there are 2^n depth trees and then $(n+1)$ possible marked edges. Therefore, if we drop the distinguished edge and the depth tree at the end of the process, we obtain a uniform distribution on 2-near-cubic maps without isthmus. This allows us to generate uniformly cubic maps without isthmus or, dually, loopless triangulations in linear time.

7.2. Kreweras walks ending at $(i,0)$ and $(i+2)$-near-cubic maps. The Kreweras walks ending at $(i,0)$ are the words w on the alphabet $\{a,b,c\}$ with $|w|_a+i=|w|_b=|w|_c$ such that any suffix w' of w satisfies $|w'|_a+i\geq|w'|_c$ and $|w'|_b\geq|w'|_c$.
There is a very nice formula [**4**] counting Kreweras walks of size n (length $3n+2i$) ending at $(i,0)$:

$$k_{n,i} = \frac{4^n(2i+1)}{(n+i+1)(2n+2i+1)}\binom{2i}{i}\binom{3n+2i}{n}. \tag{7.1}$$

There is also a similar formula [**5**] for non-separable $(i+2)$-near-cubic maps of size n ($3n+2i+1$ edges):

$$c_{n,i} = \frac{2^n(2i+1)}{(n+i+1)(2n+2i+1)}\binom{2i}{i}\binom{3n+2i}{n}. \tag{7.2}$$

In this subsection, we show that the bijection Φ (Definition 4.2) can be extended to walks ending at $(i,0)$. This gives a bijective correspondence explaining why $k_{n,i}=2^nc_{n,i}$.
Consider the tree-growing map M_T^i reduced to a vertex, a root, a head and i left legs (Figure 14). We define the image of a Kreweras walk w ending at $(i,0)$ as the map obtained by closing $\varphi_w(M_T^i)$. We get the following extension of Theorem 4.3.

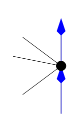


FIGURE 14. The tree-growing map M_T^i when $i=3$.

THEOREM 7.1. *The mapping Φ is a bijection between Kreweras walks of size n (length $3n+2i$) ending at $(i,0)$ and non-separable $(i+2)$-near-cubic maps of size n ($3n+2i+1$ edges) with a depth tree that contains the edge following the root in counterclockwise order around the root-vertex.*

By Theorem 6.4, there are 2^n such trees. Consequently, we obtain the following corollary:

COROLLARY 7.2. *The number $k_{n,i}$ of Kreweras walks of size n ending at $(i,0)$ and the number c_i of non-separable $(i+2)$-near-cubic maps of size n are related by the equation $k_{n,i}=2^nc_{n,i}$.*

One can define the counterpart of extended Kreweras walks in the case of walks ending at $(i, 0)$. These are the words obtained when one chooses an external edge (in a non-separable $(i+2)$-near-cubic depth-map such that the edge following the root is in the tree) and applies the mapping Ψ' which is the inverse of Φ'. We have no simple characterization of this set of words. However, it would be interesting to find a bijective proof that this set has size $\frac{4^n(2i+1)}{(2n+2i+1)}\binom{2i}{i}\binom{3n+2i}{n}$. We were not able to solve this problem yet.

Acknowledgments: I thank Mireille Bousquet-Mélou for steady encouragements and support.

References

[1] M. Bousquet-Mélou, Walks in the quarter plane: Kreweras' algebraic model, Ann. Appl. Probab. 15(2):1451-1491, 2005.
[2] T.H. Cormen, C.E. Leiserson and R.L. Rivest, Introduction to Algorithms, MIT press, 1990.
[3] I. M. Gessel, A probabilistic method for lattice path enumeration, J. Statist. Plann. Inference 14:49-58, 1986.
[4] G. Kreweras, Sur une classe de problmes liés au treillis des partitions d'entiers, Cahier du B.U.R.O. 6:5-105, 1965.
[5] R.C. Mullin, On counting rooted triangular maps, Canad. J. Math. 17:373-382, 1965.
[6] H. Niederhausen, The ballot problem with three candidates, European J. Combin. 4(2):175-188, 1983.
[7] G. Schaeffer, Conjugaison d'arbres et cartes combinatoires aléatoires, Université Bordeaux I, 1998.
[8] 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.
[9] R.P. Stanley, Clay Research Academy problems, 2005.
[10] R.P. Stanley, Enumerative combinatorics - volume 2, Cambridge university press, 1999.

LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France
E-mail address: bernardi@labri.fr

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A Proof of the q, t-Square Conjecture

Mahir Can and Nicholas Loehr

Abstract. We prove a combinatorial formula conjectured by Loehr and Warrington for the coefficient of the sign character in $\nabla(p_n)$. Here ∇ denotes the Bergeron-Garsia nabla operator, and p_n is a power-sum symmetric function. The combinatorial formula enumerates lattice paths in an $n \times n$ square according to two suitable statistics.

Résumé. Nous démontrons une formule combinatoire conjecturée par Loehr et Warrington concernant le coefficient du caratère signe dans $\nabla(p_n)$. Nous dénotons par ∇ l'operateur nabla de Bergeron-Garsia, et par p_n une fonction symmétrique en les puissances n-èmes. La formule combinatoire énumère les chemins, sur un rèseau carré de dimension $n \times n$, vérifiant deux statistiques.

1. Introduction

We begin with a quick overview of the remarkable q, t-*Catalan theorem* and the q, t-*square conjecture*. The q, t-Catalan theorem is the culmination of a series of papers by Garsia, Haglund, and Haiman [**2, 3, 4, 8, 9**]. The theorem states that, for every n, the following seemingly unrelated quantities are in fact equal:

1. the weighted sum of all *Dyck paths* of order n, weighted by *area* and *bounce score*;
2. the *Hilbert series* of the module of *diagonal harmonic alternants* of order n;
3. the n'th Garsia-Haiman q, t-*Catalan number*, which is a certain sum of complicated rational functions constructed from partitions;
4. the coefficient of the sign character in $\nabla(e_n)$, where ∇ is the *Bergeron-Garsia nabla operator* [**1, 10**], and e_n is an elementary symmetric function.

Precise definitions of the terms mentioned here will be given later (§2).

Loehr and Warrington [**12**] recently found an analogue of this theorem that involves q, t-analogues of lattice paths inside squares. Their result, which we call the q, t-*square conjecture*, states that the following five quantities are equal for every n:

1. the weighted sum of all $n \times n$ *square paths* ending in a north step, weighted by *area* and *bounce score*;
2. the weighted sum of all $n \times n$ *square paths* ending in an east step, weighted by *area* and *bounce score*;
3. a certain sum of rational functions analogous to the Garsia-Haiman q, t-Catalan number;
4. the coefficient of the sign character in $(-1)^{n-1}\nabla(p_n)$, where p_n is a power-sum symmetric function.

Again, we defer precise definitions of these quantities to §2. Loehr and Warrington proved that items 1 and 2 were equal, and also proved that items 3 and 4 were equal. Based on extensive computer calculations, they conjectured that all four items were equal.

The main theorem of our paper is a proof of this q, t-square conjecture. Here is a rough outline of the proof strategy. In light of previous results, it suffices to prove that item 1 equals item 4 for all n. We will

2000 *Mathematics Subject Classification.* Primary 05E05; Secondary 05A19.

Key words and phrases. symmetric functions, sign character, Dyck paths, square paths.

establish a refinement of this equality based on an expansion of p_n in terms of certain symmetric functions $E_{n,k}$ that appeared in the proof of the q,t-Catalan theorem. Explicitly, we will prove that

$$(-1)^{n-1}p_n = \sum_{k=1}^{n} \frac{1-q^n}{1-q^k} E_{n,k}. \tag{1.1}$$

(In contrast, $e_n = \sum_{k=1}^{n} E_{n,k}$.) Applying nabla and taking the coefficient of s_{1^n} gives

$$\langle (-1)^{n-1}\nabla(p_n), s_{1^n}\rangle = \sum_{k=1}^{n} \frac{1-q^n}{1-q^k} \langle \nabla(E_{n,k}), s_{1^n}\rangle. \tag{1.2}$$

Garsia and Haglund previously found combinatorial formulas and recursions for $\langle \nabla(E_{n,k}), s_{1^n}\rangle$. Comparing these results to recursions involving q,t-analogues of square lattice paths, we will show that each summand on the right side of (1.2) enumerates a suitable subcollection of the q,t-square paths mentioned in item 1. The equality of item 1 and item 4 will readily follow.

The rest of this paper is organized as follows. Section 2 reviews the minimal framework of definitions needed to give precise statements of the q,t-Catalan theorem and the q,t-square conjecture. Section 3 discusses some (previously known) technical results needed in our proof of the q,t-square conjecture. Section 4 uses a plethystic calculation to prove the fundamental expansion (1.1). Section 5 analyzes a combinatorial recursion that lets us identify the square q,t-lattice paths enumerated by each summand in (1.2). Section 6 concludes by discussing some natural open problems pertaining to the q,t-Catalan theorem and the q,t-square conjecture.

2. Definitions

This section reviews the definitions of the concepts appearing in the q,t-Catalan theorem and the q,t-square conjecture. Precise statements of these two results are given at the end of this section. We assume familiarity with standard background material on partitions, symmetric functions, representation theory, Macdonald polynomials, and lattice paths [**6, 13, 14, 15**]. Readers who find this section too terse may wish to consult the more leisurely treatment contained in the introduction of [**12**].

2.1. Partition Definitions. We write $\mu \vdash n$ to indicate that μ is a partition of n. The *diagram* of μ is

$$\mathrm{dg}(\mu) = \{(i,j) \in \mathbb{N}\times\mathbb{N} : 1 \le i \le \mu_j\}.$$

Let $c = (i_0, j_0)$ be a cell in $\mathrm{dg}(\mu)$. The *arm* of c is $a(c) = |\{(i,j_0) \in \mathrm{dg}(\mu) : i > i_0\}|$. The *coarm* of c is $a'(c) = |\{(i,j_0) \in \mathrm{dg}(\mu) : i < i_0\}|$. The *leg* of c is $l(c) = |\{(i_0,j) \in \mathrm{dg}(\mu) : j > j_0\}|$. The *coleg* of c is $l'(c) = |\{(i_0,j) \in \mathrm{dg}(\mu) : j < j_0\}|$. Let $n(\mu) = \sum_{c\in\mathrm{dg}(\mu)} l(c)$, and let μ' be the transpose of μ. In the ring $\mathbb{Z}[q,t] \subseteq \mathbb{Q}(q,t)$, define $M = (1-q)(1-t)$, $B_\mu = \sum_{c\in\mathrm{dg}(\mu)} q^{a'(c)}t^{l'(c)}$, $\Pi_\mu = \prod_{(1,1)\neq c\in\mathrm{dg}(\mu)}(1-q^{a'(c)}t^{l'(c)})$, $T_\mu = q^{n(\mu')}t^{n(\mu)}$, and $w_\mu = \prod_{c\in\mathrm{dg}(\mu)}[(q^{a(c)}-t^{l(c)+1})(t^{l(c)}-q^{a(c)+1})]$.

The n'th Garsia-Haiman q,t-Catalan number is defined by the formula

$$\sum_{\mu\vdash n} \frac{T_\mu^2 M B_\mu \Pi_\mu}{w_\mu} \in \mathbb{Q}(q,t). \tag{2.1}$$

This sum of rational functions evaluates to a polynomial in $\mathbb{N}[q,t]$, although this is quite hard to prove. The analogous expression appearing in the q,t-square conjecture is

$$\sum_{\mu\vdash n} \frac{T_\mu^2 M B_{(n^n)} \Pi_\mu}{w_\mu} \in \mathbb{Q}(q,t). \tag{2.2}$$

The only difference is that B_μ has been replaced by the constant $B_{(n^n)}$, where $(n^n) \vdash n^2$ consists of n parts equal to n. It is easy to see that this expression can also be written

$$(1-t^n)(1-q^n)\sum_{\mu\vdash n} \frac{T_\mu^2 \Pi_\mu}{w_\mu}. \tag{2.3}$$

2.2. Symmetric Function Definitions. We write Λ to denote the ring of symmetric functions with coefficients in the field $F = \mathbb{Q}(q,t)$. As usual, e_n will denote the n'th elementary symmetric function, p_n will denote the n'th power-sum symmetric function, s_μ will denote the Schur function indexed by a partition μ, and $\tilde{H}_\mu$ will denote the modified Macdonald polynomial indexed by μ. The Bergeron-Garsia *nabla operator* is the unique linear operator on Λ such that $\nabla(\tilde{H}_\mu) = T_\mu \tilde{H}_\mu$. The *Hall scalar product* $\langle \cdot, \cdot \rangle$ on Λ is defined by requiring that the Schur functions be an orthonormal basis. If $f \in \Lambda$ is the Frobenius character of some S_n-module M, then $\langle f, s_{1^n} \rangle$ gives the multiplicity of the sign representation in M. Accordingly, for any $f \in \Lambda$, we often call $\langle f, s_{1^n} \rangle$ the "coefficient of the sign character in f." We remark in passing that the well-known identities

$$e_n = \sum_{\mu \vdash n} (MB_\mu \Pi_\mu / w_\mu) \tilde{H}_\mu, \qquad (-1)^{n-1} p_n = \sum_{\mu \vdash n} (MB_{(n^n)} \Pi_\mu / w_\mu) \tilde{H}_\mu,$$

and $\langle \tilde{H}_\mu, s_{1^n} \rangle = T_\mu$ easily imply that $\langle \nabla(e_n), s_{1^n} \rangle$ and $\langle (-1)^{n-1} \nabla(p_n), s_{1^n} \rangle$ are given by formulas (2.1) and (2.2), respectively.

2.3. Path Definitions. A *Dyck path* of order n is a lattice path in the x,y-plane that starts at the origin, consists of n unit north steps (N) and n unit east steps (E), and always stays weakly above the line $y = x$. We let $\mathcal{DP}_n$ denote the set of Dyck paths of order n. For $P \in \mathcal{DP}_n$, define $\mathrm{area}(P)$ to be the number of complete lattice squares bounded by P and the line $y = x$. Next, we define Haglund's *bounce path* for P, which consists of certain quantities $v_i(P)$ for $i \geq 0$. We imagine a ball bouncing from (n,n) to $(0,0)$ and being deflected by P and the diagonal $y = x$. At stage $i \geq 0$, the ball is on the line $y = x$ and goes $v_i(P)$ units west until it is blocked by the upper end of a north step of P. The ball then "bounces" south $v_i(P)$ units back to the diagonal. This bouncing continues until the ball reaches $(0,0)$. The *bounce score* of P is $\mathrm{bounce}(P) = \sum_{i \geq 0} i v_i(P)$. For example, the path P encoded by NNNENNEENENENEEE lies in $\mathcal{DP}_8$ and has $\mathrm{area}(P) = 14$, $v_0(P) = 3$, $v_1(P) = 4$, $v_2(P) = 1$, and $\mathrm{bounce}(P) = 6$.

A *square path* of order n is a lattice path in the x,y-plane that starts at the origin and consists of n unit north steps and n unit east steps. We write $\mathcal{SQP}_n$, $\mathcal{SQP}_n^N$, and $\mathcal{SQP}_n^E$ to denote (respectively) the set of all such square paths, the set of all such paths ending with a north step, and the set of all such paths ending in an east step. To define analogues of bounce and area as in [**12**], we need some auxiliary concepts. Fix a square path P. Consider the diagonal lines $y = x - c$, for $c = 0, 1, 2, \ldots$. The lowest such diagonal that meets the path P is called the *base diagonal*. The lowest point on P touching the base diagonal is called the *breakpoint*. We now let $\mathrm{area}(P)$ be the number of complete lattice squares in the region bounded on the left by P, on the right by the base diagonal, on the top by $y = n$ and on the bottom by $y = 0$. For example, the path P encoded by NEENEENNENEEENNENENENEENNNENNNEE lies in $\mathcal{SQP}_{15}^E$ and has base diagonal $y = x - 3$, breakpoint $(8,5)$, and area 25.

Now we define square bounce paths. Given P, let $y = x - c$ be the base diagonal for P. This time there are two bouncing balls. The first ball starts at (n,n) and moves vertically c units south to the base diagonal. Thereafter, the ball bounces west and south as in the Dyck path case, with all southward moves terminating on the base diagonal, until it reaches the breakpoint. The second ball starts at $(0,0)$ and moves horizontally c units east to the base diagonal. This ball proceeds to bounce northeast to the breakpoint as follows. Starting at the base diagonal, the ball moves north until it is blocked by the end of an east step of P. It then moves the same distance east to reach the base diagonal. (This is *not* simply a reflected version of the bouncing policy followed by the first ball!) Let the vertical moves made by the first ball (in order) have lengths $v_0'(P) = c, v_1'(P), \ldots, v_s'(P)$, and let the vertical moves made by the second ball have lengths $v_0''(P), \ldots, v_t''(P)$. We then define $(v_0(P), v_1(P), \ldots) = (v_t''(P), \ldots, v_0''(P), v_0'(P), \ldots, v_s'(P))$ and set $\mathrm{bounce}(P) = \sum_{i \geq 0} i v_i(P)$ as before. For the specific example considered above, the first ball moves south 3, west 3, south 3, west 2, south 2, west 2, south 2, while the second ball moves east 3, north 2, east 2, north 2, east 2, north 1, east 1. Accordingly, $(v_0(P), v_1(P), \ldots) = (1,2,2,3,3,2,2)$ and $\mathrm{bounce}(P) = 49$. See Figure 1.

2.4. q-Definitions. For $n \geq 1$, define $[n]_q = 1 + q + q^2 + \cdots + q^{n-1} = (1-q^n)/(1-q)$, $(a;q)_n = (1-a)(1-aq)(1-aq^2)\cdots(1-aq^{n-1})$, and $[n]!_q = (q;q)_n/(1-q)^n = \prod_{i=1}^n [i]_q$. We also set $[0]_q = 1 = (a;q)_0$.

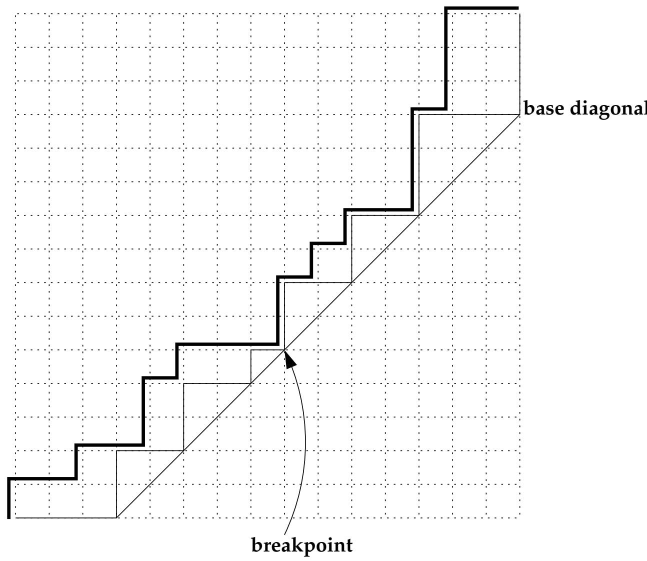


Figure 1. Square bouncing.

We define the q-binomial coefficient by setting

$$\begin{bmatrix} m+n \\ m,n \end{bmatrix}_q = \frac{(q;q)_{m+n}}{(q;q)_m (q;q)_n} = \frac{[m+n]!_q}{[m]!_q [n]!_q}.$$

It is well-known that $\begin{bmatrix} m+n \\ m,n \end{bmatrix}_q = \sum_{\mu \subseteq (m^n)} q^{|\mu|}$. Thus the q-binomial coefficient enumerates partitions (or lattice paths) contained in an $m \times n$ rectangle, weighted by area.

2.5. Precise Statements of Theorems. We summarize the preceding definitions by giving precise versions of the q,t-Catalan theorem and the q,t-square conjecture.

Theorem 2.1 (Garsia-Haglund-Haiman q,t-Catalan Theorem). *For all $n \geq 1$, we have*

$$\sum_{P \in \mathcal{DP}_n} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)} = \mathrm{Hilb}(DHA_n) = \sum_{\mu \vdash n} \frac{T_\mu^2 M B_\mu \Pi_\mu}{w_\mu} = \langle \nabla(e_n), s_{1^n} \rangle.$$

In particular, the last two expressions are elements of $\mathbb{N}[q,t]$.

Theorem 2.2 (Loehr-Warrington). *For all $n \geq 1$, we have*

$$\sum_{P \in \mathcal{SQP}_n^N} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)} = \sum_{P \in \mathcal{SQP}_n^E} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)}$$

$$\text{and} \qquad \sum_{\mu \vdash n} \frac{T_\mu^2 M B_{(n^n)} \Pi_\mu}{w_\mu} = \langle (-1)^{n-1} \nabla(p_n), s_{1^n} \rangle.$$

Conjecture 2.3 (Loehr-Warrington). For all $n \geq 1$, all four quantities in the previous theorem are equal. In particular, the last two expressions are elements of $\mathbb{N}[q,t]$.

The rest of this paper is devoted to a proof of this conjecture. More specifically, we will prove a refinement of the identity

$$\langle (-1)^{n-1} \nabla(p_n), s_{1^n} \rangle = \sum_{P \in \mathcal{SQP}_n^N} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)}.$$

3. Technical Results

This section states without proof some known results of a somewhat technical nature that will be needed to establish the q,t-square conjecture.

3.1. Plethysm. We begin with some fundamental facts about plethystic notation. This material is treated in much greater detail in [**6, 11**].

Recall that Λ can be viewed as a polynomial ring $\Lambda = F[p_1, p_2, \ldots, p_k, \ldots]$. Like any polynomial ring, Λ enjoys a universal mapping property (UMP) that says that any function g mapping the set $\{p_k : k \geq 1\}$ into an F-algebra S extends uniquely to an F-algebra homomorphism from Λ into S. This homomorphism is often called the *evaluation homomorphism* determined by g.

Now, if $f \in \Lambda$ and A is a "plethystic alphabet," the plethystic substitution $f[A]$ is defined to be the image of f under the evaluation homomorphism determined by a certain function g_A. This function g_A is itself determined by A according to the rules for interpreting plethystic alphabets. We shall only need three special cases of this definition:

1. $f[X(1-z)/(1-q)]$ is the image of f under the F-algebra homomorphism from Λ to $\Lambda[z]$ such that $p_k \mapsto p_k(1-z^k)/(1-q^k)$.
2. $f[X/(1-q)]$ is the image of f under the F-algebra homomorphism from Λ to Λ such that $p_k \mapsto p_k/(1-q^k)$.
3. $f[1-z]$ is the image of f under the F-algebra homomorphism from Λ to $F[z]$ such that $p_k \mapsto 1-z^k$.

We also have the trivial substitutions $f[X] = f$ and $f[0] = 0$ for $f \in \Lambda$.

We now state three (standard) facts about plethysm needed in our proof. First, for any alphabets A and B, we have the dual Cauchy identity

$$e_n[AB] = \sum_{\mu \vdash n} s_\mu[A] s_{\mu'}[B]. \tag{3.1}$$

Second, for all partitions μ, we have

$$s_\mu[1-z] = \begin{cases} (-z)^a(1-z) & \text{if } \mu = (n-a, 1^a) \text{ for some } a \in \{0,1,2,\ldots,n-1\}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Third, $e_n[X(1-z)/(1-q)]$ is an element of the polynomial ring $\Lambda[z]$ of degree at most n in z.

3.2. Definition of $E_{n,k}$ and $F_{n,k}$. We can now define the symmetric functions $E_{n,k}$ mentioned in the introduction. Let M be the subset of $\Lambda[z]$ consisting of polynomials of degree at most n in z. Clearly, M is a free Λ-module with basis $1, z, z^2, \ldots, z^n$. Easy degree considerations show that the set $\{(z;q)_k/(q;q)_k : 0 \leq k \leq n\}$ is also a basis for M. Combining this observation with the third fact from the last subsection, we see that there exist unique elements $E_{n,k} \in \Lambda$ such that

$$e_n\left[\frac{X(1-z)}{1-q}\right] = \sum_{k=0}^{n} \frac{(z;q)_k}{(q;q)_k} E_{n,k}. \tag{3.3}$$

Setting $z=1$, we see that $E_{n,0} = 0$, while setting $z = q$ shows that $e_n = \sum_{k=1}^n E_{n,k}$.[1]

Define $F_{n,k} = \langle \nabla(E_{n,k}), s_{1^n} \rangle \in \mathbb{Q}(q,t)$. Garsia and Haglund showed [**3, 5**] that the $F_{n,k}$ satisfy the recurrence

$$F_{n,k} = q^{k(k-1)/2} t^{n-k} \sum_{r=0}^{n-k} \begin{bmatrix} r+k-1 \\ r, k-1 \end{bmatrix}_q F_{n-k,r} \tag{3.4}$$

with initial conditions $F_{n,0} = \delta_{n0}$. On the other hand, let $\mathcal{DP}_{n,k}$ be the set of all Dyck paths P of order n that end in exactly k east steps. Equivalently, $\mathcal{DP}_{n,k} = \{P \in \mathcal{DP}_n : v_0(P) = k\}$. Let $F'_{n,k} = \sum_{P \in \mathcal{DP}_{n,k}} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)}$. By "removing the first bounce" in the bounce path for P, one easily sees that the $F'_{n,k}$ satisfy the same recurrence and initial conditions as $F_{n,k}$. Therefore, $F_{n,k} = F'_{n,k}$ for all n and k, i.e.,

$$\langle \nabla(E_{n,k}), s_{1^n} \rangle = \sum_{P \in \mathcal{DP}_{n,k}} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)} \in \mathbb{N}[q,t]. \tag{3.5}$$

This formula provides the fundamental link between the nabla operator and the combinatorics of q,t-Dyck paths.

[1] These statements, while true, require a subtle additional justification. See [**11**] for a detailed discussion.

4. Expansion of p_n via $E_{n,k}$'s

Theorem 4.1. *For all $n \geq 1$,*

$$(-1)^{n-1}p_n = \sum_{k=1}^{n} \frac{1-q^n}{1-q^k} E_{n,k}.$$

Proof. Using (3.1) and (3.2), we compute

$$\begin{aligned} e_n\left[\frac{X(1-z)}{1-q}\right] &= \sum_{\mu\vdash n} s_\mu[X/(1-q)]s_{\mu'}[1-z] \\ &= \sum_{\substack{\mu\vdash n \\ \mu'=(n-a,1^a)}} s_\mu[X/(1-q)](-z)^a(1-z) \\ &= \sum_{a=1}^{n} s_{(a,1^{n-a})}[X/(1-q)](-z)^{a-1}(1-z). \end{aligned}$$

On the other hand, using (3.3) and $E_{n,0}=0$, we get

$$e_n\left[\frac{X(1-z)}{1-q}\right] = \sum_{k=1}^{n} \frac{(z;q)_k}{(q;q)_k} E_{n,k} = \sum_{k=1}^{n} \frac{(1-z)(zq;q)_{k-1}}{(q;q)_k} E_{n,k}.$$

Comparing the two expressions for $e_n[X(1-z)/(1-q)]$ and cancelling $1-z$ in the integral domain $\Lambda[z]$, we obtain

$$\sum_{a=1}^{n}(-z)^{a-1}s_{(a,1^{n-a})}[X/(1-q)] = \sum_{k=1}^{n} \frac{(zq;q)_{k-1}}{(q;q)_k} E_{n,k}.$$

Now apply the evaluation homomorphism $\Lambda[z] \to \Lambda$ sending z to 1:

$$\sum_{a=1}^{n}(-1)^{a-1}(s_{(a,1^{n-a})}[X/(1-q)]) = \sum_{k=1}^{n} \frac{(q;q)_{k-1}}{(q;q)_k} E_{n,k} = \sum_{k=1}^{n} \frac{E_{n,k}}{1-q^k}.$$

By the Pieri rule and linearity of plethysm, the left side here is

$$\begin{aligned} \sum_{a=1}^{n}(-1)^{a-1}(s_{(a,1^{n-a})}[X/(1-q)]) &= \left(\sum_{a=1}^{n}(-1)^{a-1}s_{(a,1^{n-a})}\right)[X/(1-q)] \\ &= ((-1)^{n-1}p_n)[X/(1-q)] = (-1)^{n-1}p_n/(1-q^n). \end{aligned}$$

Putting this into the previous formula and multiplying through by $1-q^n$, we obtain the theorem. $\square$

By linearity of ∇ and the Hall scalar product, we immediately deduce the following corollary.

Corollary 4.2.

$$\langle (-1)^{n-1}\nabla(p_n), s_{1^n}\rangle = \sum_{k=1}^{n} \frac{[n]_q}{[k]_q} F_{n,k}.$$

5. Combinatorial Recursion Analysis

In this section, we will identify each summand in the last corollary as the weighted sum of a suitable subcollection of square lattice paths. Specifically, define

$$S_{n,k} = \sum_{P\in\mathcal{SQP}_n^N : v_0(P)=k} q^{\mathrm{area}(P)}t^{\mathrm{bounce}(P)}.$$

We will prove that $S_{n,k} = \frac{[n]_q}{[k]_q}F_{n,k}$. The q,t-square conjecture will easily follow from this fact and the corollary. To obtain these results, we first derive a recursion characterizing $S_{n,k}$.

Theorem 5.1. *We have $S_{n,n} = q^{n(n-1)/2} = F_{n,n}$ for all n. For all $n \geq 1$ and $1 \leq k < n$, we have*

$$S_{n,k} = F_{n,k} + q^{k(k-1)/2}t^{n-k}\sum_{r=1}^{n-k} q^k \begin{bmatrix} r-1+k \\ r-1, k \end{bmatrix}_q S_{n-k,r} \tag{5.1}$$

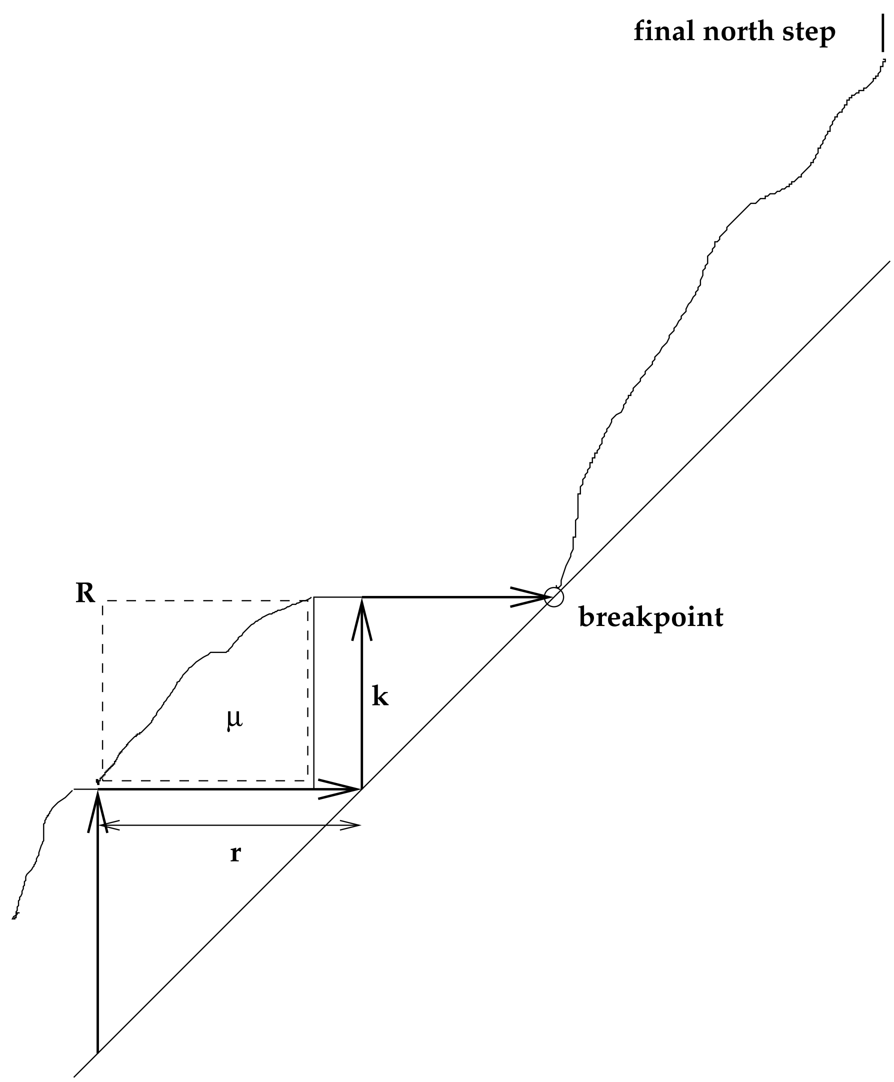


FIGURE 2. Removing the last negative bounce in case 2.

PROOF. Recall the combinatorial description of $F_{n,k}$ from §3.2:

$$F_{n,k} = \sum_{P \in \mathcal{DP}_n : v_0(P) = k} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)}.$$

The proof of the recurrence for $S_{n,k}$ is so similar to the proof of an analogous recurrence in [**12**] that we only sketch the details. (See Theorem 7 in [**12**] — the difference between the $R_{n,k}$ appearing there and the $S_{n,k}$ appearing here is that we demand that our paths end in a north step. This extra condition simplifies the recursion considerably.) Let P be a path counted by $S_{n,k}$. Observe that P is not a Dyck path, since it ends in a north step. Now consider two cases.

Case 1: The break point of P lies on the line $y = 0$. Then P must begin with k east steps. Moving these east steps to the end of the path and translating the break point to the origin, we obtain a typical path P' counted by $F_{n,k}$. (Note that $v_0(P') = k$ because P ends in a north step.) The map $P \mapsto P'$ defines a bijection between the paths P occurring in case 1 and the paths P' counted by $F_{n,k}$. Area and bounce are clearly preserved, so we have explained the first summand in (5.1).

Case 2: The break point of P lies above the line $y = 0$. We know $v_0(P) = k$; define $r = v_1(P)$, which is always the length of the horizontal move preceding the last vertical move made by the second bouncing ball.

A typical situation is pictured in Figure 2 — but note that the horizontal move of length r may also occur on the line $y = 0$. We can map P to a certain triple (r, μ, P'), where $r = v_1(P) \in \{1, 2, \ldots, n-k\}$, μ is the partition contained in the rectangle $R = ((r-1)^k)$ shown in the figure, and P' is a typical path counted by $S_{n-k,r}$. We obtain P' by merely erasing everything in the k rows immediately below the breakpoint, and then translating the part of P above the breakpoint k units down along the base diagonal. This has the effect of "removing the last bounce" made by the second bouncing ball. The rest of the bounce paths are unaffected by this shift, and it readily follows that $v_i(P') = v_{i+1}(P)$ and $\mathrm{bounce}(P) = \mathrm{bounce}(P') + n - k$. The breakpoint cannot be located at (n, n), so P' still ends in a north step. Since $P' \in \mathcal{SQP}_{n-k}^N$ and $v_0(P') = r$, P' is a path counted by $S_{n-k,r}$. It is not hard to see that $\mathrm{area}(P) = \mathrm{area}(P') + k(k-1)/2 + k + |\mu|$; here the $k(k-1)/2$ accounts for area cells to the right of the last vertical move made by the second ball, and the k accounts for the area cells in the column just left of this bounce, which is not part of μ. Finally, the passage from P (in case 2) to triples (r, μ, P') with $r \in \{1, 2, \ldots, n-k\}$, $\mu \subseteq (r-1)^k$, and $P' \in \mathcal{SQP}_{n-k}^N$ with $v_0(P') = r$ is clearly a bijection. Combining all these facts, we obtain the remaining terms in the recurrence (5.1). □

THEOREM 5.2. *For all $n \geq 1$ and all $k \leq n$, $S_{n,k} = \frac{[n]_q}{[k]_q} F_{n,k}$.*

PROOF. The theorem holds for all n when $k = n$, since $S_{n,n} = F_{n,n}$ in this case. For the remaining cases, we use induction on n. Using the induction hypothesis to replace $S_{n-k,r}$ in the recursion (5.1), we first obtain

$$S_{n,k} = F_{n,k} + q^{k(k-1)/2} t^{n-k} \sum_{r=1}^{n-k} \frac{[r+k-1]!_q}{[r-1]!_q [k]!_q} \left(\frac{[n-k]_q}{[r]_q} F_{n-k,r} \right) q^k.$$

Rearranging the q-numbers here, the right side can be written

$$F_{n,k} + \frac{q^k [n-k]_q}{[k]_q} \left(q^{k(k-1)/2} t^{n-k} \sum_{r=1}^{n-k} \begin{bmatrix} r+k-1 \\ r, k-1 \end{bmatrix}_q F_{n-k,r} \right).$$

Comparing to (3.4), we see that the term in parentheses is just $F_{n,k}$ again! So the calculation continues:

$$S_{n,k} = F_{n,k} \left(1 + \frac{q^k [n-k]_q}{[k]_q} \right) = \frac{[n]_q}{[k]_q} F_{n,k}.$$

This completes the induction step and the proof. □

Evidently $\mathcal{SQP}_n^N$ is the disjoint union of its subsets $\{P \in \mathcal{SQP}_n^N : v_0(P) = k\}$ as k ranges from 1 to n. Combining this fact with Theorem 5.2 and Corollary 4.2, we obtain our desired result:

COROLLARY 5.3.

$$\langle (-1)^{n-1} \nabla(p_n), s_{1^n} \rangle = \sum_{P \in \mathcal{SQP}_n^N} q^{\mathrm{area}(P)} t^{\mathrm{bounce}(P)}.$$

In particular, $\langle (-1)^{n-1} \nabla(p_n), s_{1^n} \rangle = \sum_{\mu \vdash n} \frac{T_\mu^2 M B_{(n^n)} \Pi_\mu}{w_\mu}$ is an element of $\mathbb{N}[q, t]$.

6. Conclusion

Comparing the q,t-Catalan theorem to the q,t-square theorem (as we shall now call it), one obvious difference is apparent: the latter theorem does not identify $\langle (-1)^{n-1} \nabla(p_n), s_{1^n} \rangle$ as the Hilbert series of some doubly graded module. Of course, one could define such a module by taking a direct sum of sign representations indexed by square paths, using the area and bounce statistics to determine the bigrading. We leave it as an open problem to find a less artificial solution, i.e., "naturally occurring" modules M_n carrying only the sign representation such that $\mathrm{Hilb}(M_n)$ is given by the quantities in the q,t-square theorem. More generally, one could seek modules whose Frobenius characters are given by $(-1)^{n-1} \nabla(p_n)$, just as the Frobenius characters of the diagonal harmonics modules DH_n are given by $\nabla(e_n)$ [**9**].

In closing, we recall that combinatorial interpretations have been proposed [**7, 12**] for the monomial expansions of the symmetric functions $\nabla(e_n)$, $\nabla(E_{n,k})$, and $\nabla(p_n)$. These conjectures involve labelled versions of Dyck paths and square lattice paths. At the time of this writing, all these conjectures are still open.

References

[1] F. Bergeron and A. Garsia, "Science Fiction and Macdonald Polynomials," *CRM Proceedings and Lecture Notes AMS VI* **3** (1999), 363—429.
[2] A. Garsia and J. Haglund, "A proof of the q, t-Catalan positivity conjecture," LACIM 2000 Conference on Combinatorics, Computer Science, and Applications (Montreal), *Discrete Math.* **256** (2002), 677—717.
[3] A. Garsia and J. Haglund, "A positivity result in the theory of Macdonald polynomials," *Proc. Nat. Acad. Sci. USA* **98** (2001), 4313—4316.
[4] A. Garsia and M. Haiman, "A remarkable q, t-Catalan sequence and q-Lagrange Inversion," *J. Algebraic Combinatorics* **5** (1996), 191—244.
[5] J. Haglund, "A proof of the q, t-Schröder conjecture," *Intl. Math. Res. Notices* **11** (2004), 525—560.
[6] J. Haglund, *The q, t-Catalan Numbers and the Space of Diagonal Harmonics.* AMS University Lecture Series, to appear.
[7] J. Haglund, M. Haiman, N. Loehr, J. B. Remmel, and A. Ulyanov, "A combinatorial formula for the character of the diagonal coinvariants," *Duke Math. J.* **126** #2 (2005), 195—232.
[8] M. Haiman, "Hilbert schemes, polygraphs, and the Macdonald positivity conjecture." *J. Amer. Math. Soc.* **14** (2001), 941—1006.
[9] M. Haiman. "Vanishing theorems and character formulas for the Hilbert scheme of points in the plane," *Invent. Math.* **149** (2002), 371—407.
[10] C. Lenart, "Lagrange inversion and Schur functions," *J. Algebraic Combin.* **11** (2000), 69—78.
[11] N. Loehr, "An exposé of Adriano Garsia's plethystic notation," in preparation.
[12] N. Loehr and G. Warrington, "Square q, t-Lattice Paths and $\nabla(p_n)$," *Trans. Amer. Math. Soc.*, in press.
[13] I. G. Macdonald, *Symmetric Functions and Hall Polynomials.* 2nd ed. Oxford University Press, 1995.
[14] B. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions.* Wadsworth and Brooks/Cole, 1991.
[15] R. Stanley, *Enumerative Combinatorics*, Vol. 2. Cambridge University Press, Cambridge, United Kingdom (1999).

Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104
E-mail address: `canm@math.upenn.edu`

Department of Mathematics, College of William & Mary, Williamsburg, WA 23187
E-mail address: `nick@math.wm.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Characterization of Eulerian binomial and Sheffer posets

Richard Ehrenborg and Margaret A. Readdy

Abstract. We completely characterize the factorial functions of Eulerian binomial posets. The factorial function $B(n)$ either coincides with $n!$, the factorial function of the infinite Boolean algebra, or 2^{n-1}, the factorial function of the infinite butterfly poset. We also classify the factorial functions for Eulerian Sheffer posets. An Eulerian Sheffer poset with binomial factorial function $B(n) = n!$ has Sheffer factorial function $D(n)$ identical to that of the infinite Boolean algebra, the infinite Boolean algebra with two new coatoms inserted, or the infinite cubical poset. Moreover, we are able to classify the Sheffer factorial functions of Eulerian Sheffer posets with binomial factorial function $B(n) = 2^{n-1}$ as the doubling of an upside-down tree with ranks 1 and 2 modified.

When we impose the further condition that a given Eulerian binomial or Eulerian Sheffer poset is a lattice, this forces the poset to be the infinite Boolean algebra $\mathbb{B}_X$ or the infinite cubical lattice $\mathbb{C}_X^{<\infty}$. We also include several poset constructions that have the same factorial functions as the infinite cubical poset, demonstrating that classifying Eulerian Sheffer posets is a difficult problem.

Résumé. Nous caractérisons complétement les fonctions factorielles des ensembles partiellement ordonnés (posets) binomiaux Eulériens. La fonction factorielle $B(n)$ coincide avec $n!$, la fonction factorielle de l'algèbre de Boole infinie, ou avec 2^{n-1}, la fonction factorielle de l'ensemble partiellement ordonné "papillon" infini. Nous classifions aussi les fonctions factorielles des ensembles partiellement ordonnés (posets) de Sheffer Eulériens. Un poset de Sheffer Eulérien dont la fonction binomiale factorielle est $B(n) = n!$ a la fonction factorielle de Sheffer $D(n)$ indentique avec celle de l'algèbre de Boole infinie, ou avec celle de l'algèbre de Boole infinie avec deux nouveaux coatômes insérés, ou avec celle de l'ensemble partiellement ordonné cubique infini. De plus, nous pouvons classifier les fonctions factorielles de Sheffer des posets de Sheffer Eulériens avec la fonction binomiale factorielle $B(n) = 2^{n-1}$ comme le doublement d'un arbre á l'envers avec les rangs 1 et 2 modifiés.

Quand nous démandons la condition additionnelle qu'un poset binomial Eulérien ou Sheffer Eulérien soit un treillis, celle-ci force l'ensemble à être l'algèbre de Boole infinie $\mathbb{B}_X$ ou le treillis cubique infini $\mathbb{C}_X^{<\infty}$. Plusieures constructions des posets sont inclus qui possèdent les mêmes fonctions factorielles que le poset cubique infini, ce qui demontre que la classification des posets de Sheffer Eulériens est une problème trés difficile.

1. Introduction

Binomial posets were introduced by Doubilet, Rota and Stanley [**5**] to explain why generating functions naturally occurring in combinatorics have certain forms. They are highly regular posets since the essential requirement is that every two intervals of the same length have the same number of maximal chains. As a result, many poset invariants are determined. For instance, the quintessential Möbius function is described

---

2000 *Mathematics Subject Classification.* Primary 05E99; Secondary 06A07.

Key words and phrases. factorial function, poset structure, cubical posets, lattices, triangular posets.

The first author was partially supported by National Science Foundation grant 0200624 and by a University of Kentucky College of Arts & Sciences Faculty Research Fellowship.

The second author was partially supported by a University of Kentucky College of Arts & Sciences Research Grant. Both authors thank Gábor Hetyei for inspiring them to study Eulerian binomial posets, the Banff International Research Station where some of the ideas for this paper were developed, and the Mittag-Leffler Institute where this paper was completed.

by the generating function identity

$$\sum_{n \geq 0} \mu(n) \cdot \frac{t^n}{B(n)} = \left(\sum_{n \geq 0} \frac{t^n}{B(n)} \right)^{-1}, \tag{1.1}$$

where $\mu(n)$ is the Möbius function of an n-interval and $B(n)$ is the factorial function, that is, the number of maximal chains in an n-interval. A binomial poset is required to contain an infinite chain so that there are intervals of any length in the poset.

A graded poset is *Eulerian* if its Möbius function is given by $\mu(x,y) = (-1)^{\rho(y)-\rho(x)}$ for all $x \leq y$ in the poset. Equivalently, every interval of the poset satisfies the Euler-Poincaré relation: the number of elements of even rank is equal to the number of elements of odd rank in the interval. The major example of Eulerian posets are face lattices of convex polytopes and more generally, the face posets of regular CW-spheres. Hence there is a large geometric and topological interest in understanding them.

A natural question arises: which binomial posets are Eulerian? By equation (1.1) it is clear that the Eulerian property can be determined by knowing the factorial function. In this paper we classify the factorial functions of Eulerian binomial posets. There are two possibilities, namely, for the factorial function to correspond to that of the infinite Boolean algebra or the infinite butterfly poset.

Notice that this classification is on the level of the factorial function, not the poset itself. There are more Eulerian binomial posets than these two essential examples. See Examples 2.9 and 2.10. However, we are able to classify the intervals of Eulerian binomial posets. They are either isomorphic to the finite Boolean algebra or the finite butterfly poset.

Sheffer posets were introduced by Reiner [**10**] and independently by Ehrenborg and Readdy [**6**]. A Sheffer poset requires the number of maximal chains of an interval $[x,y]$ of length n to be given by $B(n)$ if $x > \hat{0}$ and $D(n)$ if $x = \hat{0}$. The upper intervals $[x,y]$ where $x > \hat{0}$ have the property of being binomial. Hence the interest is to understand the Sheffer intervals $[\hat{0},y]$. Just like binomial posets, the Möbius function is completely determined:

$$\sum_{n \geq 1} \overline{\mu}(n) \frac{t^n}{D(n)} = - \left(\sum_{n \geq 1} \frac{t^n}{D(n)} \right) \cdot \left(\sum_{n \geq 0} \frac{t^n}{B(n)} \right)^{-1}, \tag{1.2}$$

where $\overline{\mu}$ is the Möbius function of a Sheffer interval of length n; see [**6, 10**].

The classic example of a Sheffer poset is the infinite cubical poset (see Example 3.6). In this case, every interval $[x,y]$ of length n, where x is not the minimal element $\hat{0}$, has $n!$ maximal chains. In fact, every such interval is isomorphic to a Boolean algebra. Intervals of the form $[\hat{0},y]$ have $2^{n-1} \cdot (n-1)!$ maximal chains and are isomorphic to the face lattice of a finite dimensional cube.

In sections 3 and 4 we completely classify the factorial functions of Eulerian Sheffer posets. The factorial function $B(n)$ follows from the classification of binomial posets. The pair of factorial functions $B(n)$ and $D(n)$ fall into three cases (see Theorem 4.1) and one infinite class (Theorem 3.10). Furthermore, for the infinite class we can describe the underlying Sheffer intervals; see Theorem 3.11. For two of the three cases in Theorem 3.11 we can also classify the Sheffer intervals. However for the third case, we construct a multitude of examples of Sheffer intervals. It is a very striking coincidence that this case corresponds to the factorial functions of an infinite cubical lattice. That is, we can find many Sheffer posets having the same factorial functions as the infinite cubical lattice, but the Sheffer intervals are not isomorphic to the finite cubical lattice; see Examples 3.9, 4.2, 4.3 and 4.4. However, if we add the extra requirement that each Sheffer interval is a lattice then we obtain that the Sheffer intervals are isomorphic to cubical lattices.

When we impose the further condition that a given Eulerian binomial or Eulerian Sheffer poset is a lattice, this forces the poset to be the infinite Boolean algebra $\mathbb{B}_X$ or the infinite cubical lattice $\mathbb{C}_X^{<\infty}$. See Examples 2.10 and 4.6.

The classification of the factorial functions hinges on the condition that the posets under consideration contain an infinite chain. In the concluding remarks, we discuss what could happen if this condition is removed. We give examples of posets having their factorial functions behave like the face lattice of the dodecahedron, but themselves are not isomorphic to this lattice.

2. Eulerian binomial posets

DEFINITION 2.1. A locally finite poset P with $\hat{0}$ is called a *binomial poset* if it satisfies the following three conditions:

(i) P contains an infinite chain.
(ii) Every interval $[x, y]$ is graded; hence P has rank function ρ. If $\rho(x, y) = n$, then we call $[x, y]$ an *n-interval.*
(iii) For all $n \in \mathbb{N}$, any two n-intervals contain the same number $B(n)$ of maximal chains. We call $B(n)$ the *factorial function* of P.

If P does not satisfy condition (i) and has a unique maximal element then we say P is a *finite binomial poset.*

For standard poset terminology, we refer the reader to [**12**]. The number of elements of rank k in an n-interval is given by $B(n)/(B(k) \cdot B(n-k))$. Especially, an n-interval has $A(n) = B(n)/B(n-1)$ atoms (and coatoms). The function $A(n)$ is called the *atom function* and expresses the factorial function as $B(n) = A(n) \cdot A(n-1) \cdots A(1)$. Directly we have $B(0) = B(1) = A(1) = 1$. Since the atoms of an $(n-1)$-interval are contained among the set of atoms of an n-interval, the inequality $A(n-1) \leq A(n)$ holds. Observe if a finite binomial poset has rank j, the factorial and atom functions are only defined up to j. For further background material on binomial posets, see [**5, 11, 12**].

EXAMPLE 2.2. Let $\mathbb{B}$ be the collection of finite subsets of the positive integers ordered by inclusion. The poset $\mathbb{B}$ is a binomial poset with factorial function $B(n) = n!$ and atom function $A(n) = n$. An n-interval is isomorphic to the Boolean algebra B_n. This example is the *infinite Boolean algebra.*

EXAMPLE 2.3. Let $\mathbb{T}$ be the *infinite butterfly poset*, that is, $\mathbb{T}$ consists of the elements $\{\hat{0}\} \cup \mathbb{P} \times \{1, 2\}$ where $(n, i) \prec (n+1, j)$ for all $i, j \in \{1, 2\}$ and $\hat{0}$ is the unique minimal element. The poset $\mathbb{T}$ is a binomial poset. It has factorial function $B(n) = 2^{n-1}$ for $n \geq 1$ and atom function $A(n) = 2$ for $n \geq 2$. Let T_n denote an n-interval in $\mathbb{T}$.

EXAMPLE 2.4. Given two ranked posets P and Q, define the *rank product* $P * Q$ by
$$P * Q = \{(x, z) \in P \times Q \ : \ \rho_P(x) = \rho_Q(z)\}.$$
Define the order relation by $(x, y) \leq_{P*Q} (z, w)$ if $x \leq_P z$ and $y \leq_Q w$. If P and Q are binomial posets then so is the poset $P * Q$. It has the factorial function $B_{P*Q}(n) = B_P(n) \cdot B_Q(n)$. This example is due to Stanley [**12**, Example 3.15.3 d]. The rank product is also known as the Segre product; see [**4**].

EXAMPLE 2.5. For $q \geq 2$ let P_q be the face poset of an q-gon. Observe that this is a finite binomial poset of rank 3 with the factorial function $B(2) = 2$ and $B(3) = 2q$. Let $q_1, \ldots, q_r$ be a list of integers with each $q_i \geq 2$. Let $P_{q_1,\ldots,q_r}$ be the poset obtained by identifying all the minimal elements of P_{q_1} through P_{q_r} and identifying all the maximal elements. This is also a binomial poset with factorial function $B(2) = 2$ and $B(3) = 2 \cdot (q_1 + \cdots + q_r)$. It is straightforward to see that each rank 3 binomial poset with $B(2) = 2$ is of this form.

The *Euler-Poincaré relation* for a finite graded poset states that it has the same number of elements of even as odd rank. A poset is called *Eulerian* if every non-singleton interval satisfies the Euler-Poincaré relation. Equivalently, a poset P is Eulerian if its Möbius function satisfies $\mu(x, y) = (-1)^{\rho(y)-\rho(x)}$ for all $x \leq y$ in P.

LEMMA 2.6. *Let P be a graded poset of odd rank such that every proper interval of P is Eulerian. Then P is an Eulerian poset.*

This lemma is implicit in the two papers [**3, 7**]. We now conclude

PROPOSITION 2.7. *To verify that a poset is Eulerian it is enough to verify that every interval of even rank satisfies the Euler-Poincaré relation.*

For an n-interval of an Eulerian binomial poset the Euler-Poincaré relation states

$$\sum_{k=0}^{n} (-1)^k \cdot \frac{B(n)}{B(k) \cdot B(n-k)} = 0. \tag{2.1}$$

Note from Proposition 2.7 this relation will only give information when n is an even integer. Also observe that $B(2) = A(2) = 2$ since every 2-interval is a diamond.

Theorem 2.8. *Let P be an Eulerian binomial poset with factorial function $B(n)$. Then either*

(i) *the factorial function $B(n)$ is given by $B(n) = n!$ and every n-interval is isomorphic to the Boolean algebra B_n, or*

(ii) *the factorial function $B(n)$ is given by $B(n) = 2^{n-1}$ and every n-interval is isomorphic to the butterfly poset T_n.*

It is tempting to state this theorem as, "There are only two Eulerian binomial posets, namely, the infinite Boolean algebra $\mathbb{B}$ and the infinite butterfly poset $\mathbb{T}$." However, this is false. The next two examples demonstrate this.

Example 2.9. Let Q be an infinite poset with a minimal element $\hat{0}$ containing an infinite chain such that every interval of the form $[\hat{0}, x]$ is a chain. Observe the poset Q is an infinite tree and, in fact, is a binomial poset with factorial function $B(n) = 1$. Thus we know that both $\mathbb{B} * Q$ and $\mathbb{T} * Q$ are Eulerian binomial posets.

Example 2.10. For each infinite cardinal κ there is a Boolean algebra consisting of all finite subsets of a set X of cardinality κ. We denote this poset by $\mathbb{B}_X$. Observe that different cardinals give rise to non-isomorphic Boolean algebras.

We now state a very useful lemma.

Lemma 2.11. *Let P and P' be two Eulerian binomial posets having atom functions $A(n)$ and $A'(n)$ which agree for $n \leq 2m$, where $m \geq 2$. Then the following equality holds:*

$$\frac{1}{A(2m+1)} \cdot \left(1 - \frac{1}{A(2m+2)}\right) = \frac{1}{A'(2m+1)} \cdot \left(1 - \frac{1}{A'(2m+2)}\right). \tag{2.2}$$

We will use Lemma 2.11 in the following manner.

Corollary 2.12. *Let P and P' be two Eulerian binomial posets satisfying the conditions in Lemma 2.11. Assume furthermore there is a lower and an upper bound for $A'(2m+2)$ of the form $L \leq A'(2m+2) < U$. Let x be the left-hand side of equation (2.2). Then we obtain a lower and an upper bound for $A'(2m+1)$, namely*

$$\frac{1}{x} \cdot \left(1 - \frac{1}{L}\right) \leq A'(2m+1) < \frac{1}{x} \cdot \left(1 - \frac{1}{U}\right). \tag{2.3}$$

We will see these bounds can be improved by using that $A'(2m+1)$ is in fact an integer.

Proposition 2.13. *Let P' be an Eulerian binomial poset with factorial function $B'(n)$ satisfying $B'(3) = 6$. Then the factorial function is given by $B'(n) = n!$.*

Proof. Let P be the infinite Boolean algebra $\mathbb{B}$ with atom function $A(n) = n$ and factorial function $B(n) = n!$. We will first prove that the two factorial functions $B(n)$ and $B'(n)$ are identical, equivalently that the two atom functions $A(n)$ and $A'(n)$ are equal.

Assume that the two atom functions A and A' agree up to $2m = j$. Since $A(n) = n$ the left-hand side of equation (2.2) is equal to $1/(j+2)$. We have the following bounds for $A'(j+2)$: $j = A'(j) \leq A'(j+2) < \infty$. Applying Corollary 2.12 we obtain the following bounds on $A'(j+1)$:

$$j + 1 - \frac{2}{j} \leq A'(j+1) < j + 2.$$

Since $A'(j+1)$ is an integer and $j \geq 4$ we conclude that $A'(j+1) = j+1$. This implies that $A'(j+2) = j+2$ and hence we conclude the two atom functions are equal. $\square$

Proposition 2.14. *Let P be a finite binomial poset of rank n with factorial function $B(k) = k!$ for $k \leq n$. Then the poset P is isomorphic to the Boolean algebra B_n.*

Proof. Directly the result is true for $n \le 2$. Assume it is true for all posets of rank $n-1$ and consider a poset P of rank n. Since P is a binomial poset with factorial function $B(k) = k!$, we know that the number of elements of rank k in P is given by $\binom{n}{k}$. Especially, the cardinality of P is given by 2^n. Let c be a coatom in the poset. Observe that the interval $[\hat{0}, c]$ is isomorphic to B_{n-1} by the induction hypothesis and hence the coatom c is greater than all but one atom a in the poset P. Similarly, the interval $[a, \hat{1}]$ is also isomorphic to B_{n-1}. Since the two intervals $[a, \hat{1}]$ and $[\hat{0}, c]$ are disjoint and have the same cardinality 2^{n-1}, the poset P is the disjoint union of these two intervals.

Using the binomial property of P, an element z of rank k in the lower interval $[\hat{0}, c]$ is covered by $n-k$ elements in the poset P and by $n-k-1$ elements in the interval $[\hat{0}, c]$. Thus there is a unique element in $[a, \hat{1}]$ that covers z. Denote this element by $\varphi(z)$. By a similar argument we obtain that φ is a bijective function from $[\hat{0}, c]$ to $[a, \hat{1}]$. Let $z \prec w$ be a cover relation in $[\hat{0}, c]$. Consider the 2-interval $[z, \varphi(w)]$. As every 2-interval is a diamond there is an element v different from w such that $z \prec v \prec \varphi(w)$. Since w is the unique element in $[\hat{0}, c]$ that is covered by $\varphi(w)$, the element v belongs to the upper interval $[a, \hat{1}]$. Also, the element $\varphi(z)$ is the unique element in the upper interval that covers z, we conclude that $v = \varphi(z)$ and especially $\varphi(w)$ covers $\varphi(z)$. Hence the function φ is order-preserving. By the symmetric argument φ^{-1} is also order-preserving. Therefore the poset P is the Cartesian product of $[\hat{0}, c]$ with the two element poset B_1 and we conclude that P is isomorphic to the Boolean algebra B_n. □

Proposition 2.15. *Let P' be an Eulerian binomial poset with factorial function $B'(n)$ satisfying $B'(3) = 4$. Then the factorial function is given by $B'(n) = 2^{n-1}$ for $n \ge 1$.*

Proof. Let P be the butterfly poset $\mathbb{T}$ and $A(n)$ its atom function, where $A(1) = 1$ and $A(n) = 2$ for $n \ge 2$. Similar to the proof of Proposition 2.13 we consider how $A(n)$ and $A'(n)$ relate.

Assume that the two atom functions agree up to $2m = j$. Now the right-hand side of equation (2.2) is equal to $1/4$. For $A'(j+2)$ we have the bounds $2 = A'(j) \le A'(j+2) < \infty$. Applying Corollary 2.12 we obtain

$$2 \le A'(j+1) < 4.$$

Consider now the possibility that $A'(j+1) = 3$. Let $[x, y]$ be a $(j+1)$-interval in P'. For $1 \le k \le j$ there are $B'(j+1)/(B'(k) \cdot B'(j+1-k)) = 3 \cdot 2^{j-1}/(2^{k-1} \cdot 2^{j-k}) = 3$ elements of rank k in this interval. Let c be a coatom. The interval $[x, c]$ has two atoms, say a_1 and a_2. Moreover, the interval $[x, c]$ has two elements of rank 2, say b_1 and b_2. Moreover we know that each b_j covers each a_i. Let a_3 and b_3 be the third atom, respectively the third rank 2 element, in the interval $[x, y]$. We know that b_3 covers two atoms in $[x, y]$. One of them must be a_1 or a_2, say a_1. But then a_1 is covered by the three elements b_1, b_2 and b_3. But this contradicts the fact that each atom is covered by exactly two elements. Hence this rules out the case $A'(j+1) = 3$.

The only remaining possibility is $A'(j+1) = 2$, implying $A'(j+2) = 2$. Hence the atom functions $A(n)$ and $A'(n)$ are equal. □

Lemma 2.16. *Let P be a finite binomial poset with factorial function $B(k) = 2^{k-1}$ for $1 \le k \le n$. Then the poset P is isomorphic to the butterfly poset T_n.*

Proof of Theorem 2.8: The atom function of an Eulerian binomial poset satisfies $2 = A(2) \le A(3)$. Hence $B(3) = A(3) \cdot B(2)$ is an even integer greater than or equal to 4. The Euler-Poincaré relation implies that

$$\frac{1}{B(4)} = \frac{1}{B(3)} - \frac{1}{8},$$

implying that $B(3) < 8$. Hence there are only two remaining cases, which are considered in Propositions 2.13 and 2.15. The corresponding structure statements are considered in Proposition 2.14 and Lemma 2.16. □

Theorem 2.17. *Let L be an Eulerian binomial poset which we furthermore require to be a lattice. Then L is isomorphic to the Boolean algebra $\mathbb{B}_X$ where X is the set of atoms of the poset P.*

Proof. Since every interval of L is a lattice we can rule out the butterfly factorial function. Hence $B(n) = n!$ and every interval $[\hat{0}, x]$ is a Boolean lattice. Let φ be the map from L to $\mathbb{B}_X$ defined by $\varphi(x) = \{a \in X \;:\; a \le x\}$. The inverse of φ is given by $\varphi(Y) = \vee_{a \in Y} a$. It is straightforward to see that both φ and φ^{-1} are order-preserving. Hence the two lattices L and $\mathbb{B}_X$ are isomorphic and the result follows. □

We end this section with a result that will be used in Section 4 when we study Eulerian Sheffer posets.

PROPOSITION 2.18. *There is no finite binomial poset P' of rank $j+1 \geq 4$ with the atom function*

$$A'(n) = \begin{cases} n & \text{if } n \leq j, \\ j+2 & \text{if } n = j+1. \end{cases}$$

PROOF. Assume that the poset P' exists. Then it has $j+2$ atoms and $j+2$ coatoms. Each atom x lies below exactly j coatoms and each coatom c lies above exactly j atoms. Moreover, by the proof of Proposition 2.13 we know that each of the intervals $[\hat{0}, c]$ and $[x, \hat{1}]$ is isomorphic to B_j.

Define a multigraph G with the $j+2$ atoms as the vertices. For each coatom c let there be an edge xy between the two unique atoms x and y such that $x, y \not\leq c$. Since each atom is not below exactly two coatoms, each vertex of the graph has degree equal to 2. Hence the graph is a disjoint union of cycles.

Pick a coatom c that corresponds to an edge xy. The coatom c is greater than the j atoms $z_1, \ldots, z_j$. Using that the interval $[\hat{0}, c]$ is a Boolean algebra, let w_i be the unique coatom in the interval $[\hat{0}, c]$ that is not greater than z_i. Let d_i be the atom in the interval $[w_i, \hat{1}] \cong B_2$ distinct from c. Observe for $i \neq k$ we have $z_i < w_k < d_k$. Hence the j coatoms $c, d_1, \ldots, \widehat{d_i}, \ldots, d_j$ are all the coatoms greater than z_i. Moreover, since $j \geq 3$ we conclude that $d_1, \ldots, d_j$ are all distinct.

Consider the j atoms below d_k. They are $z_1, \ldots, \widehat{z_k}, \ldots, z_j$ and exactly one of x and y. Thus the edge e_k corresponding to d_k intersects the edge xy. This holds for all j edges e_k. Hence we obtain the contradiction $4 = \deg(x) + \deg(y) \geq 2 + j$. Thus there is no such finite binomial poset. □

3. Eulerian Sheffer posets

Sheffer posets, also know as upper binomial posets, were first defined by Reiner [**10**] and independently discovered by Ehrenborg and Readdy [**6**].

DEFINITION 3.1. A locally finite poset P with $\hat{0}$ is called a *Sheffer poset* if it satisfies the following four conditions:

(i) P contains an infinite chain.
(ii) Every interval $[x, y]$ is graded; hence P has rank function ρ. If $\rho(x, y) = n$, then we call $[x, y]$ an *n-interval*.
(iii) Two n-intervals $[\hat{0}, y]$ and $[\hat{0}, v]$, such that $y \neq \hat{0}$, $v \neq \hat{0}$, have the same number $D(n)$ of maximal chains.
(iv) Two n-intervals $[x, y]$ and $[u, v]$, such that $x \neq \hat{0}$, $u \neq \hat{0}$, have the same number $B(n)$ of maximal chains.

As in the finite binomial poset case, if P does not satisfy condition (i) and has a unique maximal element then we say P is a *finite Sheffer poset*.

An interval of the form $[\hat{0}, y]$ is called a *Sheffer interval*, whereas an interval $[x, y]$, where $x > \hat{0}$, is called a *binomial interval*. Similarly, the functions $B(n)$ and $D(n)$ are called the *binomial*, respectively, the *Sheffer factorial function*. The number of elements of rank $k \geq 1$ in a Sheffer interval of length n is given by $D(n)/(D(k) \cdot B(n-k))$. Especially, a Sheffer interval $[\hat{0}, y]$ has $C(n) = D(n)/D(n-1)$ coatoms. The function $C(n)$ is called the *coatom function* and we have $D(n) = C(n) \cdot C(n-1) \cdots C(1)$. Observe that $D(1) = C(1) = 1$.

We will be using the following two facts to exclude possible factorial functions.

FACT 3.2. (a) The inequality $A(n-1) \leq C(n) < \infty$ holds since the set of coatoms in a Sheffer interval of rank n, say $[\hat{0}, y]$, contains the set of coatoms in an $(n-1)$-interval $[x, y]$, and there are a finite number of them.
(b) The value $B(k)$ divides $C(n) \cdots C(n-k+1)$ for $n > k$, since the number of elements of rank $n-k$ in a Sheffer interval of length n is given by $D(n)/(D(n-k) \cdot B(k)) = C(n) \cdots C(n-k+1)/B(k)$.

EXAMPLE 3.3. Every binomial poset is a Sheffer poset. The factorial functions are equal, that is, $D(n) = B(n)$ for $n \geq 1$.

EXAMPLE 3.4. The rank product $P * Q$ of two Sheffer posets P and Q is also a Sheffer poset with the factorial functions $B_{P*Q}(n) = B_P(n) \cdot B_Q(n)$ and $D_{P*Q}(n) = D_P(n) \cdot D_Q(n)$.

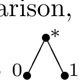
Example 3.5. For a poset P with a unique minimal element $\hat{0}$, let the *dual suspension* $\Sigma^*(P)$ be the poset P with two new elements a_1 and a_2. Let the order relations be as follows: $\hat{0} <_{\Sigma^*(P)} a_i <_{\Sigma^*(P)} y$ for all $y > \hat{0}$ in P and $i = 1, 2$. That is, the elements a_1 and a_2 are inserted between $\hat{0}$ and the atoms of P. Clearly if P is Eulerian then so is $\Sigma^*(P)$. Moreover, if P is a binomial poset then $\Sigma^*(P)$ is a Sheffer poset with the factorial function $D_{\Sigma^*(P)}(n) = 2 \cdot B(n-1)$ for $n \geq 2$.

One may extend the dual suspension Σ^* by inserting k new atoms instead of 2. Yet again it will take a binomial poset to a Sheffer poset. However we have no need of this extension since it does not preserve the Eulerian property.

For a ranked poset P (not necessarily having a unique minimal element) and a set X define the *power poset* P^X as follows. Let the underlying set be given by

$$P^X = \left\{ f : X \to P \ : \ \sum_{x \in X} \rho(f(x)) < \infty \right\}$$

and define the order relation by componentwise comparison, that is, $f \leq_{P^X} g$ if $f(x) \leq g(x)$ for all x in X.

Example 3.6. Let P be the three element poset 

and let X be an infinite set. Then the poset $\mathbb{C}_X = P^X \cup \{\hat{0}\}$, that is, the poset P^X with a new minimal element adjoined, is a Sheffer poset. This example is precisely the *infinite cubical poset* with the factorial functions $B(n) = n!$ and $D(n) = 2^{n-1} \cdot (n-1)!$. Similar to Example 2.10, for different infinite cardinalities of X we obtain non-isomorphic cubical posets. Note, however, this poset is not a lattice since the two atoms $(0, 0, \ldots)$ and $(1, 1, \ldots)$ do not have a join.

Example 3.7. Let $E_2, E_3, \ldots$ be an infinite sequence of disjoint nonempty finite sets, where E_n has cardinality e_n. Consider the poset

$$U_{e_2, e_3, \ldots} = \{\hat{0}\} \ \cup \ \bigcup_{n \geq 2} \prod_{i \geq n} E_i,$$

where $\prod$ stands for Cartesian product. We make this into a ranked poset by letting $\hat{0}$ be the minimal element, and defining the cover relation by

$$(x_n, x_{n+1}, x_{n+2}, \ldots) \prec (x_{n+1}, x_{n+2}, \ldots),$$

where $x_i \in E_i$. Thus the elements of $\prod_{i \geq n} E_i$ have rank $n-1$. This poset is a Sheffer poset with the atom function $A(n) = 1$ and coatom function is given by $C(n) = e_n$ for $n \geq 2$. We may view this poset as an "upside-down tree" with a minimal element attached.

Naturally, the previous example is not an Eulerian poset. However, we can use it to construct Eulerian Sheffer posets as the next two examples illustrate.

Example 3.8. Consider the poset $\mathbb{T} * U_{e_2, e_3, \ldots}$, where $e_2 = e_4 = e_6 = \cdots = 1$. This poset has the factorial functions $B(n) = 2^{n-1}$ and $D(n) = 2^{n-1} \cdot \prod_{i=2}^{n} e_i$. In Theorem 3.10 we will observe that the condition that $e_{2j} = 1$ implies that the poset is Eulerian.

In general the rank product $\mathbb{T} * P$ can be viewed as the "doubling" of the poset P. This notion was introduced by Bayer and Hetyei in [**2**].

Example 3.9. Let $\mathbb{B} \cup \{\hat{0}\}$ be the infinite Boolean algebra with a new minimal element adjoined. This is a Sheffer poset with factorial functions $B(n) = n!$ and $D(n) = (n-1)!$. Now consider the rank product $(\mathbb{B} \cup \{\hat{0}\}) * U_{2,2,\ldots}$. It has the factorial functions $B(n) = n!$ and $D(n) = 2^{n-1} \cdot (n-1)!$. This poset has the same factorial functions as the infinite cubical poset and hence it is an Eulerian poset.

For an Eulerian Sheffer poset of rank n, the Euler-Poincaré relation states

$$1 + \sum_{k=1}^{n} (-1)^k \cdot \frac{D(n)}{D(k) \cdot B(n-k)} = 0. \tag{3.1}$$

Again by Proposition 2.7 this relation will only give us information for n even. When $n = 2m$ we can write this relation as:

$$\frac{2}{D(2m)} + \sum_{k=1}^{2m-1} (-1)^k \cdot \frac{1}{D(k) \cdot B(2m-k)} = 0. \tag{3.2}$$

Also note that $D(2) = C(2) = 2$.

Theorem 3.10. *Let P be an Eulerian Sheffer poset with the binomial factorial function satisfying $B(0) = 1$ and $B(n) = 2^{n-1}$ for $n \geq 1$. Then the coatom function $C(n)$ and the poset P satisfy:*

(i) *$C(3) \geq 2$, and a length 3 Sheffer interval is isomorphic to a poset of the form $P_{q_1,\ldots,q_r}$ described in Example 2.5.*

(ii) *$C(2m) = 2$ for $m \geq 2$ and the two coatoms in a length $2m$ Sheffer interval cover exactly the same elements of rank $2m-2$.*

(iii) *$C(2m+1) = h$ is an even positive integer, for $m \geq 2$. Moreover, the set of h coatoms in a Sheffer interval of length $2m+1$ can be grouped into $h/2$ pairs, $\{c_1, d_1\}$, $\{c_2, d_2\}$, $\ldots$, $\{c_{h/2}, d_{h/2}\}$, such that c_i and d_i cover the same two elements of rank $2m-1$.*

Proof. Part (i) is immediate since $A(2) \leq C(3)$. Next we prove (ii). Let $j = 2m$. In this case the Euler-Poincaré relation for a Sheffer j-interval states:

$$\sum_{k=1}^{j} (-1)^k \cdot \frac{1}{D(k) \cdot 2^{j-k-1}} = 0. \tag{3.3}$$

Use equation (3.3) in the case of a $(j-2)$-interval to eliminate the first $j-2$ terms in the j-interval case of (3.3) gives the equality (ii). Since $D(j)/(D(j-2) \cdot B(2)) = D(j-1)/(D(j-2) \cdot B(1))$, the two coatoms in the Sheffer j-interval cover the same elements of rank $j-2$.

Finally, we consider (iii). Assume that $C(j+1) = h$, where $j = 2m$. Let $[\hat{0}, y]$ be a Sheffer interval of rank $j+1$. The number of elements of rank j and of rank $j-1$ are both given by h. Moreover each element of rank $j-1$ is covered by exactly 2 elements of rank j, and by part (ii), each element of rank j covers 2 elements of rank $j-1$. Hence the order relations between elements of rank $j-1$ and j are those of rank 1 and 2 in the poset $P_{q_1,\ldots,q_r}$ in Example 2.5, where $q_1 + \cdots + q_r = h$.

Let $z_1, \ldots, z_q$ be q coatoms in the Sheffer $(j+1)$-interval $[\hat{0}, y]$ such that z_i covers w_i and w_{i-1}, where we count modulo q in the indices. That is, z_1 through z_q correspond to the edges in a q-gon and w_1 through w_q to the vertices. Consider an element x of rank $j-2$ that is covered by w_1. The interval $[x, y]$ is isomorphic to T_3, that is, the interval has exactly 2 atoms and 2 coatoms. In this interval the element x is covered by one more element of rank $j-1$. Call it v. If the element v does not correspond to the elements $w_2, \ldots, w_q$, we obtain the contradiction that the interval $[x, y]$ has 4 coatoms. If v belongs to the elements $w_2, \ldots, w_q$, say w_i, then the interval $[x, y]$ has the coatoms z_1, z_2, z_i, z_{i+1}. When $q \geq 3$ the set $\{z_1, z_2, z_i, z_{i+1}\}$ has at least 3 members. Hence the only possibility is that $q = 2$ and $v = w_2$. Also the coatoms z_1 and z_2 cover the same elements of rank $j-1$.

We conclude that the only possibility is that all q_i's are equal to 2, that is, $q_1 = \cdots = q_r = 2$. Hence $r = h/2$ and h is an even integer. Moreover, we also obtain a pairing of the coatoms such that the two coatoms in each pair cover the same elements. ☐

Given a graded poset P of rank n and a subset $S \subseteq \{1, \ldots, n-1\}$, the *rank selected poset* P_S is the graded poset consisting of the elements

$$P_S = \{\hat{0}, \hat{1}\} \cup \{x \in P \ : \ \rho(x) \in S\}.$$

Combining the conclusions of Theorem 3.10, we have

Theorem 3.11. *Let P be an Eulerian Sheffer poset with the binomial factorial function satisfying $B(0) = 1$ and $B(n) = 2^{n-1}$ for $n \geq 1$ and coatom function $C(n)$. Set $e_2 = e_4 = e_6 = \cdots = 1$ and $e_{2m+1} = C(2m+1)/2$ for all $m \geq 1$. Let Q be the poset $\mathbb{T} * U_{e_2,e_3,\ldots}$ from Example 3.8. Suppose n is an integer greater than or equal to 3 and $S = \{3, 4, \ldots, n-1\}$. Then the rank selection S of the rank n Sheffer interval $[\hat{0}, y]$ in P is isomorphic to the rank selection S of the rank n Sheffer interval $[\hat{0}, z]$ in Q, that is,*

$$[\hat{0}, y]_S \cong [\hat{0}, z]_S.$$

Furthermore, the poset $[\hat{0}, y]$ is obtained by replacing every length 3 Sheffer interval in $[\hat{0}, z]$ by a rank 3 binomial poset with $C(3)$ coatoms.

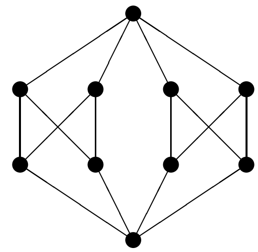


Figure 1. A finite Sheffer poset with the same factorial functions as the cubical lattice.

4. Eulerian Sheffer posets with factorial function $B(n) = n!$

In this section we will classify Eulerian Sheffer posets that have the factorial function $B(n) = n!$, that is, every interval $[x, y]$, where $x > \hat{0}$, is a Boolean algebra.

Theorem 4.1. *Let P be an Eulerian Sheffer poset with binomial factorial function $B(n) = n!$. Then the Sheffer factorial function $D(n)$ satisfies one of the following three alternatives:*

(i) $D(n) = 2 \cdot (n-1)!$. *In this case every Sheffer n-interval is of the form $\Sigma^*(B_{n-1})$.*
(ii) $D(n) = n!$. *In this case the poset is a binomial poset and hence every Sheffer n-interval is isomorphic to the Boolean algebra B_n.*
(iii) $D(n) = 2^{n-1} \cdot (n-1)!$. *If we furthermore assume that a Sheffer n-interval $[\hat{0}, y]$ is a lattice then the interval $[\hat{0}, y]$ is isomorphic to the cubical lattice C_n.*

The cubical posets of Example 3.6 and Example 3.9 demonstrate there is no classification of the non-lattice Sheffer intervals in case (iii) of Theorem 4.1. The following examples further illustrates Sheffer posets (both finite and infinite) having the same factorial functions as the cubical poset.

Example 4.2. Let C_n be the finite cubical lattice, that is, the face lattice of an $(n-1)$-dimensional cube. We are going to deform this lattice as follows. The 1-skeleton of the cube is a bipartite graph. Hence the set of atoms A has a natural decomposition as $A_1 \cup A_2$. Every rank 2 element (edge) covers exactly one atom in each A_i. Consider the poset

$$H_n = (C_n - A) \cup A_1 \times \{1, 2\}.$$

That is, we remove all the atoms and add in two copies of each atom from A_1. Define the cover relations for the new elements as follows. If a in A_1 is covered by b then let b cover both copies $(a, 1)$ and $(a, 2)$. The poset H_n is a Sheffer poset with the cubical factorial functions.

The poset in Figure 1 is the atom deformed cubical lattice H_3. This poset is also obtained as length 3 Sheffer interval in Example 3.9.

Example 4.3. Let P and Q be two Sheffer posets (finite or infinite) having the cubical factorial functions $B(n) = n!$ and $D(n) = 2^{n-1} \cdot (n-1)!$. Their diamond product, namely $P \diamond Q = (P - \{\hat{0}\}) \times (Q - \{\hat{0}\}) \cup \{\hat{0}\}$, also has the cubical factorial functions.

Example 4.4. As an extension of the previous example, let P be a Sheffer poset (finite or infinite) having the cubical factorial functions. Then for a set X the poset $(P - \{\hat{0}\})^X \cup \{\hat{0}\}$ is a Sheffer poset with the cubical factorial functions. The cubical poset (Example 3.6) is an illustration of this.

If we require the extra condition every Sheffer interval is a lattice, we obtain it is in fact the cubical lattice.

Proposition 4.5. *Let P be a finite Sheffer poset of rank n with the cubical factorial functions $B(k) = k!$ for $k \leq n-1$ and $D(k) = 2^{k-1} \cdot (k-1)!$ for $1 \leq k \leq n$. If P is a lattice then P is isomorphic to the cubical lattice C_n.*

Proof. The proof is by induction on the rank n of P. The induction base $n \leq 2$ is straightforward to verify. Assume true for all posets of rank $n-1$ and consider a rank n poset P. Using the cubical factorial functions, we know that the half open interval $(\hat{0}, \hat{1}]$ contains 3^{n-1} elements. Let c be a coatom in the poset. The interval $[\hat{0}, c]$ is isomorphic to C_{n-1} by the induction hypothesis. Now define a function $\varphi : (\hat{0}, c] \longrightarrow (\hat{0}, \hat{1}] - (\hat{0}, c]$ as follows. For z in $(\hat{0}, c]$ let $\varphi(z)$ be the unique atom in the interval $[z, \hat{1}]$ that does not belong to the interval $[z, c]$. The existence and uniqueness follows from the fact the atom function satisfies $A(k) - A(k-1) = 1$. Also note that $\varphi(z)$ covers the element z.

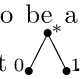
We next verify the function φ is injective. If we have $\varphi(z) = \varphi(w)$ then z and w have the same rank. Also observe that $\varphi(z) \not\leq c$ by the definition of the function φ. This contradicts that the interval $[\hat{0}, \hat{1}]$ is a lattice, since z and w have the two upper bounds $\varphi(z)$ and c.

The function φ also preserves the cover relations. If $z \prec w$ the two-interval $[z, \varphi(w)]$ contains two atoms which must be w and $\varphi(z)$. Hence $\varphi(z) \prec \varphi(w)$. Let Φ be the image of the function φ. By a similar argument the inverse function $\varphi^{-1} : \Phi \longrightarrow (\hat{0}, c]$ also preserves the cover relations. Thus as posets $(\hat{0}, c]$ and Φ are isomorphic. Moreover, the disjoint union $(\hat{0}, c] \cup \Phi$ is an upper order ideal of the poset P and has cardinality $2 \cdot 3^{n-2}$.

The poset P has $C(n) = 2n - 2$ coatoms. One of them is the coatom c. Since c covers $2n - 4$ elements there are $2n - 4$ coatoms in Φ. Hence there is a unique coatom d that does not belong to the upper order ideal $(\hat{0}, c] \cup \Phi$. Since the interval $[\hat{0}, d]$ is isomorphic to the cubical lattice C_{n-1} and has $3^{n-2} + 1$ elements, we conclude that the complement of the upper order ideal is the lower order ideal $[\hat{0}, d]$. Thus we have the partition $(\hat{0}, c] \cup \Phi \cup (\hat{0}, d]$ of $P - \{\hat{0}\}$.

It remains to show that there is a bijective function $\psi : (\hat{0}, d] \longrightarrow \Phi$ such that $\psi(z)$ covers z and ψ preserves the cover relation. Define $\psi : (\hat{0}, d] \longrightarrow (\hat{0}, y] - (\hat{0}, d] = (\hat{0}, c] \cup \Phi$ by letting $\psi(z)$ be the unique atom in the interval $[z, \hat{1}]$ that does not belong to the interval $[z, d]$. Observe that if $\psi(z) \in (\hat{0}, c]$ we obtain that $z < \psi(z) \leq c$, contradicting that $(\hat{0}, c]$ and $(\hat{0}, d]$ are disjoint. Hence the image of ψ is Φ. The remaining properties of ψ are proven just like those for the function φ.

Hence $P - \{\hat{0}\}$ is isomorphic to the Cartesian product of the three element poset  with $(\hat{0}, c] \cong C_{n-1}$. That is, the poset is isomorphic to the cubical lattice C_n. □

Example 4.6. Define $\mathbb{C}_X^{<\infty}$ to be a subposet of the cubical poset $\mathbb{C}_X = P^X \cup \{\hat{0}\}$ in Example 3.6, where P is the three element poset 

Define

$$\mathbb{C}_X^{<\infty} = \{f \in P^X \ : \ |f^{-1}(1)| < \infty\} \cup \{\hat{0}\}.$$

That is, for each function f only a finite number of elements of X take on non-zero values. Since the union of two finite sets is finite it follows that the join of the two elements is defined. It follows that $\mathbb{C}_X^{<\infty}$ is a lattice. Observe the subposet $\mathbb{C}_X^{<\infty}$ remains a Sheffer poset with the cubical factorial functions $B(n) = n!$ and $D(n) = 2^{n-1} \cdot (n-1)!$. Call this poset the *infinite cubical lattice*.

Theorem 4.7. *Let L be an Eulerian Sheffer poset that is also a lattice. Then L is either isomorphic to $\mathbb{B}_X$ where X is the set of atoms of L or L is the infinite cubical lattice $\mathbb{C}_X^{<\infty}$ where X is the set of rank 2 elements of L which are greater than some fixed atom a in L.*

Proof. Using Theorem 2.17 we know that the binomial factorial function is $B(n) = n!$. Since every Sheffer interval is a lattice there are only two choices for the Sheffer factorial function. The case $D(n) = n!$ is indeed the Boolean algebra which is the first alternative of the conclusion of the theorem. Hence let us consider the second choice $D(n) = 2^{n-1} \cdot (n-1)!$. Thus every interval $[\hat{0}, y]$ is a finite cubical lattice.

Let a be an atom of the lattice L and let X be the set of elements of rank 2 which cover a. Define the function $\varphi : L \longrightarrow \mathbb{C}_X^{<\infty}$ as follows. Set $\varphi(\hat{0}) = \hat{0}$. For $x \in L$ and $x > \hat{0}$ let y be the join of a and x. Since the interval $[\hat{0}, y]$ is a finite cubical lattice, the non-minimal elements of this interval can be encoded by functions $g : Y \longrightarrow P$, where is P is the three element poset in Example 4.6. Furthermore we may assume that the set Y is all the elements in the interval $[a, y]$ that cover a. Without loss of generality, we may choose the encoding so that the atom a is the constant function 0.

Encode the element x as such a function $g : Y \longrightarrow P$. Observe that g does not take the value 0, since that would contradict that the join of a and x is y. Now define $f : X \longrightarrow P$ by

$$f(z) = \begin{cases} g(z) & \text{if } z \in Y, \\ 0 & \text{if } z \in X - Y. \end{cases}$$

Observe that since Y is a finite set, we know that f belongs to the lattice $\mathbb{C}_X^{<\infty}$. Hence set $\varphi(x)$ to be the function f.

The inverse of φ is given as follows. For f, a non-zero element of the lattice $\mathbb{C}_X^{<\infty}$ let the set Y be defined as

$$Y = \{z \in X \ : \ f(z) \neq 0\}.$$

In the lattice L let the element y be the join $\bigvee_{z \in Y} z$. Observe that $a \leq y$. Since the interval $[\hat{0}, y]$ is isomorphic to the finite cubical lattice $\mathbb{C}_Y$, let x be the unique element corresponding to the function f restricted to Y. That is, the inverse of φ is given by $\varphi^{-1}(f) = x$. Moreover let $\varphi^{-1}(\hat{0}) = \hat{0}$.

Observe that both φ and φ^{-1} are order preserving, thus proving that the lattices L and $\mathbb{C}_X^{<\infty}$ are isomorphic. □

Note that it is enough to work with the join operation in this proof, since a locally finite join semi-lattice with unique minimal element is a lattice [**12**, Proposition 3.3.1].

5. Concluding remarks

An interesting research project is to classify the factorial functions of finite Eulerian binomial posets and finite Eulerian Sheffer posets. Two examples of finite Sheffer posets are the face lattices of the dodecahedron and the four-dimensional regular polytope known as the 120-cell. In Theorems 3.10 and 4.1 many finite possibilities for the factorial functions were excluded since there was no possibility to extend the factorial function to higher ranks. A first step in this classification is to consider these cases.

Also note the following lemma, the proof of which follows directly from Proposition 2.7.

Lemma 5.1. *Let P be an Eulerian finite binomial (Sheffer) poset of odd rank n. Let Q be the poset obtain by taking k disjoint copies of P and identifying the minimal, respectively, maximal elements. Then Q is an Eulerian finite binomial (Sheffer) poset. The only value of the factorial function(s) that changes is the one that enumerates the maximal chains, namely, $B_Q(n) = k \cdot B_P(n)$ in the binomial case, and $D_Q(n) = k \cdot D_P(n)$ in the Sheffer case.*

A larger class of posets to consider are the triangular posets [**5**]. A poset is *triangular* if every interval $[x, y]$, where x has rank n and y has rank m, has $B(n, m)$ maximal chains. Both binomial and Sheffer posets are triangular. A non-trivial Eulerian example of a finite triangular poset is the face lattice of the 4-dimensional regular polytope known as the 24-cell. Can the factorial function $B(n, m)$ be classified for Eulerian triangular posets?

Classifying finite Eulerian Sheffer posets only by their factorial functions seems to be hard as seen from the multitude of examples having the cubical factorial functions. We leave the reader with three examples of Sheffer posets with the same factorial functions as the face lattice of the dodecahedron, each of which is not isomorphic to this face lattice.

Example 5.2. *An Eulerian finite Sheffer poset with the same factorial functions as the face lattice of the dodecahedron.* For an n-gon define a CW-complex X_n as follows. First take the antiprism of the n-gon. We then have a CW-complex consisting of two n-gons and $2n$ triangles. Note that at every vertex three triangles and one n-gon meet. Now subdivide each of the two n-gons by placing a vertex in each n-gon and attaching this vertex by n new edges to the n vertices of the n-gon. Let this be the CW-complex X_n.

Observe that X_n consists of $2n + 2$ vertices, $6n$ edges and $4n$ triangles. Moreover, at $2n$ of the vertices 5 triangles meet. At the other two vertices n triangles meet. Label these two vertices a and b. Also note that X_5 is the boundary complex of an icosahedron. Observe for $n \geq 3$ that X_n is a simplicial complex. However, for $n = 2$ it is necessary to view X_2 as a CW-complex.

Construct a CW-complex Y by taking X_2 and X_3 and identifying the vertices labeled a and identifying the vertices labeled b. See Figure 2. The dual of the face poset of Y is an Eulerian Sheffer poset with factorial functions agreeing with the face lattice of a dodecahedron.

Example 5.3. For $1 \leq i \leq 3$ let Z_i be the boundary of a 3-dimensional simplex with vertices $z_{i,1}$, $z_{i,2}$, $z_{i,3}$ and $z_{i,4}$. Similarly, for $1 \leq j \leq 4$ let W_j be the spherical CW-complex consisting of two triangles sharing the three edges. Call the vertices $w_{1,j}$, $w_{2,j}$ and $w_{3,j}$. Now identify vertex $z_{i,j}$ with $w_{i,j}$. We then have a CW-complex that has 12 vertices, $3 \cdot 6 + 4 \cdot 3 = 30$ edges and $3 \cdot 4 + 4 \cdot 2 = 20$ triangles. Observe that the vertex figure of every vertex is the disjoint union of a 2-gon and a triangle. Thus the dual of the face poset is Sheffer poset with the same factorial functions as the face lattice of a dodecahedron. In fact, one may obtain several of these CW-complexes by choosing different identifications between the two classes of vertices.

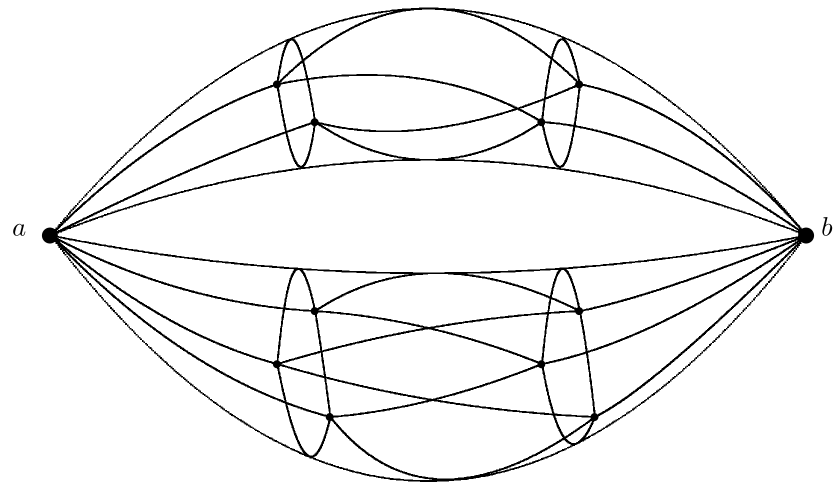


Figure 2. The CW-complex obtained by joining the complexes X_2 and X_3 at the vertices a and b.

Example 5.4. A third example is formed by taking two X_2's from Example 5.2 and the boundary of one 3-dimensional simplex, Z, from Example 5.3 and identifying vertices a_1, a_2, b_1 and b_2 with the vertices of the simplex.

A different proof of Proposition 2.14 may be given using the following result of Stanley. A graded finite poset P is a Boolean algebra if every 3-interval is a Boolean algebra and for every interval $[x, y]$ of rank of least 4 the open interval (x, y) is connected. See [**9**, Lemma 8]. Hence it is natural to ask if one can extend this result to cubical lattices. That is, a graded finite poset P is a cubical lattice if every 3-interval $[x, y]$, where $x > \hat{0}$, is a Boolean algebra, every 3-interval $[\hat{0}, y]$ is the face lattice of a square, and for every interval $[x, y]$ of rank of least 4 the open interval (x, y) is connected.

One may drop the Eulerian condition and ask to characterize Sheffer posets which are lattices. The lattice-theoretic techniques of Farley and Schmidt may be useful [**8**].

Finally, there are long-standing open questions regarding binomial posets. One such question was whether there exist two binomial posets having the same factorial function but non-isomorphic intervals. This question was very recently settled by Jörgen Backelin (personal communication). However, it is still unknown if there is a binomial poset having the atom function $A(n) = F_n$, the nth Fibonacci number. See Exercise 78b, Chapter 3 in [**12**].

References

[1] J. Backelin, Binomial posets with non-isomorphic intervals, **math.CO/0508397**, 22 August 2005.

[2] M. M. Bayer and G. Hetyei, Flag vectors of Eulerian partially ordered sets, *European J. Combin.* **22** (2001), 5–26.

[3] L. J. Billera and N. Liu, Noncommutative enumeration in graded posets, *J. Algebraic Combin.* **12** (2000), 7–24.

[4] A. Bjorner and V. Welker, Segre and Rees products of posets, with ring-theoretic applications, to appear in *J. Pure Appl. Algebra*.

[5] P. Doubilet, G.-C. Rota and R. Stanley, *On the foundation of combinatorial theory (VI). The idea of generating functions,* in Sixth Berkeley Symp. on Math. Stat. and Prob., vol. 2: Probability Theory, Univ. of California (1972), pp. 267–318.

[6] R. Ehrenborg and M. Readdy, Sheffer posets and r-signed permutations, *Annales des Sciences Mathématiques du Québec* **19** (1995), 173–196.

[7] R. Ehrenborg and M. Readdy, Homology of Newtonian coalgebras, *European J. Combin.* **23** (2002), 919–927.

[8] J. Farley and S. Schmidt, Posets that locally resemble distributive lattices, *J. Combin. Theory Ser. A* **92** (2000), 119–137.

[9] D. J. Grabiner, Posets in which every interval is a product of chains, and natural local actions of the symmetric group, *Discrete Math.* **199** (1999), 77–84.

[10] V. Reiner, Upper binomial posets and signed permutation statistics, *European J. Combin.* **14** (1993), 581–588.

[11] R. Stanley, Binomial posets, Möbius inversion, and permutation enumeration, *J. Combin. Theory Ser. A* **20** (1976), 336–356.

[12] R. Stanley, “Enumerative Combinatorics, Vol. I,” Wadsworth and Brooks/Cole, Pacific Grove, 1986.

Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA
E-mail address: jrge@ms.uky.edu
URL: http://www.ms.uky.edu/~jrge

Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA
E-mail address: readdy@ms.uky.edu
URL: http://www.ms.uky.edu/~readdy

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Bounds on the number of inference functions of a graphical model

Sergi Elizalde and Kevin Woods

Abstract. We give an upper bound on the number of inference functions of any directed graphical model. This bound is polynomial on the size of the model, for a fixed number of parameters, thus improving the exponential upper bound given in [Pachter and Sturmfels, Tropical Geometry of Statistical Models, *Proc. Natl. Acad. Sci.* 101, n. 46 (2004), 16132–16137]. Our proof reduces the problem to the enumeration of vertices of a Minkowski sum of polytopes. We also show that our bound is tight up to a constant factor, by constructing a family of hidden Markov models whose number of inference functions agrees asymptotically with the upper bound. Finally, we apply this bound to a model for sequence alignment that is used in computational biology.

Résumé. Nous donnons une limite supérieure sur le nombre de fonctions d'inférence de tout modéle graphique dirigé. Cette limite est polynômielle sur la grosseur du modèle, pour un nombre fixe de paramètres, améliorant ainsi la limite supérieure exponentielle donnée dans [Pachter and Sturmfels, Tropical Geometry of Statistical Models, *Proc. Natl. Acad. Sci.* 101, n. 46 (2004), 16132–16137]. Notre preuve réduit le problème à l'énumération de sommets d'une somme de Minkowski de polytopes. Nous montrons aussi que notre limite est serrée jusqu'à un facteur constant, en construisant une famille de modèles de Markov cachés dont le nombre de fonctions d'inférence coïncide asymptotiquement avec la limite supérieure. Finalement, nous appliquons cette limite à un modèle pour l'alignement de séquences qui est utilisé dans la biologie computationnelle.

1. Introduction

Many statistical models seek, given a set of observed data, to find the *hidden* (unobserved) data which best explains these observations. In this paper we consider graphical models, also called Bayesian networks, a broad class that includes many useful models, such as hidden Markov models (HMMs), pair hidden Markov models, and hidden tree models (background on graphical models will be given in Section 2.1). These graphical models relate the hidden and observed data probabilistically, and a natural problem is to determine, given a particular observation, what is the most likely hidden data (which is called the *explanation*). These models rely on parameters that are the probabilities relating the hidden and observed data. Any fixed values of the parameters determine a way to assign an explanation to each possible observation. This gives us a map, called an *inference function*, from observations to explanations.

An example of an inference function is the popular "*Did you mean*" feature from `google`, which could be implemented as a hidden Markov model, where the observed data is what we type into the computer, and the hidden data is what we were meaning to type. Graphical models are frequently used in these sorts of probabilistic approaches to artificial intelligence (see [**5**] for an introduction).

Inference functions for graphical models are also important in computational biology [**6**, Section 1.5]. For example, consider the *gene-finding functions*, which were discussed in [**7**, Section 5]. These inference functions (corresponding to a particular HMM) are used to identify gene structures in DNA sequences. An observation in such a model is a sequence of nucleotides in the alphabet $\Sigma' = \{\mathtt{A}, \mathtt{C}, \mathtt{G}, \mathtt{T}\}$, and an explanation is a sequence of 1's and 0's which indicate whether the particular nucleotide is in a gene or is not. We seek to use the information in the observed data (which we can find via DNA sequencing) to decide on the hidden

2000 *Mathematics Subject Classification.* Primary 62F15,52C45; Secondary 52B20,62P10,52B05.

Key words and phrases. inference functions, graphical models, sequence alignment, Newton polytope, normal fan.

information of which nucleotides are part of genes (which is hard to figure out directly). Another class of examples is that of sequence alignment models [**6**, Section 2.2]. In such models, an inference function is a map from a pair of DNA sequences to an optimal alignment of those sequences. If we change the parameters of the model, which alignments are optimal may change, and so the inference functions may change.

A surprising conclusion of this paper is that there cannot be *too many* different inference functions, though the parameters may vary continuously over all possible choices. For example, in the homogeneous binary HMM of length 5 (see Section 2.1 for some definitions; they are not important at the moment), the observed data is a binary sequence of length 5, and the explanation will also be a binary sequence of length 5. At first glance, there are

$$32^{32} = 1\ 461\ 501\ 637\ 330\ 902\ 918\ 203\ 684\ 832\ 716\ 283\ 019\ 655\ 932\ 542\ 976$$

possible maps from observed sequences to explanations. In fact, Christophe Weibel has computed that only 5266 of these possible maps are actually inference functions [**9**].

Different inference functions represent different criteria to decide what is the most likely explanation for each observation. A bound on the number of inference functions is important because it indicates how badly a model may respond to changes in the parameter values (which are generally known with very little certainty and only guessed at). Also, the polynomial bound given in Section 3 suggests that it might be feasible to precompute all the inference functions of a given graphical model, which would yield an efficient way to provide an explanation for each given observation.

This paper is structured as follows. In Section 2 we introduce some preliminaries about graphical models and inference functions, as well as some facts about polytopes. In Section 3 we present our main result. We call it the *Few Inference Functions Theorem*, and it states that in any graphical model the number of inference functions grows polynomially in the size of the model (if the number of parameters is fixed). The proof involves combinatorial tools, and a key step consists in reducing the enumeration of inference functions to the problem of counting the number of vertices of a certain polytope that is obtained as a Minkowski sum of smaller polytopes. In Section 4 we prove that our upper bound on the number of inference functions of a graphical model is sharp, up to a constant factor, by constructing a family of HMMs whose number of inference functions asymptotically matches the bound. In Section 5 we show that the bound is also asymptotically tight on a model for sequence alignment which is actually used in computational biology. In particular, this bound will be quadratic on the length of the input DNA sequences. We conclude with a few remarks and possible directions for further research.

2. Preliminaries

2.1. Graphical models. A *statistical model* is a family of joint probability distributions for a collection of discrete random variables $\mathbf{Z} = (Z_1, \ldots, Z_m)$, where each Z_i takes on values in some finite state space Σ_i. Here we will focus on directed graphical models. A *directed graphical model* (or *Bayesian network*) is a finite directed acyclic graph G where each vertex v_i corresponds to a random variable Z_i. Each vertex v_i also has an associated probability map

$$p_i : \left(\prod_{j:\ v_j \text{ a parent of } v_i} \Sigma_j \right) \to [0,1]^{|\Sigma_i|}.$$

Given the states of each Z_j such that v_j is a parent of v_i, the probability that v_i has a given state is independent of the values of all other vertices that are not descendants of v_i, and this map p_i gives that probability. In particular, we have the equality

$$\mathrm{Prob}(\mathbf{Z} = \tau) = \prod_i \mathrm{Prob}\left(Z_i = \tau_i, \text{ given that } Z_j = \tau_j \text{ for all parents } v_j \text{ of } v_i\right) = \prod_i \left[p_i\left(\tau_{j_1}, \ldots, \tau_{j_k}\right)\right]_{\tau_i},$$

where $v_{j_i}, \ldots, v_{j_k}$ are the parents of v_i. Sources in the digraph (which have no parents) are generally given the uniform probability distribution on their states, though more general distributions are possible. See [**6**, Section 1.5] for general background on graphical models.

Example 2.1. The hidden Markov model (HMM) is a model with random variables $\mathbf{X} = (X_1, \ldots, X_n)$ and $\mathbf{Y} = (Y_1, \ldots, Y_n)$. Edges go from X_i to X_{i+1} and from X_i to Y_i.

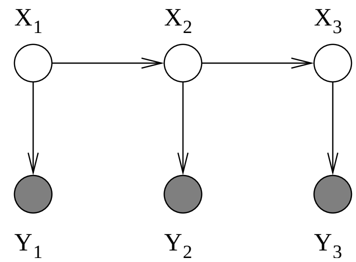


Figure 1. The graph of an HMM for $n = 3$.

Generally, each X_i has the same state space Σ and each Y_i has the same state space Σ'. An HMM is called *homogeneous* if the p_{X_i}, for $1 \leq i \leq n$ are identical and the p_{Y_i} are also identical. In this case, the p_{X_i} each correspond to the same $|\Sigma| \times |\Sigma|$ matrix $T = (t_{ij})$ (the *transition matrix*) and the p_{Y_i} each correspond to the same $|\Sigma| \times |\Sigma'|$ matrix $S = (s_{ij})$ (the *emission* matrix).

In the example, we have partitioned the variables into two sets. In general graphical models, we also have two kinds of variables: observed variables $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ and hidden variables $\mathbf{X} = (X_1, X_2, \ldots, X_q)$. Generally, the observed variables are exactly the sinks of the directed graph, but this does not need to be the case. To simplify the notation, we make the assumption, which is often the case in practice, that all the observed variables take their values in the same finite alphabet Σ', and that all the hidden variables are on the finite alphabet Σ.

Notice that for given Σ and Σ' the homogeneous HMMs in this example depend only on a fixed set of parameters, t_{ij} and s_{ij}, even as n gets large. These are the sorts of models we are interested in. Let the number of parameters be a fixed integer d. We will name our parameters $\theta_1, \theta_2, \ldots, \theta_d$. By a graphical model with d parameters, we mean a graphical model such that each probability $\left[p_i\left(\tau_{j_1}, \ldots, \tau_{j_k}\right)\right]_{\tau_i}$ is a monomial in our parameters, and furthermore the degree of this monomial is bounded by the number of parents of v_i. This is a natural assumption, because this probability is usually a product of one parameter for each edge incoming to v, as long as the parameters affect the probability of state v_i independently. This bound on degrees encompasses most interesting and useful graphical models. For example, in the homogeneous HMM, each v_i has only one parent, and the coordinates of p_i are degree one monomials (one of t_{ij} or s_{ij}).

In what follows we denote by E the number of edges of the underlying graph of a graphical model, by n the number of observed random variables, and by q the number of hidden random variables. The observations, then, are sequences in $(\Sigma')^n$ and the explanations are sequences in Σ^q. Let $l = |\Sigma|$ and $l' = |\Sigma'|$.

For each observation τ and hidden variables $\mathbf{h}$, $\mathrm{Prob}\left(\mathbf{X} = \mathbf{h},\ \mathbf{Y} = \tau\right)$ is a monomial $f_{\mathbf{h},\tau}$ of degree at most E in the parameters $\theta_1, \theta_2, \ldots, \theta_d$. Then for each observation $\tau \in (\Sigma')^n$, the observed probability $\mathrm{Prob}(\mathbf{Y} = \tau)$ is the sum over all hidden data $\mathbf{h}$ of $\mathrm{Prob}\left(\mathbf{X} = \mathbf{h},\ \mathbf{Y} = \tau\right)$, and so $\mathrm{Prob}(\mathbf{Y} = \tau)$ is the polynomial $f_\tau = \sum_{\mathbf{h}} f_{\mathbf{h},\tau}$ in the parameters $\theta_1, \theta_2, \ldots, \theta_d$. The degree of f_τ is at most E.

Note that we have not assumed that the appropriate probabilities sum to 1. It turns out that the analysis is much easier if we do not place that restriction on our probabilities. At the end of the analysis, these restrictions may be added if desired (there are many models in use, however, which never place that restriction; these can no longer be properly called "probabilistic" models, but in fact belong to a more general class of "scoring" models which our analysis encompasses).

2.2. Inference functions. For fixed values of the parameters, the basic inference problem is to determine, for each given observation τ, the value $\mathbf{h} \in \Sigma^q$ of the hidden data that maximizes $\mathrm{Prob}(\mathbf{X} = \mathbf{h} \mid \mathbf{Y} = \tau)$. A solution to this optimization problem is denoted $\widehat{\mathbf{h}}$ and is called an *explanation* of the observation τ. Each choice of parameter values $(\theta_1, \theta_2, \ldots, \theta_d)$ defines an *inference function* $\tau \mapsto \widehat{\mathbf{h}}$ from the set of observations $(\Sigma')^n$ to the set of explanations Σ^q.

It is possible that there is more than one value of $\widehat{\mathbf{h}}$ attaining the maximum of $\mathrm{Prob}(\mathbf{X} = \mathbf{h} \mid \mathbf{Y} = \tau)$. In this case, for simplicity, we will pick only one such explanation, according to some consistent tie-breaking rule decided ahead of time. For example, we can pick the least such $\widehat{\mathbf{h}}$ in some given total order of the set Σ^q of hidden states. Another alternative would be to define inference functions as maps from $(\Sigma')^n$ to subsets of Σ^q. This would not affect the results of this paper, so for the sake of simplicity, we consider only inference functions as defined above.

It is interesting to observe that the total number of maps $(\Sigma')^n \longrightarrow \Sigma^q$ is $(l^q)^{(l')^n} = l^{q(l')^n}$, which is doubly-exponential in the length n of the observations. However, most of these maps are not inference functions for any values of the parameters. Before our results, the best upper bound in the literature was an exponential bound given in [**8**, Corollary 10]. In Section 3 we give a polynomial upper bound on the number of inference functions of a graphical model.

2.3. Polytopes. Here we review some facts about convex polytopes, and we introduce some notation. Recall that a polytope is a bounded intersection of finitely many closed halfspaces, or equivalently, the convex hull of a finite set of points. For the basic definitions about polytopes we refer the reader to [**10**].

Given a polynomial $f(\theta) = \sum_{i=1}^N \theta_1^{a_{1,i}} \theta_2^{a_{2,i}} \cdots \theta_d^{a_{d,i}}$, its *Newton polytope*, denoted by $\mathrm{NP}(f)$, is defined as the convex hull in $\mathbb{R}^d$ of the set of points $\{(a_{1,i}, a_{2,i}, \ldots, a_{d,i}) : i = 1, \ldots, N\}$. For example, if $f(\theta_1, \theta_2) = 2\theta_1^3 + 3\theta_1^2\theta_2^2 + \theta_1\theta_2^2 + 3\theta_1 + 5\theta_2^4$, then its Newton polytope $\mathrm{NP}(f)$ is given in Figure 2.

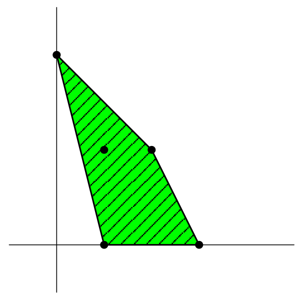


Figure 2. The Newton polytope of $f(\theta_1, \theta_2) = 2\theta_1^3 + 3\theta_1^2\theta_2^2 + \theta_1\theta_2^2 + 3\theta_1 + 5\theta_2^4$.

Given a polytope $P \subset \mathbb{R}^d$ and a vector $w \in \mathbb{R}^d$, the set of all points in P at which the linear functional $x \mapsto x \cdot w$ attains its maximum determines a *face* of P. It is denoted

$$\mathrm{face}_w(P) \quad = \quad \big\{ x \in P \;:\; x \cdot w \geq y \cdot w \text{ for all } y \in P \big\}.$$

Faces of dimension 0 (consisting of a single point) are called *vertices*, and faces of dimension 1 are called *edges*. If d is the dimension of the polytope, then faces of dimension $d-1$ are called *facets*.

Let P be a polytope and F a face of P. The *normal cone* of P at F is

$$N_P(F) \quad = \quad \big\{ w \in \mathbb{R}^d : \mathrm{face}_w(P) = F \big\}.$$

The collection of all cones $N_P(F)$ as F runs over all faces of P is denoted $\mathcal{N}(P)$ and is called the *normal fan* of P. Thus the normal fan $\mathcal{N}(P)$ is a partition of $\mathbb{R}^d$ into cones. The cones in $\mathcal{N}(P)$ are in bijection with the faces of P, and if $w \in N_P(F)$ then the linear functional $x \cdot w$ is maximized on F. Figure 3 shows the normal fan of a polytope.

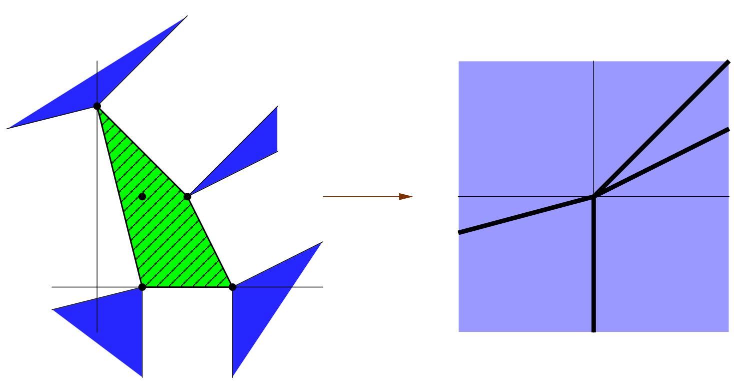


Figure 3. The normal fan of a polytope.

The *Minkowski sum* of two polytopes P and P' is defined as

$$P + P' := \{\mathbf{x} + \mathbf{x}' : \mathbf{x} \in P,\, \mathbf{x}' \in P'\}.$$

The *common refinement* of two or more normal fans is the collection of cones obtained as the intersection of a cone from each of the individual fans. For polytopes $P_1, P_2, \ldots, P_k$, the common refinement of their normal fans is denoted $\mathcal{N}(P_1) \wedge \cdots \wedge \mathcal{N}(P_k)$. The following lemma states the well-known fact that the normal fan of a Minkowski sum of polytopes is the common refinement of their individual fans (see [**10**, Proposition 7.12] or [**2**, Lemma 2.1.5]):

LEMMA 2.2. $\mathcal{N}(P_1 + \cdots + P_k) = \mathcal{N}(P_1) \wedge \cdots \wedge \mathcal{N}(P_k)$.

We finish with a result of Gritzmann and Sturmfels that will be useful later. It gives a bound on the number of vertices of a Minkowski sum of polytopes.

THEOREM 2.3 ([**2**]). *Let $P_1, P_2, \ldots, P_k$ be polytopes in $\mathbb{R}^d$, and let m denote the number of non-parallel edges of $P_1, \ldots, P_k$. Then the number of vertices of $P_1 + \cdots + P_k$ is at most*

$$2\sum_{j=0}^{d-1} \binom{m-1}{j}.$$

Note that this bound is independent of the number k of polytopes.

3. An upper bound on the number of inference functions

For fixed parameters, the inference problem of finding the explanation $\widehat{\mathbf{h}}$ that maximizes $\mathrm{Prob}(\mathbf{X} = \mathbf{h} \mid \mathbf{Y} = \tau)$ is equivalent to identifying the monomial $f_{\mathbf{h},\tau} = \theta_1^{a_{1,\mathbf{h}}}\theta_2^{a_{2,\mathbf{h}}} \cdots \theta_d^{a_{d,\mathbf{h}}}$ of f_τ with maximum value. Since the logarithm is a monotonically increasing function, the desired monomial also maximizes the quantity

$$\begin{aligned} \log(\theta_1^{a_{1,\mathbf{h}}}\theta_2^{a_{2,\mathbf{h}}} \cdots \theta_d^{a_{d,\mathbf{h}}}) &= a_{1,\mathbf{h}}\log(\theta_1) + a_{2,\mathbf{h}}\log(\theta_2) + \cdots + a_{d,\mathbf{h}}\log(\theta_d) \\ &= a_{1,\mathbf{h}}v_1 + a_{2,\mathbf{h}}v_2 + \cdots + a_{d,\mathbf{h}}v_d, \end{aligned}$$

where we replace $\log(\theta_i)$ with v_i. This is equivalent to the fact that the corresponding point $(a_{1,\mathbf{h}}, a_{2,\mathbf{h}}, \ldots, a_{d,\mathbf{h}})$ maximizes the linear expression $v_1x_1 + \cdots + v_dx_d$ on the Newton polytope $\mathrm{NP}(f_\tau)$. Thus, the inference problem for fixed parameters becomes a linear programming problem.

Each choice of the parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_d)$ determines an inference function. If $\mathbf{v} = (v_1, v_2, \ldots, v_d)$ is the vector in $\mathbb{R}^d$ with coordinates $v_i = \log(\theta_i)$, then we denote the corresponding inference function by

$$\Phi_{\mathbf{v}} : (\Sigma')^n \longrightarrow \Sigma^q.$$

For each observation $\tau \in (\Sigma')^n$, its explanation $\Phi_{\mathbf{v}}(\tau)$ is given by the vertex of $\mathrm{NP}(f_\tau)$ that is maximal in the direction of the vector $\mathbf{v}$. Note that for certain values of the parameters (if $\mathbf{v}$ is perpendicular to a positive-dimensional face of $\mathrm{NP}(f_\tau)$) there may be more than one vertex attaining the maximum. It is also possible that a single point $(a_{1,\mathbf{h}}, a_{2,\mathbf{h}}, \ldots, a_{d,\mathbf{h}})$ in the polytope corresponds to several different values of the hidden data. In both cases, we pick the explanation according to the tie-breaking rule determined ahead of time. This simplification does not affect the asymptotic number of inference functions.

Different values of θ yield different directions $\mathbf{v}$, which can result in distinct inference functions. We are interested in bounding the number of different inference functions that a graphical model can have. The next theorem gives an upper bound which is polynomial in the size of the graphical model. In fact, very few of the $l^{q(l')^n}$ functions $(\Sigma')^n \longrightarrow \Sigma^q$ are inference functions.

THEOREM 3.1 (The Few Inference Functions Theorem). *Let d be a fixed positive integer. Consider a graphical model with d parameters, and let E be the number of edges of the underlying graph. Then, the number of inference functions of the model is at most $O(E^{d(d-1)})$.*

Before proving this theorem, observe that the number E of edges depends on the number n of observed random variables. In most graphical models of interest, E is a linear function of n, so the bound becomes $O(n^{d(d-1)})$. For example, the hidden Markov model has $E = 2n-1$ edges. The only property of the number E that we actually need in the proof is that it is a bound on the degrees of the monomials $f_{\mathbf{h},\tau}$.

Proof. In the first part of the proof we will reduce the problem of counting inference functions to the enumeration of the vertices of a certain polytope. We have seen that an inference function is specified by a choice of the parameters, which is equivalent to choosing a vector $\mathbf{v} \in \mathbb{R}^d$. The function is denoted $\Phi_{\mathbf{v}} : (\Sigma')^n \longrightarrow \Sigma^q$, and the explanation $\Phi_{\mathbf{v}}(\tau)$ of a given observation τ is determined by the vertex of $\mathrm{NP}(f_\tau)$ that is maximal in the direction of $\mathbf{v}$. Thus, cones of the normal fan $\mathcal{N}(\mathrm{NP}(f_\tau))$ correspond to sets of vectors $\mathbf{v}$ that give rise to the same explanation for the observation τ. Non-maximal cones (i.e., those contained in another cone of higher dimension) correspond to directions $\mathbf{v}$ for which more than one vertex is maximal. Since ties are broken using a consistent rule, we disregard this case for simplicity. Thus, in what follows we consider only maximal cones of the normal fan.

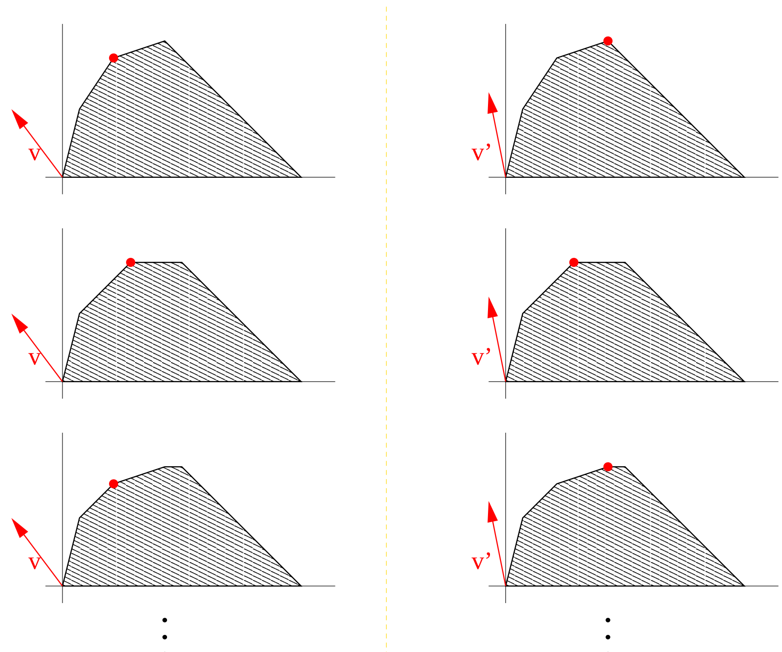


Figure 4. Two different inference functions, $\Phi_{\mathbf{v}}$ (left column) and $\Phi_{\mathbf{v}'}$ (right column). In each row is the Newton polytope corresponding to a different observation. The respective explanations are given by the marked vertices in each polytope.

Let $\mathbf{v}' = (v'_1, v'_2, \ldots, v'_d)$ be another vector corresponding to a different choice of parameters (see Figure 4). By the above reasoning, $\Phi_{\mathbf{v}}(\tau) = \Phi_{\mathbf{v}'}(\tau)$ if and only if $\mathbf{v}$ and $\mathbf{v}'$ belong to the same cone of $\mathcal{N}(\mathrm{NP}(f_\tau))$. Thus, $\Phi_{\mathbf{v}}$ and $\Phi_{\mathbf{v}'}$ are the same inference function if and only if $\mathbf{v}$ and $\mathbf{v}'$ belong to the same cone of $\mathcal{N}(\mathrm{NP}(f_\tau))$ for all observations $\tau \in (\Sigma')^n$. Consider the common refinement of all these normal fans, $\bigwedge_{\tau\in(\Sigma')^n} \mathcal{N}(\mathrm{NP}(f_\tau))$. Then, $\Phi_{\mathbf{v}}$ and $\Phi_{\mathbf{v}'}$ are the same inference function exactly when $\mathbf{v}$ and $\mathbf{v}'$ lie in the same cone of this common refinement. This implies that the number of inference functions equals the number of cones in $\bigwedge_{\tau\in(\Sigma')^n} \mathcal{N}(\mathrm{NP}(f_\tau))$. By Lemma 2.2, this common refinement is the normal fan of $\mathrm{NP}(\mathbf{f}) = \sum_{\tau\in(\Sigma')^n} \mathrm{NP}(f_\tau)$, the Minkowski sum of the polytopes $\mathrm{NP}(f_\tau)$ for all observations τ. It follows that enumerating inference functions is equivalent to counting vertices of $\mathrm{NP}(\mathbf{f})$. In the remaining part of the proof we give an upper bound on the number of vertices of $\mathrm{NP}(\mathbf{f})$.

Note that for each τ, the polytope $\mathrm{NP}(f_\tau)$ is contained in the hypercube $[0, E]^d$, since each parameter θ_i can appear as a factor of a monomial of f_τ at most E times. Also, the vertices of $\mathrm{NP}(f_\tau)$ have integral coordinates, because they are exponent vectors. Polytopes whose vertices have integral coordinates are called *lattice polytopes*. It follows that the edges of $\mathrm{NP}(f_\tau)$ are given by vectors where each coordinate is an integer between $-E$ and E. There are only $(2E+1)^d$ such vectors, so this is an upper bound on the number of different directions that the edges of the polytopes $\mathrm{NP}(f_\tau)$ can have.

This property of the Newton polytopes of the coordinates of the model will allow us to give an upper bound on the number of vertices of their Minkowski sum $\mathrm{NP}(\mathbf{f})$. The last ingredient that we need is Theorem 2.3. In our case we have a sum of polytopes $\mathrm{NP}(f_\tau)$, one for each observation $\tau \in (\Sigma')^n$, having at most $(2E+1)^d$ non-parallel edges in total. Hence, by Theorem 2.3, the number of vertices of $\mathrm{NP}(\mathbf{f})$ is at most

$$2\sum_{j=0}^{d-1}\binom{(2E+1)^d-1}{j}.$$

As E goes to infinity, the dominant term of this expression is

$$\frac{2^{d^2-d+1}}{(d-1)!}\,E^{d(d-1)}.$$

Thus, we get an $O(E^{d(d-1)})$ upper bound on the number of inference functions of the graphical model. $\square$

In the next section we show that the bound given in Theorem 3.1 is tight up to a constant factor.

4. A lower bound

As before, we fix d, the number of parameters in our model. The Few Inference Functions Theorem (Theorem 3.1) tells us that the number of inference functions is bounded from above by some function $cE^{d(d-1)}$, where c is a constant (depending only on d) and E is the number of edges in the graphical model. Here we show that this bound is tight up to a constant, by constructing a family of graphical models whose number of inference functions is at least $\tilde{c}E^{d(d-1)}$, where $\tilde{c}$ is another constant. In fact, we will construct a family of hidden Markov models with this property. To be precise, we have the following theorem.

Theorem 4.1. *Fix d. There is a constant $c' = c'(d)$ such that, given $n \in \mathbb{Z}_+$, there exists an HMM of length n, with d parameters, $2d+2$ hidden states, and 2 observed states, such that there are at least $c'n^{d(d-1)}$ distinct inference functions. (For HMMs, $E = 2n-1$, so this also gives us the lower bound in terms of E).*

In the proof of this theorem, we will state several lemmas that must be used. We omit the proofs of some of them here due to lack of space. Given n, we first construct the appropriate HMM, $\mathcal{M}_n$, using the following lemma.

Lemma 4.2. *Given $n \in \mathbb{Z}_+$, there is an HMM, $\mathcal{M}_n$, of length n, with d parameters, $2d+2$ hidden states, and 2 observed states, such that for any $a = (a_1, \ldots, a_n) \in \mathbb{Z}_+^d$ with $\sum_i a_i < n$, there is an observed sequence which has one explanation if*

$$a_1\log(\theta_1) + \cdots + a_d\log(\theta_d) > 0$$

and another explanation if

$$a_1\log(\theta_1) + \cdots + a_d\log(\theta_d) < 0.$$

Proof. Given d and n, define a length n HMM with parameters $\theta_1, ..., \theta_d$, as follows. The observed states will be S and C (for "start of block," and "continuing block," respectively). The hidden states will be s_i, s'_i, c_i, and c'_i, for $1 \le i \le d+1$ (think of s_i and s'_i as "start of the ith block" and c_i and c'_i as "continuing the ith block").

Here's the idea of what we want this HMM to do: if the observed sequence has S's in position 1, a_1+1, a_1+a_2+1, …, and $a_1+\cdots+a_d+1$ and C's elsewhere, then there will be only two possibilities for the sequence of hidden states, either

$$t = s_1\,\underbrace{c_1\cdots c_1}_{a_1-1}\,s_2\,\underbrace{c_2\cdots c_2}_{a_2-1}\cdots s_d\,\underbrace{c_d\cdots c_d}_{a_d-1}\,s_{d+1}\,\underbrace{c_{d+1}\cdots c_{d+1}}_{n-a_1-\cdots-a_d-1}$$

or

$$t' = s'_1\,\underbrace{c'_1\cdots c'_1}_{a_1-1}\,s'_2\,\underbrace{c'_2\cdots c'_2}_{a_2-1}\cdots s'_d\,\underbrace{c'_d\cdots c'_d}_{a_d-1}\,s'_{d+1}\,\underbrace{c'_{d+1}\cdots c'_{d+1}}_{n-a_1-\cdots-a_d-1}.$$

We will also make sure that t has a priori probability

$$\theta_1^{a_1}\cdots\theta_d^{a_d}$$

and t' has a priori probability 1. Then t is the explanation if $a_1\log(\theta_1)+\cdots+a_d\log(\theta_d) > 0$ and t' is the explanation if $a_1\log(\theta_1)+\cdots+a_d\log(\theta_d) < 0$. Remember that we are not constraining our probability sums to be 1. A very similar HMM could be constructed that obeys that constraint, if desired. But to simplify notation it will be more convenient to treat the transition probabilities as parameters that do not necessarily sum to one at each vertex, even if this forces us to use the term "probability" somewhat loosely.

Here is how we set up the transitions/emmisions. Let s_i and s'_i, for $1 \le i \le d+1$, all emit S with probability 1 and C with probability 0. Let c_i and c'_i emit C with probability 1 and S with probability 0. Let s_i, for $1 \le i \le d$, transition to c_i with probability θ_i and transition to everything else with probability 0. Let s_{d+1} transition to c_{d+1} with probability 1 and to everything else with probability 0. Let s'_i, for $1 \le i \le d+1$, transition to c'_i with probability 1 and to everything else with probability 0. Let c_i, for $1 \le i \le d$, transition to c_i with probability θ_i, to s_{i+1} with probability θ_i, and to everything else with probability 0. Let c_{d+1} transition to c_{d+1} with probability 1, and to everything else with probability 0. Let c'_i, for $1 \le i \le d$ transition to c'_i with probability 1, to s_{i+1} with probability 1, and to everything else with probability 0. Let c'_{d+1} transition to c'_{d+1} with probability 1 and to everything else with probability 0.

Starting with the uniform probability distribution on the first hidden state, this does exactly what we want it to: given the correct observed sequence, t and t' are the only explanations, with the correct probabilities. □

This means that, for the HMM provided by this lemma, the decomposition of (log-)parameter space into inference cones includes all of the hyperplanes $\{x : \langle a, x\rangle = 0\}$ such that $a \in \mathbb{Z}^d_+$ with $\sum_i a_i < n$. Call the arrangement of these hyperplanes $\mathcal{H}_n$. It suffices to show that the arrangement $\mathcal{H}_n$ consists of at least $c'n^{d(d-1)}$ chambers (full dimensional cones determined by the arrangement). There are $c_1 n^d$ ways to choose one of the hyperplanes from $\mathcal{H}_n$, for some constant c_1. Therefore there are $c_1^{d-1} n^{d(d-1)}$ ways to choose $d-1$ of the hyperplanes; their intersection is, in general, a 1-dimensional face of $\mathcal{H}_n$ (that is, the intersection is a ray which is an extreme ray for the cones it is contained in). It is quite possible that two different ways of choosing $d-1$ hyperplanes give the same extreme ray. The following lemma says that some constant fraction of these choices of extreme rays are actually distinct.

Lemma 4.3. *Fix d. Given n, let $\mathcal{H}_n$ be the hyperplane arrangement consisting of the hyperplanes of the form $\{x : \langle a, x\rangle = 0\}$ with $a \in \mathbb{Z}^d_+$ and $\sum_i a_i < n$. Then the number of 1-dimensional faces of $\mathcal{H}_n$ is $c_2 n^{d(d-1)}$, for some constant c_2.*

Each chamber will have a number of these extreme rays on its boundary. The following lemma gives a constant bound on this number.

Lemma 4.4. *Fix d. Given n, define $\mathcal{H}_n$ as above. Each chamber of $\mathcal{H}_n$ has at most $2^{d(d-1)}$ extreme rays.*

Conversely, each ray is an extreme ray for at least 1 chamber. Therefore there are at least $\frac{c_2}{2^{d(d-1)}} n^{d(d-1)}$ chambers, and Theorem 4.1 is proved.

5. Inference functions for sequence alignment

In this section we give an application of Theorem 3.1 to a basic model for sequence alignment. Sequence alignment is one of the most frequently used techniques in determining the similarity between biological sequences. In the standard instance of the sequence alignment problem, we are given two sequences (usually DNA or protein sequences) that have evolved from a common ancestor via a series of mutations, insertions and deletions. The goal is to find the best alignment between the two sequences. The definition of "best" here depends on the choice of scoring scheme, and there is often disagreement about the correct choice. In *parametric sequence alignment*, this problem is circumvented by instead computing the optimal alignment as a function of *variable* scores. Here we consider one such scheme, in which all matches are equally rewarded, all mismatches are equally penalized and all spaces are equally penalized. Efficient parametric sequence alignment algorithms are known (see for example [**6**, Chapter 7]). Here we are concerned with the different inference functions that car arise when the parameters vary. For a detailed treatment on the subject of sequence alignment, we refer the reader to [**3**].

Given two strings σ^1 and σ^2 of lengths n_1 and n_2 respectively, an *alignment* is a pair of equal length strings (μ^1, μ^2) obtained from σ^1, σ^2 by inserting dashes "$-$" in such a way that there is no position in which both μ^1 and μ^2 have a dash. A *match* is a position where μ^1 and μ^2 have the same character, a *mismatch* is a position where μ^1 and μ^2 have different characters, and a *space* is a position in which one of μ^1 and μ^2 has a dash. A simple scoring scheme consists of two parameters α and β denoting mismatch and space penalties respectively. The reward of a match is set to 1. The score of an alignment with z matches, x mismatches, and y spaces is then $z - x\alpha - y\beta$. Observe that these numbers always satisfy $2z + 2x + y = n_1 + n_2$.

This model for sequence alignment is a particular case of a so-called pair hidden Markov model. The problem of determining the highest scoring alignment for given values of α and β is equivalent to the inference problem in the pair hidden Markov model. In this setting, an observation is a pair of sequences $\tau = (\sigma^1, \sigma^2)$, and the number of observed variables is $n = n_1 + n_2$. The values of the hidden variables in an explanation indicate the positions of the spaces in the optimal alignment. We will refer to this as the *2-parameter model for sequence alignment.*

For each pair of sequences τ, the Newton polytope of the polynomial f_τ is the convex hull of the points (x, y, z) whose coordinates are the number of mismatches, spaces, and matches, respectively, of each possible alignment of the pair. This polytope is only two dimensional, as it lies on the plane $2z + 2x + y = n_1 + n_2$. No information is lost by considering its projection onto the xy-plane instead. This projection is just the convex hull of the points (x, y) giving the number of mismatches and spaces of each alignment. For any alignment of sequences of lengths n_1 and n_2, the corresponding point (x, y) lies inside the square $[0, n]^2$, where $n = n_1 + n_2$. Therefore, since we are dealing with lattice polygons inside $[0, n]^2$, it follows from the proof of the Few Inference Functions Theorem (Theorem 3.1) that the number of inference functions of this model is $O(n^{2(2-1)})) = O(n^2)$. Next we show that this quadratic bound is tight, even in the case of the binary alphabet.

PROPOSITION 5.1. *Consider the 2-parameter model for sequence alignment for two observed sequences of length n and let $\Sigma' = \{0, 1\}$ be the binary alphabet. Then, the number of inference functions of this model is $\Theta(n^2)$.*

PROOF. The above argument shows that $O(n^2)$ is an upper bound on the number of inference functions of the model. To prove the proposition, we will argue that there are at least $\Omega(n^2)$ such functions.

Since the two sequences have the same length, the number of spaces in any alignment is even. For convenience, we define $y' = y/2$ and $\beta' = 2\beta$, and we will work with the coordinates (x, y', z) and the parameters α and β'. The value y' is called the number of insertions (half the number of spaces), and β' is the insertion penalty. For fixed values of α and β', the explanation of an observation $\tau = (\sigma^1, \sigma^2)$ is given by the vertex of $\mathrm{NP}(f_\tau)$ that is maximal in the direction of the vector $(-\alpha, -\beta', 1)$. In this model, $\mathrm{NP}(f_\tau)$ is the convex hull of the points (x, y', z) whose coordinates are the number of mismatches, insertions and matches of the alignments of σ^1 and σ^2.

The argument in the proof of Theorem 3.1 shows that the number of inference functions of this model is the number of cones in the common refinement of the normal fans of $\mathrm{NP}(f_\tau)$, where τ runs over all pairs of sequences of length n in the alphabet Σ'. Since the polytopes $\mathrm{NP}(f_\tau)$ lie on the plane $x + y' + z = n$, it is equivalent to consider the normal fans of their projections onto the $y'z$-plane. These projections are lattice polygons contained in the square $[0, n]^2$. We denote by P_τ the projection of $\mathrm{NP}(f_\tau)$ onto the $y'z$-plane.

We will construct, for any relatively prime positive integers u and v with $u < v$ and $6v - 2u \le n$, a pair $\tau = (\sigma^1, \sigma^2)$ of binary sequences of length n such that P_τ has an edge of slope u/v. Such an edge gives rise to the line $u \cdot \alpha + v \cdot \beta' = 0$ separating regions in the normal fan $\mathcal{N}(P_\tau)$ and hence in $\bigwedge_\tau \mathcal{N}(P_\tau)$, where τ ranges over all pairs of binary sequences of length n. The number of such choices u, v is $\Omega(n^2)$ (this relies on the fact, see [**1**, Chapter 3], that a positive fraction of choices of $(u, v) \in \mathbb{Z}^2$ have u and v relatively prime). This implies that the number of different inference functions is $\Omega(n^2)$.

Thus, it only remains to construct such a τ, given positive integers u and v as above. Let $a := 2v$, $b := v - u$. Assume first that $n = 6v - 2u = 2a + 2b$. Consider the sequences $\sigma^1 = 0^a 1^b 0^b 1^a$, $\sigma^2 = 1^a 0^b 1^b 0^a$, where 0^a indicates that the symbol 0 is repeated a times. Let $\tau = (\sigma^1, \sigma^2)$. Then, it is not hard to see that the polygon P_τ for this pair of sequences has four vertices: $v_0 = (0, 0)$, $v_1 = (b, 3b)$, $v_2 = (a + b, a + b)$ and $v_3 = (n, 0)$. The slope of the edge between v_1 and v_2 is $\frac{a-2b}{a} = \frac{u}{v}$.

If $n > 6v - 2u = 2a + 2b$, we just append $0^{n-2a-2b}$ to both sequences σ^1 and σ^2. In this case, the vertices of P_τ are $(0, n - 2a - 2b)$, $(b, n - 2a + b)$, $(a + b, n - a - b)$, $(n, 0)$ and $(n - 2a - 2b, 0)$.

Note that if $v - u$ is even, the construction can be done with sequences of length $n = 3v - u$ by taking $a := v$, $b := \frac{v-u}{2}$. Figure 5 shows the alignment graph and the polygon P_τ for $a = 7$, $b = 2$. □

In most cases, one is interested only in those inference functions that are biologically meaningful. This corresponds to parameter values with $\alpha, \beta \ge 0$, which means that mismatches and spaces are penalized instead of rewarded. Sometimes one also requires that $\alpha \le \beta$, which means that a mismatch should be penalized less than two spaces. It is interesting to observe that our construction in the proof of Proposition 5.1 not only shows that the total number of inference functions is $\Omega(n^2)$, but also that the number of biologically

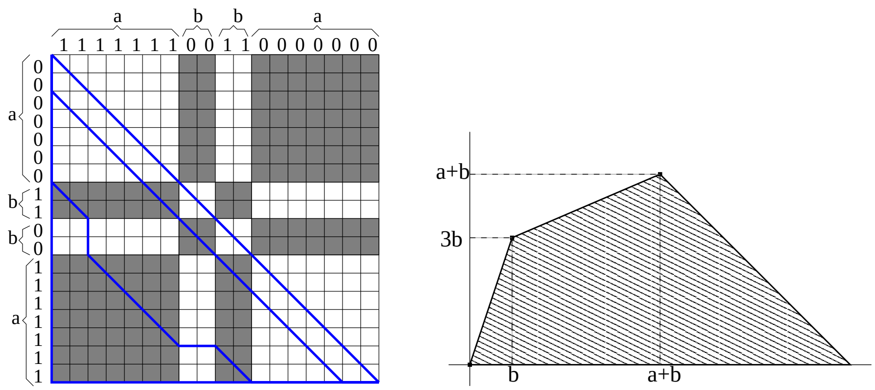


FIGURE 5. A pair of binary sequences of length 18 giving the slope 3/7 in their alignment polytope. The four paths in the *alignment graph* on the left correspond to the four vertices; a right step in the graph corresponds to a space in σ^1, a down step to a space in σ^2, and a diagonal step to a match or mismatch. See [**6**, Section 2.2] for a full definition of the alignment graph.

meaningful ones is still $\Omega(n^2)$. This is because the different rays created in our construction have a biologically meaningful direction in the parameter space.

6. Final remarks

An interpretation of Theorem 3.1 is that the ability to change the values of the parameters of a graphical model does not give as much freedom as it may appear. There is a very large number of possible ways to assign an explanation to each observation. However, only a tiny proportion of these come from a consistent method for choosing the most probable explanation for a certain choice of parameters. Even though the parameters can vary continuously, the number of different inference functions that can be obtained is at most polynomial in the number of edges of the model, assuming that the number of parameters is fixed.

In the case of sequence alignment, the number of possible functions that associate an alignment to each pair of sequences of length n is doubly-exponential in n. However, the number of functions that pick the alignment with highest score in the 2-parameter model, for some choice of the parameters α and β, is only $\Theta(n^2)$. Thus, most ways of assigning alignments to pairs of sequences do not correspond to any consistent choice of parameters. If we use a model with more parameters, say d, the number of inference functions may be larger, but still polynomial in n, namely $O(n^{d(d-1)})$.

Having shown that the number of inference functions of a graphical model is polynomial in the size of the model, an interesting next step would be to find an efficient way to precompute all the inference functions for given models. This would allow us to give the answer (the explanation) to a query (an observation) very quickly. Theorem 3.1 suggests that it might be computationally feasible to precompute the polytope $\mathrm{NP}(\mathbf{f})$, whose vertices correspond to the inference functions. However, the difficulty arises when we try to describe a particular inference function efficiently. The problem is that the characterization of an inference function involves an exponential number of observations.

References

[1] T.M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, New York, 1976.
[2] P. Gritzmann, B. Sturmfels, Minkowski addition of polytopes: Computational complexity and applications to Gröbner bases, *SIAM Journal of Discrete Mathematics* 6 (1993), 246–269.
[3] D. Gusfield, *Algorithms on Strings, Trees, and Sequences*, Cambridge University Press, 1997.
[4] D. Gusfield, K. Balasubramanian, D. Naor, Parametric optimization of sequence alignment, *Algorithmica* 12 (1994), 312–326.
[5] F. Jensen, *Bayesian Networks and Decision Graphs*, Springer, 2001.
[6] L. Pachter, B. Sturmfels, editors, *Algebraic Statistics for Computational Biology*, Cambridge University Press, 2005.
[7] L. Pachter, B. Sturmfels, The Mathematics of Phylogenomics, submitted.
[8] L. Pachter, B. Sturmfels, Tropical Geometry of Statistical Models, *Proc. Natl. Acad. Sci.* 101, n. 46 (2004), 16132–16137.
[9] Christophe Weibel, personal commnuication.

[10] G.M. Ziegler, *Lectures on Polytopes*, Graduate Texts in Mathematics 152, Springer, New York, 1995.

Department of Mathematics, Dartmouth College, Hanover, NH 03755
E-mail address: sergi.elizalde@dartmouth.edu

Department of Mathematics, University of California, Berkeley, CA 94720
E-mail address: kwoods@math.berkeley.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

From Orbital Varieties to Alternating Sign Matrices

P. Di Francesco and P. Zinn-Justin

ABSTRACT. We study a one-parameter family of vector-valued polynomials associated to each simple Lie algebra. When this parameter q equals -1 one recovers Joseph polynomials, whereas at q cubic root of unity one obtains ground state eigenvectors of some integrable models with boundary conditions depending on the Lie algebra; in particular, we find that the sum of its entries is related to numbers of Alternating Sign Matrices and/or Plane Partitions in various symmetry classes.

RÉSUMÉ. Nous étudions une famille à un paramètre de polynômes à valeurs vectorielles qui est associée à chaque algèbre de Lie simple. Quand ce paramètre q vaut -1 on retrouve les polynômes de Joseph, tandis que quand q est racine cubique de l'unité on obtient les états fondamentaux de certain modèles intégrables avec des conditions aux bords dépendant de l'algèbre de Lie ; en particulier, nous trouvons que la somme de ses composantes est reliée aux nombres de Matrices de Signe Alterné et/ou de Partitions Planes dans diverses classes de symétrie.

1. Introduction

Recently, a remarkable connection between integrable models and combinatorics has emerged. It first appeared in a series of papers concerning the XXZ spin chain and the Temperley–Lieb (TL) loop model [**1, 2**] and which culminated with the so-called Razumov–Stroganov (RS) conjecture [**3**]. One of the main observations of [**1**], a weak corollary of the RS conjecture, is that the sum of entries of the properly normalized ground state vector of the TL(1) loop model is (unexpectedly!) equal to the number of Alternating Sign Matrices. This result was eventually proved in [**4**] by using the *integrability* of the TL loop model in the following way: the model is generalized by introducing N complex numbers (spectral parameters, or inhomogeneities) in the problem, where N is the size of the system. The ground state entries become polynomials in these variables, and integrability provides many new tools for analyzing them, leading eventually to the exact computation of their sum, identified as the so-called Izergin–Korepin (IK) determinant, known to specialize to the number of Alternating Sign Matrices in the homogeneous limit [**5**]. Note that in this work, the meaning of the spectral parameters is not very transparent; in particular, it is unclear how to generalize the full RS conjecture in their presence.

Next, it was observed in [**6**] that the polynomials obtained above really belong to a one-parameter family of solutions of a certain set of linear equations, in which the parameter q has been set equal to a cubic root of unity. This observation is not obvious because the equations for generic q are not a simple eigenvector equation; in fact, as explained in [**7**], they are precisely the quantum Knizhnik–Zamolodchikov (qKZ) equations at level 1 for the algebra $U_q(\widehat{\mathfrak{sl}(2)})$. Furthermore, in the "rational" limit $q \to -1$, these polynomials have a remarkable geometric interpretation: they are equivariant Hilbert polynomials (or "multidegrees") of A_{N-1} orbital varieties $M^2 = 0$ ([**7**], see also [**8**]), which are extensions of the Joseph polynomials [**11**]. Note that here, the spectral parameters quite naturally appear as the basis of weights of $\mathfrak{gl}(N)$. In [**7**], these ideas were generalized to higher algebras $U_q(\widehat{\mathfrak{sl}(k)})$, which correspond to the orbital varieties $M^k = 0$.

Key words and phrases. algebraic combinatorics, alternating sign matrices, integrable models.

The authors acknowledge the support of the European networks "ENIGMA" MRT-CT-2004-5652, "ENRAGE" MRTN-CT-2004-005616, and of the Geocomp project (ACI Masse de Données).

Here, we pursue a *different* type of generalization: we investigate orbital varieties corresponding to the other infinite series of simple Lie algebras: B_r, C_r, D_r; but we stick to the $U_q(\widehat{\mathfrak{sl}(2)})$ case by choosing the orbital varieties $M^2 = 0$, M a complex matrix in the fundamental representation. Indeed, we show below that such orbital varieties are related to the same loop model, but with different *boundary conditions* (corresponding to variants of the Temperley–Lieb algebra). Furthermore, one can now q-deform the resulting polynomials to produce solutions of qKZ equations of type B, C, D and set q to be a cubic root of unity. Taking the homogeneous limit, the entries become integer numbers, which we conjecture to be related to *symmetry classes* of Alternating Sign Matrices and/or Plane Partitions; in particular we identify the sums of entries.

In what follows we state most results without proofs; some will appear in a joint paper with A. Knutson [**15**] on a closely related subject.

2. General setup

2.1. Orbital varieties. Let $\mathfrak{g}$ be a simple complex Lie algebra of rank r, $\mathfrak{b}$ a Borel subalgebra. $\mathfrak{b} = \mathfrak{t} \oplus \mathfrak{n}$ where $\mathfrak{t}$ is the corresponding Cartan subalgebra and $\mathfrak{n}$ is the space of nilpotent elements of $\mathfrak{b}$. B and T are Borel and Cartan subgroups. Let W denote the Weyl group of $\mathfrak{g}$, and s_α its standard generators, where α runs over the set of simple roots of $\mathfrak{g}$.

Fixing an orbit $G{\cdot}x$, with $x \in \mathfrak{n}$ and G acting by conjugation, one can consider the irreducible components of $\overline{\mathfrak{b} \cap (G \cdot x)}$, which are called orbital varieties.

Even though much of what follows can be done for any orbital varieties, we focus below on the following special case: we fix an irreducible representation ρ (of dimension N) and consider the scheme $E = \{x \in \mathfrak{b} \mid \rho(x)^2 = 0\}$. The underlying set is precisely a $\overline{\mathfrak{b} \cap (G \cdot x)}$, where x is any element of E such that $\rho(x)$ is of maximal rank. In some sense, its components are the "simplest possible" orbital varieties.

2.2. Hotta construction. It is known that there exists a representation of the Weyl group W on the vector space V of formal linear combinations of orbital varieties (Springer/Joseph representation); for each G-orbit, it is an irreducible representation. We use the following explicit form of the representation: note that orbital varieties are invariant under $T \times \mathbb{C}^\times$, where T acts by conjugation and $\mathbb{C}^\times$ acts by overall scaling. We can therefore consider equivariant cohomology $H^*_{T\times\mathbb{C}^\times}(\cdot)$ and in particular via the inclusion map from each orbital variety π to the space $\mathfrak{n}$, the unit of $H^*_{T\times\mathbb{C}^\times}(\pi)$ is pushed forward to some cohomology class Ψ_π in $H^*_{T\times\mathbb{C}^\times}(\mathfrak{n}) = \mathbb{C}[\mathfrak{t}, A]$, that is a polynomial in $r+1$ variables $\alpha_1, \ldots, \alpha_r$, A (the r simple roots plus the $\mathbb{C}^\times$ weight), sometimes called *multidegree* of π. Suppressing the $\mathbb{C}^\times$ action, that is setting $A = 0$, one recovers the Joseph polynomials [**11**].

The way that W acts on these polynomials can be described explicitly, by extending slightly the results of Hotta [**12**] to include the additional $\mathbb{C}^\times$ action. One starts by associating to each simple root α a certain geometric construction, which we briefly recall. For $x \in \mathfrak{b}$ write $x = \sum_\alpha x_\alpha e_\alpha$ where α runs over positive roots, $e_\alpha \in \mathfrak{g}$ being a vector of weight α. Define $\mathfrak{b}_\alpha = \{x \in \mathfrak{b} \mid x_\alpha = 0\}$, and L_α to be Lévy subgroup whose Lie algebra is $\mathfrak{b} \oplus \mathbb{C}e_{-\alpha}$. Starting from an orbital variety π, we distinguish two cases:

- $\pi \subset \mathfrak{b}_\alpha$. Then set $s_\alpha \pi = \pi$.
- $\pi \not\subset \mathfrak{b}_\alpha$. Then let L_α acts by conjugation: the top-dimensional components of $L_\alpha \cdot (\pi \cap \mathfrak{b}_\alpha)$ are again orbital varieties; set $s_\alpha \pi = -\pi - \sum_{\pi'} {\mu_\alpha}_\pi^{\pi'} \pi'$ where ${\mu_\alpha}_\pi^{\pi'}$ is the multiplicity of π' in $L_\alpha \cdot (\pi \cap \mathfrak{b}_\alpha)$.

These elementary operations have a counterpart when acting on multidegrees, and a simple calculation shows that both cases are covered by a single formula:

$$s_\alpha \Psi_\pi = (-\tau_\alpha + A\partial_\alpha)\Psi_\pi \tag{2.1}$$

where τ_α is the reflection orthogonal to the root α in $\mathbb{C}[\alpha_1, \ldots, \alpha_r, A]$, and $\partial_\alpha = \frac{1}{\alpha}(\tau_\alpha - 1)$ is the associated *divided difference operator*, whereas on the left hand side s_α implements right action on the Ψ_π, namely $s_\alpha \Psi_\pi := -\Psi_\pi - \sum_{\pi'} {\mu_\alpha}_\pi^{\pi'} \Psi_{\pi'}$. One can check that $s_\alpha \mapsto -\tau_\alpha + A\partial_\alpha$ is a representation of the Weyl group W on polynomials. Note that at $A = 0$, we recover the natural action of W (up to a sign, with our conventions).

2.3. Yang–Baxter equation and integrable models. Let us define the operator

$$R_\alpha(u) := \frac{A - u s_\alpha}{A + u} \tag{2.2}$$

which acts in the space $V \otimes \mathbb{C}[\alpha_1, \ldots, \alpha_r, A]$, u being a formal parameter. Rewriting slightly the relation (2.1) above we find that τ_α acts as $R_\alpha(\alpha)$. Using the fact that τ_α, just like the s_α, satisfy the Weyl group relations, we find that the operators $\tau_\alpha R_\alpha(\alpha)$ also satisfy those. In the case of non-exceptional Lie algebras, there are only 2 types of edges in the Dynkin diagram, and therefore we have Coxeter relations of the form $(s_\alpha s_\beta)^{m_{\alpha\beta}} = 1$, where $m_{\alpha\beta} = 1, 2, 3, 4$ depending on whether $\alpha = \beta$, there is no edge, a single or a double edge between α and β. Writing these relations for $\tau_\alpha R_\alpha$ and eliminating the τ_α, we find that relations with $m_{\alpha\beta} = 1, 3, 4$ correspond respectively to the *unitarity* equation:

$$R_\alpha(\alpha) R_\alpha(-\alpha) = 1 \ , \tag{2.3}$$

the *Yang–Baxter* equation:

$$R_\alpha(\alpha) R_\beta(\alpha+\beta) R_\alpha(\beta) = R_\beta(\beta) R_\alpha(\alpha+\beta) R_\beta(\alpha) \qquad \alpha \ \circ\!\!-\!\!\!-\!\!\circ\ \beta \tag{2.4}$$

and the *boundary Yang–Baxter* (or reflection) equation:

$$R_\alpha(\alpha) R_\beta(\beta+\alpha) R_\alpha(\alpha+2\beta) R_\beta(\beta) = R_\beta(\beta) R_\alpha(\alpha+2\beta) R_\beta(\beta+\alpha) R_\alpha(\alpha) \qquad \alpha \ \circ\!\!\Rightarrow\!\!\circ\ \beta \tag{2.5}$$

whereas the case $m_{\alpha\beta} = 2$ expresses a simple commutation relation for distant vertices. Indeed one recognizes in $R_\alpha(u)$ a standard form of the rational solution of the Yang–Baxter equation, the parameter u playing the role of difference of spectral parameters. Thus the multidegrees Ψ_α are closely connected to integrable models with rational dependence on spectral parameters, as will be discussed now.

Before doing so, let us remark that in the special case investigated here of orbital varieties associated to $M^2 = 0$, the s_α obey more than just the Coxeter relations. In the A_r case they actually generate a quotient of the symmetric group algebra S_{r+1} known as the Temperley–Lieb algebra $TL_{r+1}(2)$ (here 2 is the value of the parameter in the definition of the algebra, as will be explained below). The same type of phenomena will be described for other simple Lie algebras, and will lead to variants of the Temperley–Lieb algebra; in particular, the "bulk" (i.e. everything but a finite number of edges at the boundary) of the Dynkin diagrams being sequences of simple edges, these variants will only differ at the level of "boundary conditions" of the model.

2.4. Affinization and rational qKZ equation. Let us now discuss the meaning of the equation

$$R_\alpha(\alpha)\Psi = \tau_\alpha \Psi \tag{2.6}$$

where τ_α is the reflection associated to the root α acting on the "spectral parameters" α_1, …, α_r, $R_\alpha(\alpha)$ is a certain linear operator defined above acting in the space $V \otimes \mathbb{C}[\alpha_1, \ldots, \alpha_r, A]$ and $\Psi = \sum_\pi \pi \otimes \Psi_\pi$ is a vector in that space.

When $R_\alpha(u)$ is the R-matrix (or boundary R-matrix) of some integrable model, such equations are satisfied by eigenvectors of the corresponding integrable transfer matrix. More generally, these equations appear in the context of the quantum Knizhnik–Zamolodchikov (qKZ) equation, in connection with the representation theory of affine quantum groups [**13**]. In either case, it is known that we need an additional equation to fix the Ψ_π entirely.

Define $\hat{W}$ to be the semi-direct product of W and of the weight lattice of $\mathfrak{g}$. It contains as a finite index subgroup the usual affine Weyl group defined as the Coxeter group of the affinized Dynkin diagram. Just like the affine Weyl group, it has a natural action on $\mathfrak{t}$ and therefore on $\mathbb{C}[\alpha_1, \ldots, \alpha_r, A]$ which extends the action of W generated by the reflections τ_i; by definition, in this representation, an element of the weight lattice acts as translation in $\mathfrak{t}$ of the weight multiplied by $3A$ ($3 = l + \check{h}$ where $l = 1$ is the level of the qKZ equation and $\check{h} = 2$ is the dual Coxeter number of $\mathfrak{sl}(2)$).

Then we claim that one can extend the representation of W on $V \otimes \mathbb{C}[\alpha_1, \ldots, \alpha_r, A]$ (the operators $\tau_\alpha R_\alpha(\alpha)$) into a representation of $\hat{W}$, in such a way that each element of $\hat{W}$ is the product of its natural action on $\mathbb{C}[\alpha_1, \ldots, \alpha_r, A]$ and of a $\mathbb{C}[\alpha_1, \ldots, \alpha_r, A]$-linear operator. Describing here the geometric procedure that leads to this action is beyond the scope of this paper. The action will however be described explicitly in each of the cases below. An important property is that if one sets $A = 0$ the representation of $\hat{W}$ factors through the projection $\hat{W} \to W$. So the $\mathbb{C}^\times$ action actually produces the affinization.

Imposing that Ψ be invariant under the action of the whole group $\hat{W}$ leads to a full set of equations, which are precisely equivalent to the so-called rational qKZ equation (or more precisely, a generalization of it for arbitrary Dynkin diagram, the original qKZ equation corresponding to the case A_r) at level 1; and it turns out that they have a unique polynomial solution of the prescribed degree (up to multiplication by a scalar).

2.5. q-deformation and Razumov–Stroganov point. The integrability suggests how to q-deform the above construction. Indeed, we have considered thus far R-matrices that form so-called *rational* solutions of the Yang–Baxter Equation, and Ψ's that are solutions of the rational qKZ equation. It is known however that the trigonometric R-matrices are a special degeneration of a one-parameter family of *trigonometric* solutions of the Yang–Baxter Equation, depending on a parameter q. Setting $q = -e^{-\hbar A/2}$, one customarily uses exponentiated "multiplicative" spectral parameters of the form $e^{-\hbar\alpha_i}$. We then look for polynomial solutions Ψ of these parameters, to the corresponding trigonometric qKZ equations. The rational solutions are then recovered from the trigonometric ones via the limit $\hbar \to 0$, at the first non-trivial order in $\hbar$. The details of the bulk and boundary R-matrices will be given below for the cases A_r, B_r, C_r and D_r. We thus obtain, for any q, a representation of the group $\hat{W}$, the W relations satisfied by the $\tau_\alpha R_\alpha(\alpha)$ and more generally the $\hat{W}$ relations being undeformed.

In terms of the new variables $e^{-\hbar\alpha_i}$ living in T, the natural action of an element of the weight lattice ω (as the abelian subgroup of $\hat{W}$) is the multiplication by $q^{6\omega}$. Since for all simple Lie algebras, ω has half-integer coordinates, we reach the important conclusion that when $q^3 = 1$, this action becomes trivial. Therefore, all operators associated to the weight lattice by the procedure outlined in the previous section become $\mathbb{C}[\alpha_1, \dots, \alpha_r, A]$-linear (i.e. correspond to finite-dimensional operators on V after evaluation of the parameters α_1, …, α_r, A). In this case they are simply the scattering matrices of [**19**], and they commute with the usual (inhomogeneous) integrable transfer matrix of the model. This implies that Ψ is an eigenvector of the latter; in fact, we can call it "ground state eigenvector" because in the physical situation where the transfer matrix elements are positive, the Perron–Frobenius theorem applies and the eigenvalue 1 of Ψ is the largest eigenvalue in modulus.

The value $q = e^{2i\pi/3}$ (also called "Razumov–Stroganov point") is henceforth quite special and deserves a particular study. In particular, in the homogeneous limit where the spectral parameters α_i are specialized to zero, Ψ can be normalized so that its entries are all *non-negative integers*, and we are interested in their combinatorial significance, in relation to the counting of Alternating Sign Matrices and/or Plane Partitions. We do not claim to have a full understanding of the general correspondence principle between simple Lie algebras and these combinatorial problems, but we will perform a case-by-case study for A_r, B_r, C_r and D_r.

A last remark is in order. As we shall see, it is simple to see that the solutions Ψ to the A, B, C, D qKZ equations obey recursion relations, that allow to obtain the rank r case from rank $r+1$, hence we will content ourselves with the detailed description for r with a given parity, namely A_{2n-1}, B_{2n}, C_{2n+1}, D_{2n+1}.

3. A_r case

We review the A_r case, already explored in [**7**]. We set $\alpha_i = z_i - z_{i+1}$, $i = 1, \dots, r$. The fact that there are $r + 1 \equiv N$ of these new variables z_i, the spectral parameters, as opposed to the r simple roots, is a reflection of the usual embedding $\mathfrak{sl}(N) \subset \mathfrak{gl}(N)$. $\mathfrak{b}$ (resp. $\mathfrak{n}$) is simply the space of upper triangular (resp. strictly upper triangular) matrices of size N, and the orbital varieties under consideration are the irreducible components of the scheme $\{M \in \mathfrak{n} \mid M^2 = 0\}$. We also restrict ourselves to the case of $N = 2n$ even, which is technically simpler.

3.1. Orbital varieties and Temperley–Lieb algebra. In general, $\mathfrak{sl}(N)$ nilpotent orbits are classified by their Jordan decomposition type, which can be expressed as a Young diagram; the orbital varieties are then indexed by Standard Young Tableaux (SYT). The condition $M^2 = 0$ ensures that only Young diagrams with at most 2 rows can appear (blocks in the Jordan decomposition are of size at most 2), and it is easy to check that all orbits are in the closure of the largest orbit, whose Young diagram is of the form (n, n). It is convenient to describe the corresponding SYT by "link patterns", that is N points on a line connected in the upper-half plane via n non-intersecting arches, see fig. 1. The numbers in the first (resp. second) row of the SYT are the labels of the openings (resp. closings) of the arches. There are $\frac{(2n)!}{n!(n+1)!}$ such configurations.

In this language, one has a rather convenient description of orbital varieties [**25, 26**], which we mention for the sake of completeness. Indeed, to each orbital variety π we associate the upper triangular matrix $\pi^<$ with $\pi^<_{ij} = 1$ if points labelled i and j are connected by an arch, $i < j$, 0 otherwise. Then $\pi = \overline{B \cdot \pi^<}$,

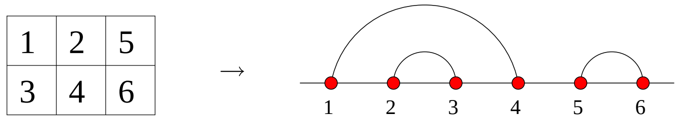


FIGURE 1. A Standard Young Tableau and the corresponding link pattern.

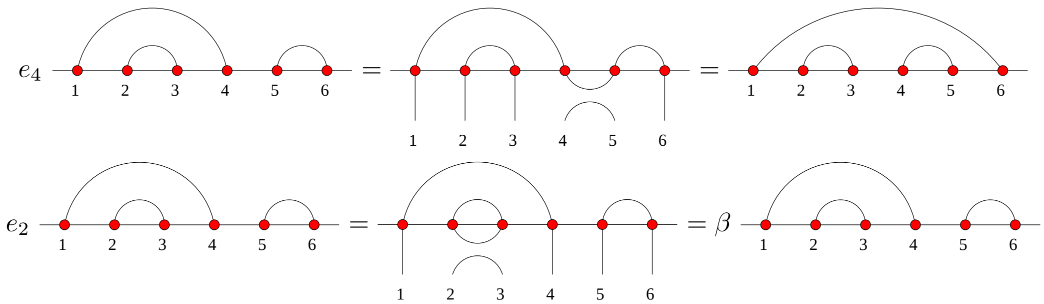


FIGURE 2. Action of the Temperley–Lieb algebra $TL(\beta)$ on link patterns.

B acting by conjugation. Equivalently, π is given by the following set of equations: (i) $M^2 = 0$ and (ii) $r_{ij}(M) \le r_{ij}(\pi^<)$, $i, j = 1, \ldots, N$, where r_{ij} is the rank of the $i \times j$ lower-left rectangle.

It is equally simple to describe the action of the Weyl group, namely the symmetric group S_N. Rather than the generators corresponding to the simple roots: $s_i \equiv s_{\alpha_i}$, $i = 1, \ldots, r$ used so far, it proves simpler to consider the action of the projectors $e_i = 1 - s_i$ in the symmetric group algebra. The operator e_i acts on link patterns π by connecting the arches ending at i and $i+1$ and creates a new little arch between these 2 points; this action is described on Fig. 2. When a closed loop is formed, it is erased but contributes a weight $\beta = 2$. The q-deformed version of this is obtained by attaching a weight $\beta = -(q + q^{-1})$ to each erased loop, thus leading to the following (pictorially clear) relations:

$$e_i^2 = \beta e_i \qquad e_i = e_i e_{i\pm 1} e_i \qquad [e_i, e_j] = 0 \quad |i - j| > 1 \tag{3.1}$$

all indices taking values in $1, \ldots, r$. These are the defining relations of the Temperley–Lieb algebra $TL_{r+1}(\beta)$. When $q = -1$, i.e. $\beta = 2$, it is simply a quotient of the symmetric group algebra. Alternatively, the deformed generators $s_i = -q^{-1} - e_i$ satisfy the usual relations of the Hecke algebra (of which the Temperley–Lieb algebra is a quotient).

In what follows, one special element of $TL_N(\beta)$ will be needed: it is the cyclic rotation S. Its effect is to rotate the endpoints of the link patterns: $1 \to 2 \to \cdots \to N \to 1$ without changing their connectivity. It can also be expressed as: $S = q^{n-2} s_1 \cdots s_{N-1}$.

3.2. qKZ equation. For each simple root α_i, we have the trigonometric R-matrix:

$$R_i(w) \equiv R_{\alpha_i}(w) = \frac{(qw - q^{-1}) + (w - 1)e_i}{q - q^{-1}w}, \tag{3.2}$$

where the $e_i = -q^{-1} - s_i$ generate $TL_N(\beta)$ and act in the space of link patterns as explained above. We first write the system of equations:

$$R_i(w_{i+1}/w_i)\Psi = \tau_i \Psi \qquad i = 1, \ldots, N-1 \tag{3.3}$$

where $\tau_i \equiv \tau_{\alpha_i}$ acts by interchanging multiplicative spectral parameters $w_i := e^{-\hbar z_i}$ and w_{i+1} in the polynomial Ψ of the w's, homogeneous of degree $n(n-1)$.

These equations are supplemented by the "affinized" equation satisfied by Ψ. Since the affine Dynkin diagram $A_r^{(1)}$ is a circular chain, this equation quite naturally involves the cyclic rotation S. Define the operator ρ on $\mathbb{C}[w_1, \ldots, w_N]$ which shifts the variables w_i according to the rule: $w_i \to w_{i+1}$, $i = 1, \ldots, N-1$ and $w_N \to q^6 w_1$. Then the additional equation is

$$q^{3(n-1)} S^{-1} \Psi = \rho \Psi \tag{3.4}$$

Together with this equation, the above system forms the so-called level one qKZ equation.

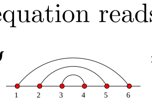
We claim that the $\mathbf{R}_i := \tau_i R_i(w_{i+1}/w_i)$ and $\mathbf{S} := q^{3(1-n)}\rho S$ generate together $\hat{W}$. In order to see that, it is sufficient to build the N generators $\mathbf{T}_i$ of the abelian subgroup (the lattice of weights). They are given by $\mathbf{T}_i = \mathbf{R}_{i-1}\mathbf{R}_{i-2}\cdots\mathbf{R}_1\mathbf{S}\mathbf{R}_{N-1}\cdots\mathbf{R}_{i+1}\mathbf{R}_i$, $i = 1,\ldots,N$. The original definition of the qKZ equation is in fact the eigenvector equation for these "scattering" matrices; with reasonable assumptions it is equivalent to the above system. Also, note that if one defines $\mathbf{R}_N := \mathbf{S}^{-1}\mathbf{R}_1\mathbf{S}$, then the $\mathbf{R}_i$, $i = 1,\ldots,N$ generate the usual affine Weyl group (a subgroup of order N of $\hat{W}$).

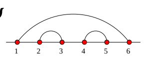
The minimal degree polynomial solution of the level one qKZ equation was obtained in [**6, 7**], and is characterized by its "base" entry Ψ_{π_0} corresponding to the link pattern π_0 that connects points $i \leftrightarrow 2n+1-i$, with the value

$$\Psi_{\pi_0} = \prod_{1\le i<j\le n}(qw_i - q^{-1}w_j)\prod_{n+1\le i<j\le 2n}(qw_i - q^{-1}w_j) \tag{3.5}$$

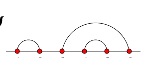
in which all factors are a direct consequence of the $\tau_i\Psi = R_i\Psi$ equations. It is then easy to prove that all the other entries of Ψ may be obtained from Ψ_{π_0} in a triangular way.

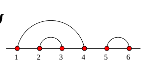
Example: at $N = 6$, there are 5 link patterns. The minimal degree polynomial solution of the level one qKZ equation reads:

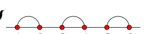
Ψ  $= (qw_1 - q^{-1}w_2)(qw_2 - q^{-1}w_3)(qw_1 - q^{-1}w_3)(qw_4 - q^{-1}w_5)(qw_5 - q^{-1}w_6)(qw_4 - q^{-1}w_6)$

Ψ  $= (qw_1 - q^{-1}w_2)(qw_3 - q^{-1}w_4)(qw_5 - q^{-1}w_6)$

$\times\Big((w_1+w_2)(q^2w_3w_4 - q^{-2}w_5w_6) - (w_3+w_4)(q^4w_1w_2 - q^{-4}w_5w_6) + (w_5+w_6)(q^2w_1w_2 - q^{-2}w_3w_4)\Big)$

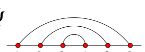
Ψ  $= (qw_2 - q^{-1}w_3)(qw_2 - q^{-1}w_4)(qw_3 - q^{-1}w_4)(qw_5 - q^{-1}w_6)(q^{-2}w_6 - q^2w_1)(q^{-2}w_5 - q^2w_1)$

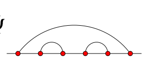
Ψ  $= (qw_1 - q^{-1}w_2)(qw_3 - q^{-1}w_4)(qw_4 - q^{-1}w_5)(qw_3 - q^{-1}w_5)(q^{-2}w_6 - q^2w_1)(q^{-2}w_6 - q^2w_2)$

Ψ  $= (qw_2 - q^{-1}w_3)(qw_4 - q^{-1}w_5)(q^{-2}w_6 - q^2w_1)$

$\times\Big((q^3w_1 + q^{-3}w_6)(q^2w_2w_3 - q^{-2}w_4w_5) - (w_2+w_3)(qw_1w_6 - q^{-1}w_4w_5) - (w_4+w_5)(qw_2w_3 - q^{-1}w_1w_6)\Big)$

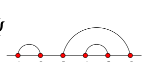
Performing the rational limit $\hbar \to 0$, $z_i = e^{-\hbar w_i}$, $q = -e^{-\hbar A/2}$ yields the following multidegrees:

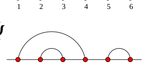
Ψ  $= (A+z_1-z_2)(A+z_2-z_3)(A+z_1-z_3)(A+z_4-z_5)(A+z_5-z_6)(A+z_4-z_6)$

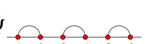
Ψ  $= (A+z_1-z_2)(A+z_3-z_4)(A+z_5-z_6)\Big(4A^3 + 3A^2(z_1+z_2-z_5-z_6) +$

$+A(2(z_1z_2 - 2z_3z_4 - z_1z_5 - z_2z_5 - z_1z_6 - z_2z_6 + z_5z_6) + (z_3+z_4)(z_1+z_2+z_5+z_6))$

$+(z_1+z_2)(z_5z_6 - z_3z_4) + (z_3+z_4)(z_1z_2 - z_5z_6) + (z_5+z_6)(z_3z_4 - z_1z_2)\Big)$

Ψ  $= (A+z_2-z_3)(A+z_2-z_4)(A+z_3-z_4)(A+z_5-z_6)(2A+z_1-z_6)(2A+z_1-z_5)$

Ψ  $= (A+z_1-z_2)(A+z_3-z_4)(A+z_4-z_5)(A+z_3-z_5)(2A+z_1-z_6)(2A+z_2-z_6)$

Ψ  $= (A+z_2-z_3)(A+z_4-z_5)(2A+z_1-z_6)\Big(5A^3 + 3A^2(z_1+z_2+z_3-z_4-z_5-z_6) +$

$+A(2z_1(z_2+z_3-z_6) + z_2z_3 + z_4z_5 - (z_2+z_3)z_6 + (z_4+z_5)(2z_6 - z_1 - z_2 - z_3))$

$+(z_1+z_6)(z_2z_3 - z_4z_5) + (z_2+z_3)(z_4z_5 - z_1z_6) + (z_4+z_5)(z_1z_6 - z_2z_3)\Big)$

and in particular the degrees $1, 4, 4, 4, 10$ respectively, upon taking $z_i = 0$ and $A = 1$.

3.3. Razumov–Stroganov point and ASM. At $q = e^{2i\pi/3}$, Ψ becomes the ground state eigenvector of the integrable transfer matrix with periodic boundary conditions and inhomogeneities $w_1, \ldots, w_N$, or equivalently of the scattering matrices $\mathbf{T}_i = R_{i-1}(w_{i-1}/w_i)\cdots R_1(w_1/w_i)SR_{N-1}(w_{N-1}/w_i)\cdots R_i(w_{i+1}/w_i)$. Consider now the particular case $w_1 = \cdots = w_N = 1$, when Ψ is the Perron–Frobenius eigenvector of the

Hamiltonian $H = e_1 + \cdots + e_N$ where $e_N = S^{-1} e_1 S$. Note that the periodic boundary conditions mean that H is cyclic-invariant: $SH = HS$. Normalizing Ψ so that its smallest entry Ψ_{π_0} is 1, we have the following

THEOREM. [**4**] The sum of entries $\sum_\pi \Psi_\pi$ is equal to the number of Alternating Sign Matrices, $A(n)$.

The result of [**4**] is actually much more general, as the sum $\sum_\pi \Psi_\pi$ was evaluated in the presence of all the spectral parameters w_i, and identified with proper normalization to the so-called Izergin–Korepin determinant [**20, 21**], also equal to a particular Schur function [**22**]. Still unproven, however, is the

CONJECTURE. [**1**] The largest entry of Ψ, with arches connecting consecutive points, is $A(n-1)$.

For instance, plugging $w_i = 1$ and $q = e^{2i\pi/3}$ into the above example, we get for $N = 6$, $\Psi = (1, 2, 1, 1, 2)$ and $\sum_\pi \Psi_\pi = 7 = A(3)$, the total number of 3×3 ASMs.

4. B_r case

We now develop the B_r case, which allows us to recover and interpret geometrically the results of [**16**]. We concentrate on the even case $r = 2n$. We parametrize as usual the roots $\alpha_i = z_i - z_{i+1}$ for $i = 1, 2, \ldots, r-1$ and $\alpha_r = z_r$.

We consider matrices that square to zero in the fundamental representation of dimension $N = 2r + 1$: a possible choice is to select upper triangular matrices satisfying $M^T J + JM = 0$, J antidiagonal matrix with 1's on the second diagonal. It turns out that the orbital varieties are indexed by the same link patterns as before, of size r; and that the Weyl group representation is actually a representation of the same quotient, the Temperley–Lieb algebra $TL_r(\beta)$, the additional reflection s_r being represented by a multiple of the identity.

4.1. B-type qKZ equation. According to the dicusssion above, the B qKZ system reads:

$$R_i(w_{i+1}/w_i)\Psi = \tau_i \Psi, \quad i = 1, 2, ..., r-1 \tag{4.1}$$

$$w_r^{-m_r} \frac{q^{-1} w_r - q}{q^{-1} - q w_r} \Psi = \tau_r \Psi \tag{4.2}$$

where τ_r stands for the inversion of the last spectral parameter, namely $\tau_r \Psi(w_1, ...w_{r-1}, w_r) = \Psi(w_1, ..., w_{r-1}, 1/w_r)$ and m_r is the degree of Ψ in w_r.

Finally, these equations are to be supplemented by the affinization relation. The latter is expressed by considering the reflection with respect to the extra root z_1. One finds that

$$(q^3 w_1)^{-m_1} \frac{q^{-2} - q^2 w_1}{q w_1 - q^{-1}} \Psi(w_1, w_2, ..., w_r) = \Psi\Big(\frac{1}{q^6 w_1}, w_2, ..., w_r\Big) \tag{4.3}$$

where m_1 is the degree of Ψ in z_1.

Introducing the boundary operators $\mathbf{K}_1$ and $\mathbf{K}_r$ so that Eqs. (4.2–4.3) reduce to $\mathbf{K}_1 \Psi = \mathbf{K}_2 \Psi = \Psi$, as well as the usual $\mathbf{R}_i = \tau_i R_i(w_{i+1}/w_i)$, the generators of the weight lattice (as abelian subgroup of $\hat{W}$) are: (i) $\mathbf{T}_i = \mathbf{R}_i \mathbf{R}_{i+1} \cdots \mathbf{R}_{r-1} \mathbf{K}_r \mathbf{R}_{r-1} \cdots \ \mathbf{R}_1 \mathbf{K}_1 \mathbf{R}_1 \cdots \mathbf{R}_{i-1}$ that implements $w_i \to q^6 w_i$ and (ii) one additional generator implementing $w_i \to q^3 w_i$ simultaneously for all i. The latter is a combination of $\mathbf{R}$ and $\mathbf{K}$ matrices as well as an additional operator implementing the reflection $w_i \leftrightarrow q^{-3}/w_{r+1-i}$ for all i.

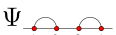
The minimal polynomial solution to the system (4.1–4.3) has degree $m_1 = m_r = r - 1 = 2n - 1$ in each spectral parameter and total degree $n(3n-1)$. As before it has a simple factorized base entry

$$\Psi_{\pi_0} = C \prod_{1 \le i < j \le n} (q w_i - q^{-1} w_j)(q^{-2} - q^2 w_i w_j) \prod_{n+1 \le i < j \le 2n} (q w_i - q^{-1} w_j)(q w_i w_j - q^{-1}) \tag{4.4}$$

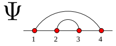
where $C = 2^n \prod_{i=1}^r (q w_i - q^{-1})$ is a common (symmetric) factor to all entries of Ψ. All other entries may be obtained from this one in a triangular manner.

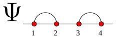
Example: For B_4, there are 2 link patterns as for the case A_3. The minimal degree polynomial solution of the level one B_4 qKZ equation reads:

Ψ  $= C(q w_1 - q^{-1} w_2)(q^{-2} - q^2 w_1 w_2)(q w_3 - q^{-1} w_4)(q w_3 w_4 - q^{-1})$

Ψ  $= C(q w_2 - q^{-1} w_3)(q^{-1} w_1 - q w_1 w_2 w_3 - q^{-5} w_4 - q w_1^2 w_4 + (q^{-1} - q) w_1 (w_2 + w_3) w_4$
$+ q^{-1} w_2 w_3 w_4 + q^5 w_1^2 w_2 w_3 w_4 + q^{-1} w_1 w_4^2 - q w_1 w_2 w_3 w_4^2)$

As before, we get the corresponding multidegrees upon taking the rational limit, with the result:

Ψ  $= C'(A+z_1-z_2)(2A+z_1+z_2)(A+z_3-z_4)(A+z_3+z_4)$

Ψ  $= C'(A+z_2-z_3)\Big(5A^3+3A^2(2z_1+z_2+z_3)+A(2z_1^2+3z_1(z_2+z_3)+z_2z_3-z_4^2)$
$+(z_2+z_3)(z_1^2-z_4^2)\Big)$

with $C'=4(A+z_1)(A+z_2)(A+z_3)(A+z_4)$; hence the degrees $4\times 2, 4\times 5$ for $A=1$ and $z_i=0$.

4.2. RS point, VSASM and CSTCPP. As explained in Sect. 2, the case $q=e^{2i\pi/3}$ is special in that the problem admits a transfer matrix, and its solution Ψ in the homogeneous limit where all $w_i=1$ is the groundstate of a Hamiltonian

$$H_B = e_1+e_2+\ldots+e_{N-1} \tag{4.5}$$

which is the open boundary version of the A_r Hamiltonian H.

As shown in [**17**], at the RS point $q=e^{2i\pi/3}$, and in the homogeneous limit where $w_i=1$ for all i, and in which Ψ is normalized so that its smallest entry is $\Psi_{\pi_0}=1$, we have the following

THEOREM. [**16**] The sum of entries $\sum_\pi \Psi_\pi$ is equal to the number of Vertically Symmetric Alternating Sign Matrices (VSASM), $A_V(2n+1)$.

This was actually proved in the same spirit as for the A_r case, by identifying the sum of components including all spectral parameters w_i as yet another determinant, which takes the form of a particular symplectic Schur function. A similar result holds for the case of odd $r=2n-1$, namely once properly normalized, the sum of entries $\sum_\pi \Psi_\pi$ is equal to an integer we call $A_V(2n)$ by analogy. It turns out that $A_V(2n)=N_8(2n)$ is the number of Cyclically Symmetric Transpose Complement Plane Partitions (CSTCPP) in an hexagon of size $2n\times 2n\times 2n$ [**24**]. The numbers $A_V(i)$ both have determinant formulae, namely $A_V(2n)=\det\binom{i+j}{2i-j}_{0\le i,j\le n-1}$, and $A_V(2n+1)=\det\binom{i+j+1}{2i-j}_{0\le i,j\le n-1}$.

As in the A case, we have the

CONJECTURE. [**1**] The largest entry of Ψ, with arches connecting consecutive points, is $A_V(r)$.

Example: for $r=2n=4$, taking $w_i\to 1$ and $q=e^{2i\pi/3}$ in the above expressions, we get the components $\Psi=(1,2)$, which sum to $3=A_V(5)$, the number of 5×5 VSASMs, and the maximal entry of Ψ is $2=N_8(4)$.

5. C_r case

The simple roots of C_r are $\alpha_i=z_i-z_{i+1}$, $i=1,2,\ldots,r-1$ and $\alpha_r=2z_r$. We concentrate on the odd case $r=2n+1$, and consider the fundamental representation of dimension $N=2r$. One choice is to select upper triangular matrices satisfying $M^TJ+JM=0$, J antidiagonal matrix with 1's (resp. -1's) in the upper (resp. lower) triangle.

5.1. Orbital varieties and C-type Temperley–Lieb algebra. There are $\binom{r}{\lfloor\frac{r+1}{2}\rfloor}$ orbital varieties, which are now indexed by *open link patterns*, that is configurations of r points on a line connected in the upper-half plane either in pairs via (closed) arches or to infinity via half-lines (open arches).

The representation of the Weyl group on these open link patterns takes the form of a modified Temperley–Lieb algebra. We describe now its q-deformed version, $CTL(\beta)$ (see also [**23**] for other variants of Temperley–Lieb algebra). The generators $e_1,e_2,\ldots,e_{r-1}$ obey the standard $TL(\beta)$ relations (3.1) and the additional "boundary" generator e_r satisfies: $e_r^2=\beta e_r$, $e_{r-1}e_re_{r-1}=2e_{r-1}$.

These generators act on open link patterns as follows. Open link patterns are represented with their open arches connected to a vertical line on the right. The e_i, $i=1,2,...,r-1$ act as usual, and e_r like the left half of an e, connecting the point $2n+1$ to the vertical line (first line of Fig. 3). The rule is that any loop may be erased and replaced by a factor β. Moreover, whenever a connection between points on the vertical line (consecutive open arches) is created, they may also be erased and replaced by a factor β (resp. 2) if this is created by the action of some e_{2i-1} (resp. e_{2i}). As r is odd, the loop created by e_r^2 yields a weight β, while that created by $e_{r-1}e_re_{r-1}$ yields a weight 2, hence the result $2e_{n-1}$ (second line of Fig. 3).

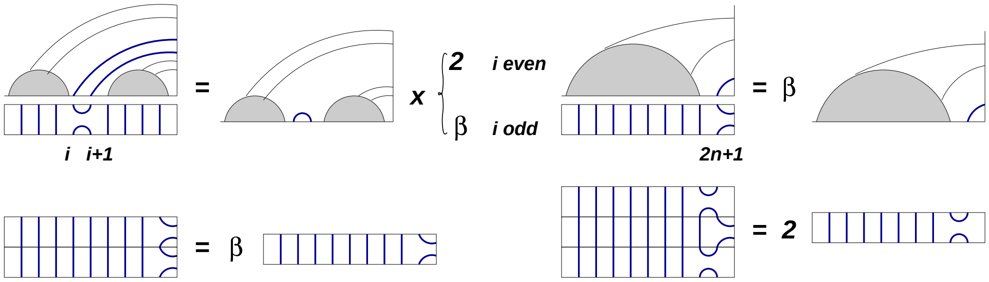


FIGURE 3. The rule for erasing arches at infinity when acting with e_i: they are replaced by a factor 2 (resp. β) according to whether the index i is even (resp. odd). We have also represented the case $i = 2n+1$ (first line), and the resulting boundary relations $e_r^2 = \beta e_r$ and $e_{r-1}e_re_{r-1} = 2e_{r-1}$ (second line).

We shall also need an additional operator e_1' satisfying the relations: $(e_1')^2 = \beta e_1'$ and $e_1e_1' = e_1'e_1 = e_1'e_2e_1' - e_1' = e_2e_1'e_2 - e_2 = 0$. It is defined as $e_1' = se_1s$, where s is the involution acting on link patterns as follows: (i) $s\pi = \pi$ if the arch connected to point 1 is open, and (ii) $s\pi = -\pi + \pi'$ otherwise, where π' is the link pattern in which the closed arch connected to 1 is cut into two open arches.

5.2. C-type qKZ equation. To each simple root we attach respectively the standard trigonometric R-matrices $R_i(w_{i+1}/w_i)$, $i = 1, 2, \ldots, r-1$ of Eq. (3.2), and the boundary R-matrix $R_r(1/w_r^2) \equiv R_{\alpha_r}$, with the same expression.

The level one C qKZ equation consists of the following system

$$R_i(w_{i+1}/w_i)\Psi \quad = \tau_i\Psi \tag{5.1}$$

$$w_r^{-m_r}R_r(1/w_r^2)\Psi \quad = \tau_r\Psi \tag{5.2}$$

where as usual τ_i acts by interchanging the spectral parameters w_i and w_{i+1}, $i = 1, 2, ..., r-1$ and τ_r acts on Ψ by letting $w_r \to 1/w_r$, and m_r is the degree of Ψ in w_r.

These are finally supplemented by the affinization relation, obtained by considering an extra root, say $\alpha_1' = -z_1 - z_2$, and the associated boundary operator $R_1'(q^6w_1w_2)$:

$$R_1'(q^6w_1w_2)\Psi = \tau_1'\Psi \tag{5.3}$$

where τ_1' interchanges w_2 and $1/(q^6w_1)$, and R_1' is of the form of Eq. (3.2) with e_1' in place of e_i. Using $R_1'(w) = sR_1(w)s$, the relation can also be recast into

$$(q^3z_1)^{-m_1}s\Psi(w_1, \ldots, w_r) = \Psi\Big(\frac{1}{q^6w_1}, w_2, \ldots, w_r\Big) \tag{5.4}$$

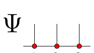
The generators of the weight lattice (as abelian subgroup of $\hat{W}$) are very similar to the generators (i) of the case B_r: the only change concerns the boundary operators $\mathbf{K}_1$ and $\mathbf{K}_r$ now implementing Eqs. (5.2) and (5.4).

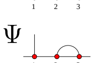
The polynomial solution Ψ to the level one C_r qKZ system has degree $m_1 = m_r = 2n$ in each variable, total degree $n(2n+1)$ and base entry

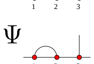
$$\Psi_{\pi_0} = \prod_{1\le i<j\le 2n+1} (qz_i - q^{-1}z_j) \tag{5.5}$$

and all the other entries of Ψ may be obtained in a triangular way from this one.

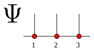
Example: for $r = 3$, we have the following minimal polynomial solution to the level one C_3 qKZ system:

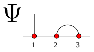
Ψ  $= (qw_1 - q^{-1}w_2)(qw_1 - q^{-1}w_3)(qw_2 - q^{-1}w_3)$

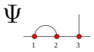
Ψ  $= (qw_1 - q^{-1}w_2)(q^2w_1w_2 - q^{-2})(q^{-1} - qw_3^2)$

Ψ  $= (q^3w_1^2 - q^{-3})(qw_2 - q^{-1}w_3)(qw_2w_3 - q^{-1})$

which, upon taking the rational limit yields the multidegrees:

Ψ  $= (A + z_1 - z_2)(A + z_1 - z_3)(A + z_2 - z_3)$

Ψ  $= (A + z_1 - z_2)(2A + z_1 + z_2)(A + 2z_3)$

Ψ  $= (3A + 2z_1)(A + z_2 - z_3)(A + z_2 + z_3)$

and the degrees $\Psi = (1, 2, 3)$ for $A = 1$ and $z_i = 0$.

5.3. RS point and CSSCPP. At the point $q = e^{2i\pi/3}$, Ψ may be viewed as the ground state eigenvector of a transfer matrix, corresponding in the homogeneous limit to the Hamiltonian

$$H_C = \frac{e_1 + e_1'}{2} + \sum_{i=2}^{r-1} e_i + e_r \tag{5.6}$$

Normalizing Ψ so that its smallest entry $\Psi_{\pi_0} = 1$, we have been able to compute the sum of entries to be $A(n)A(n+1)$. In the case of even $r = 2n$, the above may be repeated almost identically: in the presence of spectral parameters, the even case may be recovered from the odd one by taking $w_{2n+1} \to -q^{-1}$, and dividing out the result by $\prod_{1 \le i \le 2n}(1 + q^3 w_i)$. Indeed, this specialization leaves us with only non-vanishing components whith an open arch at the rightmost point, in bijection with open link patterns with that point erased, hence the projection onto the case of size one less. This leads us to the

Conjecture.

$$\sum_\pi \Psi_\pi = A(\lfloor r/2 \rfloor) A(\lceil r/2 \rceil) \tag{5.7}$$

Note that the sum in the even case, $A(n)^2$, also counts the Cyclically Symmetric Self-Complementary Plane Partitions (CSSCPP) in an hexagon of size $2n \times 2n \times 2n$ [**24**]. Also note the determinant formulae $A(n)^2 = \det\left(\binom{i+j}{2i-j-1} + \binom{i+j+1}{2i-j}\right)_{0 \le i,j \le n-1}$ and $A(n)A(n+1) = \det\left(\binom{i+j+1}{2i-j} + \binom{i+j+2}{2i-j}\right)_{0 \le i,j \le n-1}$.

Furthermore, consider the left eigenvector v of H_C with the same eigenvalue (r for r odd, $r + 1/2$ for r even). Normalize v so that its entries are coprime positive integers. We have found empirically the following

Conjecture.

$$\sum_\pi v_\pi \Psi_\pi = A(r) \ . \tag{5.8}$$

Finally, we formulate the

Conjecture. The largest entry of Ψ for C_r is the sum of entries for C_{r-1}.

Example: at $r = 5$, $\Psi = (1, 2, 3, 3, 0, 1, 4, 0, 0, 0)$, $v = (48, 36, 28, 34, 24, 23, 25, 18, 17, 14)$, $\sum_\pi \Psi_\pi = 14 = 2 \times 7 = A(2)A(3)$, $\sum_\pi v_\pi \Psi_\pi = 429 = A(5)$, and the maximal entry of Ψ is $4 = A(2)^2$.

6. D_r case

The simple roots of D_r are $\alpha_i = z_i - z_{i+1}$ for $i = 1, 2, \ldots, n-1$ and $\alpha_r = z_{r-1} + z_r$. We concentrate on the odd case $r = 2n + 1$, and consider again the fundamental representation of dimension $N = 2r$. Just like in the B_r case, one choice is to select upper triangular matrices satisfying $M^T J + JM = 0$, J antidiagonal matrix with 1's on the second diagonal.

6.1. Orbital varieties and D-type Temperley–Lieb algebra. Just as in the case C, there are $\binom{r}{\lfloor \frac{r+1}{2} \rfloor}$ orbital varieties, indexed by open link patterns.

We now deal with D-type Temperley–Lieb algebras, denoted $DTL(\beta)$, with generators e_i, $i = 1, 2, ..., r-1$ obeying the $TL(\beta)$ relations (3.1) and an extra generator e'_{r-1}, satisfying the relations:

$$(e'_{r-1})^2 = \beta e_{r-1}, \qquad e_{r-1}e'_{r-1} = e'_{r-1}e_{r-1} = e_{r-2}e'_{r-1}e_{r-2} - e_{r-2} = e'_{r-1}e_{r-2}e'_{r-1} - e'_{r-1} = 0 \tag{6.1}$$

These operators act on open link patterns as follows. The e_i, $i = 1, 2, \ldots, r-1$ act in the usual way, by creating a little arch between points i and $i + 1$ and by gluing the two former points. To describe the

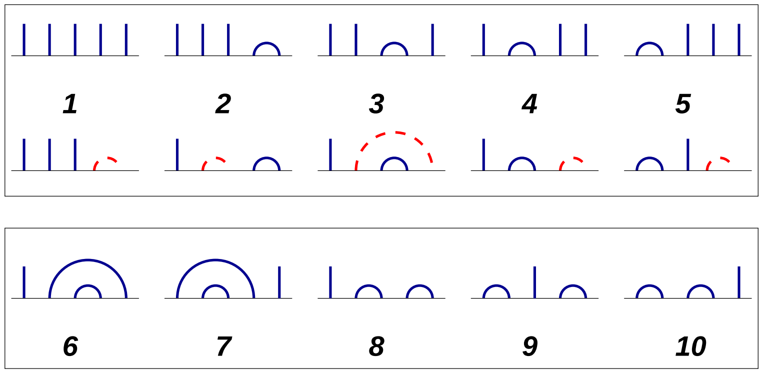


Figure 4. The ten open link patterns for D_5. In the second line, we have transformed the open link patterns by connecting the two rightmost open arches into a (dashed) arch. The involution s simply switches the color of the rightmost arch (if it is closed) in this representation, namley exchanges $1 \leftrightarrow 2$, $3 \leftrightarrow 6$, $4 \leftrightarrow 8$, $5 \leftrightarrow 9$, and leaves 7 and 10 invariant (as their rightmost arch is open).

action of e'_{r-1}, let us first connect the open arches of the open link patterns by pairs of consecutive open arches from the left to the right, and represent the newly formed arches in a different color (dashed lines, cf Fig. 4 for the D_5 example). We then define an involution s on open link patterns that simply switches the color (solid $\leftrightarrow$ dashed) of the rightmost arch if it is closed, and leaves it invariant if it is open. Then $e'_{r-1} = s e_{r-1} s$.

Finally, we introduce an extra boundary operator e_0, which is the right half of an e (like a reflected e_r of C_r), with its open end connected to the vertical line, and acts as such, with the same rules as for C_r, but upon reflection of indices $i \leftrightarrow r-i$. It satisfies the relations: $e_0^2 = \beta e_0$ and $e_1 e_0 e_1 = 2e_1$.

6.2. ***D*-type *q*KZ equation.** We associate to the roots the R-matrices $R_i(w_{i+1}/w_i)$ of Eq. (3.2), and $R_r(1/(w_r w_{r-1}))$ defined by the same equation in which e_i is replaced with e'_{r-1}, so that $R_r(w) = sR_{r-1}(w)s$.

The level one D qKZ equation consists of the following system

$$\begin{aligned} R_i(w_{i+1}/w_i)\Psi &= \tau_i \Psi, \qquad i = 1, 2, ..., r-1 \\ R_r(1/(w_r w_{r-1}))\Psi &= \tau'_{r-1}\Psi \end{aligned}$$

where as usual τ_i acts by interchanging the spectral parameters w_i and w_{i+1}, $i = 1, 2, ..., r-1$ and τ'_r acts on Ψ by interchanging w_{r-1} and $1/w_r$. Upon using the above relation $e'_{r-1} = s e_{r-1} s$, the latter equation may be equivalently replaced by

$$z_r^{-m_r} s\Psi(z_1, \ldots, z_r) = \Psi\Big(z_1, \ldots, z_{r-1}, \frac{1}{z_r}\Big) \tag{6.2}$$

These are finally supplemented by the affinization relation, obtained by considering the extra root $\alpha_0 = -2z_1$, and the associated boundary operator $R_0(q^6 w_1^2)$ involving the extra operator e_0:

$$w_1^{-m_1} R_0(q^6 w_1^2)\Psi = \tau_0 \Psi \tag{6.3}$$

where $\tau_0 f(w_1) = f(1/(q^6 w_1))$ and m_1 the degree of Ψ in w_1.

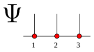
The construction of the abelian subgroup of $\hat{W}$ is similar to the cases B and C, and is skipped for the sake of brevity.

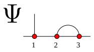
The minimal degree polynomial solution to the level one D_r qKZ system has total degree $r(r-1)/2$ and partial degree $m_1 = m_r = r-1$ in all variables. Its base entry, corresponding to the open link pattern π_0 with only open arches reads

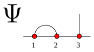
$$\Psi_{\pi_0} = \prod_{1 \le i < j \le 2n+1} (qz_i - q^{-1} z_j) \tag{6.4}$$

and all the other entries of Ψ may be obtained in a triangular way from this one.

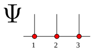
Example: for $r = 3$, we have the following minimal polynomial solution to the level one D_3 qKZ system:

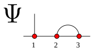
Ψ  $= (qw_1 - q^{-1}w_2)(qw_1 - q^{-1}w_3)(qw_2 - q^{-1}w_3)$

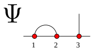
Ψ  $= (qw_1 - q^{-1}w_2)(qw_1w_3 - q^{-1})(qw_2w_3 - q^{-1})$

Ψ  $= (q^{-2} - q^2w_1^2)(qw_2 - q^{-1}w_3)(qw_2w_3 - q^{-1})$

which, upon taking the rational limit gives the multidegrees:

Ψ  $= (A + z_1 - z_2)(A + z_1 - z_3)(A + z_2 - z_3)$

Ψ  $= (A + z_1 - z_2)(A + z_1 + z_3)(A + z_2 + z_3)$

Ψ  $= 2(A + z_1)(A + z_2 - z_3)(A + z_2 + z_3)$

and the degrees $\Psi = (1, 1, 2)$ for $A = 1$ and $z_i = 0$.

6.3. RS point and HTASM. At the point $q = e^{2i\pi/3}$, Ψ may be viewed as the Perron–Frobenius eigenvector of a transfer matrix, corresponding in the homogeneous limit to the Hamiltonian

$$H_D = e_0 + \sum_{i=1}^{r-2} e_i + \frac{e_{r-1} + e'_{r-1}}{2} \tag{6.5}$$

Note that upon the reflection $e_i \to e_{r-i}$, this Hamiltonian is mapped onto H_C: we are dealing with the same algebra, but in different representations.

Going to the RS point $q = e^{2i\pi/3}$ and taking the homogeneous limit $w_i = 1$ for all i, and normalizing Ψ so that its smallest entry is $\Psi_{\pi_0} = 1$, we have found the

Conjecture. The sum of entries $\sum_\pi \Psi_\pi$ is the number of Half-Turn Symmetric Alternating Sign Matrices of size r, $A_{HT}(r)$.

This conjecture also works in the even case $r = 2n$, which may be obtained from the odd one by taking $z_1 = -q^{-2}$, shifting all remaining spectral parameters $w_i \to w_{i-1}$, $i = 2, 3, ..., 2n+1$, and dividing out by $\prod_{1\le i\le 2n}(1 + z_i)$. Note the formulae $A_{HT}(2n) = \det\left(\binom{i+j}{2i-j} + \binom{i+j+1}{2i-j}\right)_{0\le i,j\le n-1}$ and $A_{HT}(2n+1) = \det\left(\binom{i+j+1}{2i-j} + \binom{i+j+2}{2i-j+1}\right)_{0\le i,j\le n-1}$.

Introduce as before the left Perron–Frobenius eigenvector v of H_D with coprime positive integer entries.

Conjecture.

$$\sum_\pi v_\pi \Psi_\pi = A(r)\ . \tag{6.6}$$

Finally, we also find the

Conjecture. The largest entry of Ψ for D_r is the sum of entries for C_{r-1}.

Example: at $r = 5$, $\Psi = (1, 1, 3, 4, 2, 3, 1, 4, 2, 4)$, $v = (10, 10, 17, 14, 18, 17, 23, 14, 18, 25)$, $\sum_\pi \Psi_\pi = 25 = A_{HT}(5)$, $\sum_\pi v_\pi \Psi_\pi = 429 = A(5)$, and the maximal entry of Ψ is $4 = A(2)^2$, the sum of the components of the C_4 solution.

References

[1] M.T. Batchelor, J. de Gier and B. Nienhuis, The quantum symmetric XXZ chain at $\Delta = -1/2$, alternating sign matrices and plane partitions, *J. Phys.* A34 (2001) L265–L270, cond-mat/0101385.

[2] A.V. Razumov and Yu.G. Stroganov, Spin chains and combinatorics, *J. Phys* A34 (2001), 3185, cond-mat/0012141; Spin chains and combinatorics: twisted boundary conditions, *J. Phys* A34 (2001), 5335, cond-mat/0012247.

[3] A.V. Razumov and Yu.G. Stroganov, Combinatorial nature of ground state vector of $O(1)$ loop model, *Theor. Math. Phys.* **138** (2004) 333–337; *Teor. Mat. Fiz.* 138 (2004) 395–400, math.CO/0104216.

[4] P. Di Francesco and P. Zinn-Justin, Around the Razumov–Stroganov conjecture: proof of a multi-parameter sum rule, *E. J. Combi.* 12 (1) (2005), R6, math-ph/0410061.

[5] G. Kuperberg, Another proof of the alternating sign matrix conjecture, *Int. Math. Research Notes* (1996) 139–150, math.CO/9712207.
[6] V. Pasquier, Quantum incompressibility and Razumov Stroganov type conjectures, cond-mat/0506075.
[7] P. Di Francesco and P. Zinn-Justin, Quantum Knizhnik–Zamolodchikov equation, generalized Razumov–Stroganov sum rules and extended Joseph polynomials, to appear in *J. Phys.* A, math-ph/0508059.
[8] A. Knutson and P. Zinn-Justin, A scheme related to the Brauer loop model, math.AG/0503224.
[9] P. Di Francesco and P. Zinn-Justin, Inhomogeneous model of crossing loops and multidegrees of some algebraic varieties, to appear in *Commun. Math. Phys.* (2005), math-ph/0412031.
[10] G. Kuperberg, Symmetry classes of alternating-sign matrices under one roof, *Ann. of Math.* (2) 156 (2002), no. 3, 835–866, math.CO/0008184.
[11] A. Joseph, On the variety of a highest weight module, *J. Algebra* 88 (1) (1984), 238–278.
[12] R. Hotta, On Joseph's construction of Weyl group representations, Tohoku Math. J. Vol. 36 (1984), 49–74.
[13] I.B. Frenkel and N. Reshetikhin, Quantum affine Algebras and Holonomic Difference Equations, *Commun. Math. Phys.* 146 (1992), 1–60.
[14] M. Jimbo and T. Miwa, Algebraic analysis of Solvable Lattice Models, CBMS Regional Conference Series in Mathematics vol. 85, American Mathematical Society, Providence, 1995.
[15] P. Di Francesco, A. Knutson and P. Zinn-Justin, Extended Orbital Varieties and the Yang–Baxter equation, work in progress.
[16] P. Di Francesco, Boundary qKZ equation and generalized Razumov–Stroganov sum rules for open IRF models, math-ph/0509011.
[17] P. Di Francesco, Inhomogenous loop models with open boundaries, *J. Phys.* A 38 (2005), 6091–6120, math-ph/0504032.
[18] N. Chriss and V. Ginzburg, Representation Theory and Complex Geometry, Birkhauser 1997.
[19] V. Pasquier, Scattering matrices and Affine Hecke Algebras, q-alg/9508002.
[20] A. Izergin, Partition function of the six-vertex model in a finite volume, *Sov. Phys. Dokl.* **32** (1987) 878-879.
[21] V. Korepin, Calculation of norms of Bethe wave functions, *Comm. Math. Phys.* **86** (1982) 391-418.
[22] S. Okada, Enumeration of Symmetry Classes of Alternating Sign Matrices and Characters of Classical Groups, math.CO/0408234.
[23] R.M. Green, Generalized Temperley–Lieb algebras and decorated tangles, *Journal of Knot Theory and its Ramifications* 7 (1998), 155–171.
[24] D. Bressoud, Proofs and confirmations. The story of the alternating sign matrix conjecture, Cambridge University Press (1999).
[25] B. Rothbach, unpublished.
[26] A. Melnikova, Description of B-orbit closures of order 2 in upper-triangular matrices, math.RT/0312290, to appear in *Transformation Groups*.

Service de Physique Théorique de Saclay, CEA/DSM/SPhT, URA 2306 du CNRS, C.E.A.-Saclay, F-91191 Gif sur Yvette Cedex, France

Laboratoire de Physique Théorique et Modèles Statistiques, UMR 8626 du CNRS, Université Paris-Sud, Bâtiment 100, F-91405 Orsay Cedex, France

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Central Delannoy numbers, Legendre polynomials, and a balanced join operation preserving the Cohen-Macaulay property

Gábor Hetyei

Abstract. We introduce a new join operation on colored simplicial complexes that preserves the Cohen-Macaulay property. An example of this operation puts the connection between the central Delannoy numbers and Legendre polynomials in a wider context.

Résumé. Nous introduisons une nouvelle opération qui joint des complexes simpliciaux équilibrés d'une telle manière que la propriété de Cohen-Macaulay est preservée. Une exemple de cette opération remette la rélation entre les nombres Delannoy centraux et les polynomiaux de Legendre dans un contexte plus large.

Introduction

The Delannoy numbers, introduced by Henri Delannoy [**7**] more than a hundred years ago, became recently subject of renewed interest, mostly in connection with lattice path enumeration problems. It was also noted for more than half a century, that a somewhat mysterious connection exists between the central Delannoy numbers and Legendre polynomials. This relation was mostly dismissed as a "coincidence" since the Legendre polynomials do not seem to appear otherwise in connection with lattice path enumeration questions.

In our work we attempt to lift a corner of the shroud covering this mystery. First we observe that a variant of table A049600 in the On-Line Encyclopedia of Integer Sequences [**13**] embeds the central Delannoy numbers into another, asymmetric table, and the entries of this table may be expressed by a generalization of the Legendre polynomial substitution formula: the non-diagonal entries are connected to Jacobi polynomials. Then we show that the lattice path enumeration problem associated to these asymmetric Delannoy numbers is naturally identifiable with a 2-colored lattice path enumeration problem (Section 2). This variant helps represent each asymmetric Delannoy number as the number of facets in the *balanced join* of a simplex and the order complex of a fairly transparent poset which we call a Jacobi poset. The balanced join operation takes two balanced simplicial complexes colored with the same set of colors as its input and yields a balanced simplicial complex colored with the same set of colors as its output. It is introduced in Section 3, which also describes the Jacobi posets.

The balanced join operation we were lead to introduce turns out to be fairly interesting by its own merit. According to a famous result of Stanley [**14**], the h-vector of a balanced Cohen-Macaulay is the f-vector of another colored complex. (The converse, and the generalization to flag numbers was shown by Björner, Frankl, and Stanley [**2**].) Since the proof is algebraic, it is usually hard to construct the colored complex explicitly. Using the balanced join operation, we may construct balanced simplicial complexes as the balanced join of a balanced complex and a simplex such that the h-vector of the join is the f-vector of the original colored complex. This applies even if the balanced join does not have the Cohen-Macaulay property. Our main result is Theorem 4.3, stating that the balanced join of two balanced Cohen-Macaulay simplicial complexes is Cohen-Macaulay.

2000 *Mathematics Subject Classification.* Primary 13F55; Secondary 05A15, 16E65, 33C45.
Key words and phrases. balanced simplicial complex, Delannoy numbers, Cohen-Macaulay property.
On leave from the Rényi Mathematical Institute of the Hungarian Academy of Sciences.

In Section 5 we return to the Jacobi posets introduced in Section 3 and prove that their order complex is Cohen-Macaulay, thus our main result is applicable to the example that inspired it. The proof consists of providing an EL-labeling from which the the Cohen-Macaulay property follows by the results of Björner and Wachs [**3**] and [**4**]. By the results of Björner, Frankl, and Stanley [**2**] the flag h-vector of a Jacobi poset is the flag f-vector of a colored complex. We find this colored complex as the order complex of a *strict direct product* of two chains. We define the strict direct product of two posets by requiring a strict inequality in both coordinates.

Since removing the top and bottom elements from a Jacobi poset yields a "half-strict" direct product of two chains, there is another potentially interesting bivariate operation looming on the horizon. In Section 6 we introduce a right-strict direct product on posets that allows to assign to a pair (P, Q) of an arbitrary poset P and a graded poset Q a graded poset of the same rank as Q. There is reason to suspect that this product too, preserves the Cohen-Macaulay property, as we can show that the flag h-vector of a right-strict product is positive if the order complex of P has a positive h-vector, and Q has a positive flag h-vector.

In the concluding Section 7 we point out the impossibility of two seemingly plausible generalizations, and highlight the question in commutative algebra that arises when we try to generalize our main result, Theorem 4.3.

The journey taken will hopefully convince more mathematicians that Delannoy numbers are interesting, since they lead to some interesting results and questions in commutative algebra and algebraic combinatorics.

Acknowledgments

I wish to thank my father, Gábor Hetyei, for fruitful conversations, and my former thesis advisor, Richard Stanley, for valuable mathematical information, advice, and encouragement.

1. Preliminaries

1.1. Delannoy numbers. The *Delannoy array* $(d_{i,j} : i, j \in \mathbb{Z})$ was introduced by Henri Delannoy [**7**] in the nineteenth century. This array may be defined by the recursion formula

$$d_{i,j} = d_{i-1,j} + d_{i,j-1} + d_{i-1,j-1} \tag{1.1}$$

with the conditions $d_{0,0} = 1$ and $d_{i,j} = 0$ if $i < 0$ or $j < 0$. For $i, j \geq 0$ the number $d_{i,j}$ represents the number of lattice walks from $(0, 0)$ to (i, j) with steps $(1, 0)$, $(0, 1)$, and $(1, 1)$ The significance of these numbers is explained within a historic context in the paper "Why Delannoy numbers?" [**1**] by Banderier and Schwer. The diagonal elements $(d_{n,n} : n \geq 0)$ in this array are the *(central) Delannoy numbers* (A001850 of Sloane [**13**]). These numbers are known through the books of Comtet [**6**] and Stanley [**16**], but it is Sulanke's paper [**17**] that gives the most complete list of all known uses of Delannoy numbers (a total of 29 configurations). For more information and a detailed bibliography we refer the reader to the above mentioned sources.

1.2. Balanced simplicial complexes and the Cohen Macaulay property. A *simplicial complex* $\triangle$ on the vertex set V is a family of subsets of V, such that $\{v\} \in \triangle$ for all $v \in V$ and every subset of a $\sigma \in \triangle$ belongs to $\triangle$. An element $\sigma \in \triangle$ is a *face* and $|\sigma| - 1$ is its *dimension.* The dimension of $\triangle$ is the maximum of the dimensions of its faces. A maximal face is a *facet* and $\triangle$ is *pure* if all its facets have the same dimension. The number of i-dimensional faces is denoted by f_i. An equivalent encoding of the *f-vector* $(f_{-1}, \ldots, f_{n-1})$ of an $(n-1)$-dimensional simplicial complex is its h-vector *h-vector* $(h_0, \ldots, h_n)$ given by $h_i = \sum_{j=0}^{i} (-1)^{i-j} \binom{n-j}{i-j} f_{j-1}$. An $(n-1)$-dimensional simplicial complex $\triangle$ is *balanced* if its vertices may be colored using n colors such that every face has all its vertices colored differently. (See [**15**, 4.1 Definition].[1]) It is always assumed that a fixed coloring is part of the structure of a balanced complex. For such a complex we may refine the notions of f-vector and h-vector, as follows. Assume we use the set of colors $\{1, 2, \ldots, n\}$. For any $S \subseteq \{1, 2, \ldots, n\}$ let f_S be the number of faces whose vertices are colored exactly with the colors from S. The vector $(f_S : S \subseteq \{1, 2, \ldots, n\})$ is called the *flag f-vector* of the colored complex. The *flag h-vector* is then the vector $(h_S : S \subseteq \{1, 2, \ldots, n\})$ whose entries are given by

$$h_S = \sum_{T \subseteq S} (-1)^{|S \setminus T|} f_T.$$

[1]In the original definition of a balanced complex (occurring in [**14**]) it was also assumed that the complex is pure, but as it was observed by Stanley in [**15**, §III.4], "there is no real reason for this restriction".

A fundamental theorem on balanced simplicial complexes is Stanley's result [**14**, Corollary 4.5].

THEOREM 1.1 (Stanley). *The h-vector of a balanced Cohen-Macaulay simplicial complex is the f-vector of some other simplicial complex.*

The definition of the Cohen-Macaulay property is fairly involved, we refer the reader to Stanley [**15**]. To prove our main result, we use Reisner's criterion [**15**, Chapter II, Corollary 4.1] which characterizes Cohen-Macaulay simplicial complexes in terms of the homology groups of each *link*. The link $\mathrm{lk}_{\triangle}(\tau)$ of a face $\tau \in \triangle$ is defined by

$$\mathrm{lk}_{\triangle}(\tau) := \{\sigma \in \triangle : \sigma \cap \tau = \emptyset, \sigma \cup \tau \in \triangle\}.$$

The homology used is *simplicial homology* [**15**, Chapter 0, Section 4]. An oriented j-simplex in $\triangle$ is a j-face $\sigma = \{v_0, \ldots, v_j\} \in \triangle$, enriched with an equivalence class of orderings, two orderings being equivalent if they differ by an even permutation of vertices. We write $[v_0, v_1, \ldots, v_j]$ for the oriented simplex associated to the equivalence class of the linear order $v_0 < \ldots < v_j$. The k-module $C_j(\triangle)$ (for $j = -1, \ldots, \dim(\triangle)$, where k is a field) is then the free k-module generated by all oriented j-simplices modulo the relations $[\sigma_1] + [\sigma_2] = 0$ whenever $[\sigma_1]$ and $[\sigma_2]$ are different oriented simplices corresponding to the j-simplex. These modules, together with the boundary maps $\partial_j : C_j(\triangle) \to C_{j-1}(\triangle)$, given by

$$\partial_j [v_0, \ldots, v_j] = \sum_{i=0}^{j} (-1)^i [v_0, \ldots, \widehat{v}_i, \ldots, v_j]$$

form the *oriented chain complex* of $\triangle$. (As usual, $\widehat{v}_i$ indicates omitting v_i.) Reisner's criterion is then the following.

THEOREM 1.2 (Reisner). *The simplicial complex $\triangle$ is Cohen-Macaulay over k if and only if for all $\sigma \in \triangle$ and $i < \dim \mathrm{lk}_{\triangle}(\sigma)$ we have $\widetilde{H}_i(\mathrm{lk}_{\triangle}(\sigma), k) = 0$. Here $\widetilde{H}_i$ denotes the i-th reduced homology group of the appropriate oriented chain complex.*

Rephrasing results of his work with Björner and Frankl [**2**], Stanley refined Theorem 1.1 to flag numbers as follows [**15**, Chapter III, Theorem 4.6].

THEOREM 1.3 (Björner-Frankl-Stanley). *A vector $(\beta_S : S \subseteq \{1, 2, \ldots, n\})$ is the flag h-vector of some $(n-1)$-dimensional balanced Cohen-Macaulay simplicial complex if and only if it is the flag f-vector of some other colored simplicial complex.*

REMARK 1.4. *Although Stanley uses the term "balanced" twice in his statement* [**15**, Chapter III, Theorem 4.6]*, it is clear from his proof that the second complex only needs to be colored with the same color set as the first. The number of colors thus used may exceed the size of the largest face in the second complex. For example, an $(n-1)$-simplex is balanced and Cohen-Macaulay, all entries in its flag f-vector are 1's. The flag h-entries are all zero except for $h_{\emptyset} = 1$. Thus the second complex must have only one face, the empty set. This complex may be trivially colored using n colors (without actually using any of them).*

An important example of a balanced simplicial complex is the *order complex* $\triangle(P \setminus \{\widehat{0}, \widehat{1}\})$ of a *graded partially ordered set* P. The order complex $\triangle(Q)$ of any poset Q is the simplicial complex on the vertex set Q whose faces are the chains of Q. A poset is graded if it has a unique minimum $\widehat{0}$, a unique maximum $\widehat{1}$, and a rank function ρ. Since all saturated chains of P have the same cardinality, $\triangle(P \setminus \{\widehat{0}, \widehat{1}\})$ is pure, and coloring every element with its rank makes $\triangle(P \setminus \{\widehat{0}, \widehat{1}\})$ balanced.

2. Central Delannoy numbers and Legendre polynomials

The following connection between the central Delannoy numbers and Legendre polynomials has been known for at least half a century [**8**], [**10**], [**11**]:

$$d_{n,n} = P_n(3), \tag{2.1}$$

where $P_n(x)$ is the n-th Legendre polynomial. To date there seems to be a consensus that this link is not very relevant. Banderier and Schwer [**1**] note that there is no "natural" correspondence between Legendre polynomials and the original lattice path enumeration problem associated to the Delannoy array, while Sulanke [**17**] states that "the definition of Legendre polynomials does not appear to foster any combinatorial interpretation leading to enumeration".

Without disagreeing with these statements concerning the *original* lattice path-enumeration problem, in this section we point out the existence of a *modified* lattice path enumeration problem whose solution yields a modified Delannoy array $\widetilde{d}_{m,n}$ satisfying $\widetilde{d}_{n,n} = d_{n,n}$ and

$$\widetilde{d}_{m,n} = P_n^{(0,m-n)}(3) \quad \text{for } m \geq n. \tag{2.2}$$

Here $P_n^{(\alpha,\beta)}(x)$ is the n-th Jacobi polynomial of type (α,β) defined by

$$P_n^{(\alpha,\beta)}(x) = (-2)^{-n}(n!)^{-1}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}\left((1-x)^{n+\alpha}(1+x)^{n+\beta}\right).$$

Since the polynomials $P_n^{(0,0)}(x)$ are the Legendre polynomials [**5**, Chapter V, (2.2)], substituting $m = n = 0$ into (2.2) yields (2.1). In section 3 we present a face-enumeration problem associated to Jacobi polynomials that is related our modified lattice path enumeration problem.

The lattice path enumeration problem in question is essentially identical to the one of A049600 in the On-Line Encyclopedia of Integer Sequences [**13**].

Definition 2.1. For any $(m,n) \in \mathbb{N}\times\mathbb{N}$ let us denote by $\widetilde{d}_{m,n}$ the number of lattice paths from $(0,0)$ to $(m,n+1)$ having steps $(x,y) \in \mathbb{N}\times\mathbb{P}$. (Here $\mathbb{P}$ denotes the set of positive integers.) We call the numbers $\widetilde{d}_{m,n}$ $(m,n \geq 0)$ the *asymmetric Delannoy numbers.*

$\widetilde{d}_{m,n} :=$

$m \backslash n$	0	1	2	3	4
0	1	2	4	8	16
1	1	3	8	20	48
2	1	4	13	38	104
3	1	5	19	63	192
4	1	6	26	96	321

Table 1. The asymmetric Delannoy numbers $\widetilde{d}_{m,n}$ for $0 \leq m,n \leq 4$.

It is immediate from our definition that $\widetilde{d}_{m,n} = T(n+1,m)$ for the array T given in A049600. As a consequence we get $\widetilde{d}_{n,n} = T(n+1,n)$ which is the central Delannoy number $d_{n,n}$, as noted in A049600. Compared to A049600, we shifted the rows up by 1 to move the central Delannoy numbers to the main diagonal, and then we reflected the resulting table to its main diagonal, since this will allow picturing the partially ordered sets in section 3 the "usual" way, i.e., with the larger elements being above the smaller ones. As an immediate consequence of the definition we obtain the following:

Lemma 2.2. *The asymmetric Delannoy numbers satisfy*

$$\widetilde{d}_{m,n} = \sum_{j=0}^{n}\binom{n}{j}\binom{m+j}{j}.$$

It is worth noting that substituting $n = m$ into Lemma 2.2 yields a well-known representation of the central Delannoy number $d_{n,n}$. (See, Sulanke [**17**, Example 1].) We may also easily verify (2.2), as follows. A Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$ with nonnegative integer parameters α, β may be also given in the form

$$P_n^{(\alpha,\beta)}(x) = \sum_j \binom{n+\alpha+\beta+j}{j}\binom{n+\alpha}{j+\alpha}\left(\frac{x-1}{2}\right)^j \tag{2.3}$$

See, e.g., Wilf [Chapter 4, Exercise 15 (b)][**18**]. Substituting $\alpha = 0$ we obtain

$$P_n^{(0,\beta)}(x) = \sum_j \binom{n+\beta+j}{j}\binom{n}{j}\left(\frac{x-1}{2}\right)^j, \tag{2.4}$$

from which (2.2) follows by setting $\beta = m-n$ and $x = 3$.

Corollary 2.3. *The asymmetric Delannoy numbers satisfy (2.2).*

Remark 2.4. *It is usually required that $\alpha, \beta > -1$ in the definition of the Jacobi polynomials "for integrability purposes", cf. Chihara* [**5**, Chapter V., section 2.]. *That said, using (2.4) we may extend the definition of $P_n^{(0,\beta)}(x)$ to any integer $\beta \geq -n$. Using this extended definition, we may state (2.2) for any $m, n \geq 0$.*

A combinatorial interpretation of (2.2) will be facilitated by the following reinterpretation of our lattice path enumeration problem.

Proposition 2.1. *For any $(m,n) \in \mathbb{N} \times \mathbb{N}$ the number $\widetilde{d}_{m,n}$ also enumerates all 2-colored lattice paths from $(0,0)$ to (m,n) satisfying the following:*

(i) *Each step is either a blue $(0,1)$ or a red $(x,y) \in \mathbb{N} \times \mathbb{N} \setminus \{(0,0)\}$.*
(ii) *At least one of any two consecutive steps is a blue $(0,1)$.*

Proof. It is easy to verify directly that the number of all 2-colored lattice paths from $(0,0)$ to (m,n) with the above properties is $\sum_{j=0}^{n} \binom{n}{j}\binom{m+j}{j}$, and so we get $\widetilde{d}_{m,n}$ by Lemma 2.2. (j is the number of blue steps). But only a little more effort is necessary to find a fairly plausible bijection between the corresponding sets of lattice paths. Consider first a lattice path from $(0,0)$ to $(m,n+1)$ satisfying Definition 2.1. Replace

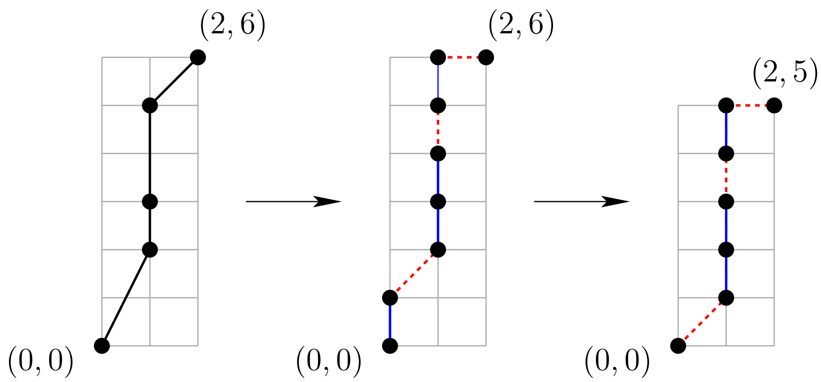


Figure 1. Transforming a lattice path into a 2-colored lattice path.

each step $(x,y) \in \mathbb{N} \times \mathbb{P}$ with one or two colored steps as follows. If $(x,y) \neq (0,1)$ then replace it with a blue $(0,1)$ followed by a red $(x,y-1)$. If $(x,y) = (0,1)$ then replace it with a blue $(0,1)$. The resulting 2-colored lattice path from $(0,0)$ to $(m,n+1)$ satisfies conditions (i) and (ii), moreover it starts always with a blue $(0,1)$. Remove this first blue step and shift the colored lattice path down by 1 unit. We obtain a 2-colored lattice path from $(0,0)$ to (m,n) satisfying the conditions of our proposition. Fig. 1 shows the two stages of such a transformation. (In the picture, $m = 2$ and $n = 5$. Red edges are marked with dashed lines.)

Finding the inverse of this transformation is easy. Given a valid 2-colored lattice path from $(0,0)$ to (m,n), let us first shift the path up by 1 and prepend a blue $(0,1)$ step from $(0,0)$ to $(0,1)$. Thus we obtain a 2-colored lattice path from $(0,0)$ to (m,n) satisfying (i), while condition (ii) may be strengthened to stating that every red step is preceded by a blue $(0,1)$ step. Replace each red step (x,y) and the blue step preceding it with a single colorless step $(x,y+1)$. After this, remove the blue color of the remaining $(0,1)$-steps.

We leave it to the reader to verify the fact that the two operations described above are inverses of each other. □

Using the equivalent definition of Proposition 2.1 it is easy to verify the following additional property of the asymmetric Delannoy numbers.

Lemma 2.5. *The asymmetric Delannoy numbers satisfy the recursion formula*

$$\widetilde{d}_{m,n} = \widetilde{d}_{0,0} + \sum_{i=0}^{m} \sum_{j=0}^{n-1} \widetilde{d}_{i,j} \quad \textit{for all } m,n \geq 0.$$

Here the second sum is empty if $n = 0$.

As a consequence any entry in Table 1 may be obtained by adding 1 to the sum of the entries in the preceding columns, up to the row of the selected entry.

Corollary 2.6. *The asymmetric Delannoy numbers satisfy the recursion formula*

$$\widetilde{d}_{m,n} = 2 \cdot \widetilde{d}_{m,n-1} + \widetilde{d}_{m-1,n} - \widetilde{d}_{m-1,n-1}.$$

3. Jacobi posets and balanced joins

The blue steps of a valid 2-colored path introduced in Proposition 2.1 form increasing chains in a partially ordered set. In this section we investigate this partial order.

Definition 3.1. Given any integer β and $n \geq \max(0, -\beta)$, we call the *Jacobi poset P_n^β of type β and rank $n+1$* the following graded poset.

(*i*) For each $q \in \{1, \ldots, n\}$, P_n^β has $n + \beta + 1$ elements of rank q, they are labeled $(0, q)$, $(1, q)$, $\ldots$, $(n + \beta, q)$.

(*ii*) Given (p, q) and (p', q') in $P_n^\beta \setminus \{\widehat{0}, \widehat{1}\}$ we set $(p, q) < (p', q')$ iff. $p \leq p'$ and $q < q'$.

(We also require $\widehat{0}$ to be the minimum element and $\widehat{1}$ to be the maximum element.)

We may think of the elements of $P_n^\beta \setminus \{\widehat{0}, \widehat{1}\}$ as the endpoints of all possible blue $(0, 1)$ steps when we enumerate all valid 2-colored lattice paths from $(0, 0)$ to $(n + \beta, n)$. We have $(p, q) < (p', q')$ if and only of if there is a valid 2-colored lattice path containing both $(p, q-1) - (p, q)$ and $(p', q'-1) - (p', q')$ as blue steps, such that the first blue step precedes the second in the path. Fig. 2 represents the Jacobi poset P_5^{-3}, which may be associated to enumerating the valid 2-colored lattice paths from $(0, 0)$ to $(2, 5)$. In the picture of the poset we marked the elements corresponding to the blue edges with empty circles. Given any valid

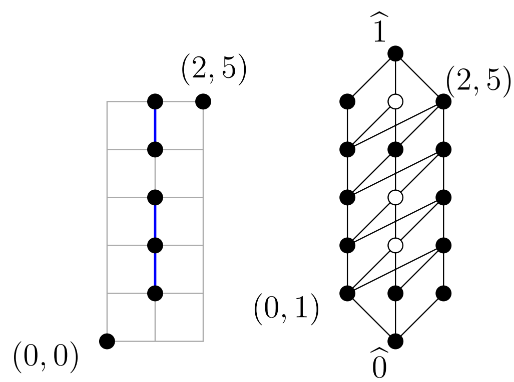


Figure 2. The Jacobi poset P_5^{-3} and the partial chain encoding the lattice path in Fig.1

2-colored path from $(0, 0)$ to $(n + \beta, n)$, the set of its blue edges correspond to a partial chain in $P_n^\beta \setminus \{\widehat{0}, \widehat{1}\}$ and, conversely, any partial chain of $P_n^\beta \setminus \{\widehat{0}, \widehat{1}\}$ encodes a set of blue edges that may be uniquely completed to a valid 2-colored path by adding the appropriate red edges.

Obviously, the face numbers of the order complex of $P_n^\beta \setminus \{\widehat{0}, \widehat{1}\}$ satisfy

$$f_{j-1}\left(\triangle\left(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\}\right)\right) = \binom{n}{j}\binom{n+\beta+j}{j}. \tag{3.1}$$

As a consequence of this equality and (2.4) we obtain that

$$\sum_{j=0}^{n} f_{j-1}\left(\triangle\left(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\}\right)\right) \cdot \left(\frac{x-1}{2}\right)^j = P^{(0,\beta)}(x) \quad \text{for } \beta \geq 0. \tag{3.2}$$

Note that, for negative values of β, (3.2) still holds in the extended sense of Remark 2.4. As another consequence of (3.2), the asymmetric Delannoy number $\widetilde{d}_{m,n}$ equals the number of all partial chains (including the empty chain) in the Jacobi poset $P_n^{m-n} \setminus \{\widehat{0}, \widehat{1}\}$. This fact is also "visually obvious" in terms of the enumeration problem presented in Proposition 2.1, since any valid 2-colored lattice path may be uniquely reconstructed from its blue steps. This visualization inspires the following definition.

Definition 3.2. Let $\triangle_1$ and $\triangle_2$ be pure balanced simplicial complexes of the same dimension. Let us fix a pair of colorings $\lambda = (\lambda_1, \lambda_2)$ such that λ_i colors the vertices of $\triangle_i$ $(i = 1, 2)$ in a balanced way, and the set of colors is the same in both colorings. We call the simplicial complex

$$\triangle_1 *_\lambda \triangle_2 := \{\sigma \cup \tau : \sigma \in \triangle_1, \tau \in \triangle_2, \lambda_1(\sigma) \cap \lambda_2(\tau) = \emptyset\}$$

the *balanced join* of $\triangle_1$ and $\triangle_2$ with respect to λ.

Example 3.3. *Let $\triangle_1$ and $\triangle_2$ be both $(n-1)$-dimensional simplices. These have essentially one balanced coloring and, independently of the choice of $\lambda = (\lambda_1, \lambda_2)$, the free join $\triangle_1 *_\lambda \triangle_2$ is isomorphic to the boundary complex of an n-dimensional cross-polytope.*

Using the notion of the balanced join we may express the relation between the asymmetric Delannoy numbers and Jacobi posets as follows.

Theorem 3.4. *Given $m, n \geq 0$, let λ_1 be the coloring of the order complex of $P_n^{m-n} \setminus \{\widehat{0}, \widehat{1}\}$ induced by the rank function, and let λ_2 be any balanced coloring of an $(n-1)$-dimensional simplex $\triangle^{n-1}$ with the color set $\{1, 2, \ldots, n\}$, Then, for $\lambda = (\lambda_1, \lambda_2)$, the asymmetric Delannoy number $\widetilde{d}_{m,n}$ is the number of facets in the balanced join $\triangle \left(P_n^{m-n} \setminus \{\widehat{0}, \widehat{1}\}\right) *_\lambda \triangle^{n-1}$.*

In fact, Theorem 3.4 may be generalized to any graded poset P of rank $n+1$ as follows.

Theorem 3.5. *Given any graded poset P of rank $n+1$, let λ_1 be the coloring of the order complex of $P \setminus \{\widehat{0}, \widehat{1}\}$ induced by the rank function, and let λ_2 be any balanced coloring of an $(n-1)$-dimensional simplex $\triangle^{n-1}$ with the color set $\{1, 2, \ldots, n\}$, Then, for $\lambda = (\lambda_1, \lambda_2)$, the number of facets in the balanced join $\triangle \left(P \setminus \{\widehat{0}, \widehat{1}\}\right) *_\lambda \triangle^{n-1}$ equals the total number of partial chains in $P \setminus \{\widehat{0}, \widehat{1}\}$.*

Theorem 3.5 is an immediate consequence of the fact that $\triangle^{n-1}$ must have precisely one vertex of each color, thus any partial chain from $P \setminus \{\widehat{0}, \widehat{1}\}$ may be uniquely complemented to a facet of $\triangle \left(P \setminus \{\widehat{0}, \widehat{1}\}\right) *_\lambda \triangle^{n-1}$ by inserting exactly those vertices of the simplex which are colored by the ranks missed by the partial chain. Theorem 3.4 is then a consequence of Theorem 3.5, Lemma 2.2 and equation (3.1).

4. Properties of the balanced join operation

In this section we take a closer look at the balanced join operation introduced at the end of the previous section. Let us point out first that the operation does depend on the colorings chosen.

Example 4.1. *Consider the "star graph" $\triangle$ shown in Fig. 3. This is a 1-dimensional simplicial complex*

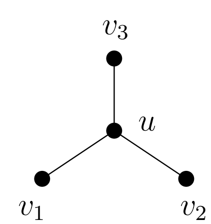


Figure 3. A "star graph" that may be colored in essentially one way

which has essentially one balanced coloring with 2 colors: v_1, v_2, and v_3 must have the same color, and u must have the other color. Yet, when we fix the set $\{1, 2\}$ to be the set of our colors, we have 2 options to chose the color of u (the rest of the coloring is then uniquely determined). Thus in a balanced join $\triangle *_\lambda \triangle$ we may choose $\lambda = (\lambda_1, \lambda_2)$ in such a way that $\lambda_1 \neq \lambda_2$, or we may use the same coloring twice. If $\lambda_1 = \lambda_2$ then the complex $\triangle *_\lambda \triangle$ has $1 \cdot 3 + 3 \cdot 1 = 6$ edges, while in the case $\lambda_1 \neq \lambda_2$ the complex $\triangle *_\lambda \triangle$ has $1 \cdot 1 + 3 \cdot 3 = 10$ edges.

However, the observation made in Example 3.3 may be generalized as follows. Given any balanced simplicial complex $\triangle$ of dimension $(n-1)$, and a balanced coloring λ_1 of it, the balanced join $\triangle *_\lambda \triangle^{n-1}$ with an $(n-1)$-simplex $\triangle^{n-1}$ obviously does not depend on the choice of the its coloring λ_2. Moreover, we have the following fact:

Theorem 4.2. *The flag h-vector of the balanced join $\triangle *_\lambda \triangle^{n-1}$ of a balanced $(n-1)$-dimensional simplicial complex with an $(n-1)$-simplex $\triangle^{n-1}$ satisfies*

$$h_S\left(\triangle *_\lambda \triangle^{n-1}\right) = f_S\left(\triangle\right)$$

for any subset S of the colors used.

Proof. We may assume that the set of colors is $\{1, 2, \ldots, n\}$. Any face of $\triangle *_\lambda \triangle^{n-1}$ of color S is a disjoint union $\sigma \cup \tau$ with $\sigma \in \triangle$, $\tau \in \triangle^{n-1}$. The set $T := \lambda_1(\sigma)$ must be a subset of S, and $\lambda_2(\tau)$ must be equal to $S \setminus T$. There is precisely one $\tau \in \triangle^{n-1}$ with this property, hence we obtain

$$f_S\left(\triangle *_\lambda \triangle^{n-1}\right) = \sum_{T \subseteq S} f_T\left(\triangle\right).$$

The statement now follows by the sieve formula. □

The proof of Theorem 4.2 is almost trivial, but the statement provides an interesting "constructive reason" for a situation that occurs in Theorem 1.3. The proof of Theorem 1.3 is algebraic, and it is usually hard to find a balanced simplicial complex combinatorially whose flag f-vector is the flag h-vector of the given Cohen-Macaulay balanced simplicial complex. An example of a combinatorial explanation in the special case of order complexes of certain distributive lattices is given by Skandera [**12**, Theorem 3.2]. The above construction is fairly "rigid", but it yields also examples of balanced simplicial complexes without the Cohen-Macaulay property. On the other hand, the balanced join preserves the Cohen-Macaulay property if we apply it to Cohen-Macaulay complexes.

Theorem 4.3. *Assume that $\triangle_1$ and $\triangle_2$ are balanced Cohen-Macaulay complexes of dimension $(n-1)$, and that their colorings λ_1 and λ_2 use the same set of colors $\{1, 2, \ldots, n\}$. Then the balanced join $\triangle_1 *_\lambda \triangle_2$ is also a Cohen-Macaulay simplicial complex.*

Proof. We prove the Cohen-Macaulay property by induction on the size of $\triangle_1 \cup \triangle_2$, using Reisner's Theorem (Theorem 1.2 in the Preliminaries). For that purpose, consider the link of any face $\tau_1 \cup \tau_2$ where $\tau_1 \in \triangle_1$, $\tau_2 \in \triangle_2$ and $\lambda_1(\tau_1) \cap \lambda_2(\tau_2) = \emptyset$. The faces of $\mathrm{lk}_{\triangle_1 *_\lambda \triangle_2}(\tau_1 \cup \tau_2)$ are precisely the faces of the form $\sigma_1 \cup \sigma_2$ where $\sigma_i \in \mathrm{lk}_{\triangle_i}(\tau_i)$ for $i = 1, 2$, and the sets of colors $\lambda_1(\sigma_1)$, $\lambda_2(\sigma_2)$, $\lambda_1(\tau_1)$, and $\lambda_2(\tau_2)$ are pairwise disjoint. Using this description it is easy to deduce

$$\mathrm{lk}_{\triangle_1 *_\lambda \triangle_2}(\tau_1 \cup \tau_2) = \mathrm{lk}_{\triangle_1}(\tau_1)_{\{1,\ldots,n\}\setminus\lambda_2(\tau_2)} *_\lambda \mathrm{lk}_{\triangle_2}(\tau_2)_{\{1,\ldots,n\}\setminus\lambda_1(\tau_1)}.$$

Here the simplicial complexes $\mathrm{lk}_{\triangle_1}(\tau_1)_{\{1,\ldots,n\}\setminus\lambda_2(\tau_2)}$ and $\mathrm{lk}_{\triangle_2}(\tau_2)_{\{1,\ldots,n\}\setminus\lambda_1(\tau_1)}$ are both balanced and the appropriate restrictions of λ_1 resp. λ_2 color both with the same color set $\{1, \ldots, n\} \setminus (\lambda_1(\tau_1) \cup \lambda_2(\tau_2))$. By Reisner's theorem, the link of every face in a Cohen-Macaulay complex is Cohen-Macaulay. According to Stanley's theorem [**15**, Chapter III, Theorem 4.5], every rank-selected subcomplex of a balanced Cohen-Macaulay complex is Cohen-Macaulay. Thus, whenever at least one of τ_1 and τ_2 is not the empty set, we may apply our induction hypothesis to the balanced join of $\mathrm{lk}_{\triangle_1}(\tau_1)_{\{1,\ldots,n\}\setminus\lambda_2(\tau_2)}$ and $\mathrm{lk}_{\triangle_2}(\tau_2)_{\{1,\ldots,n\}\setminus\lambda_1(\tau_1)}$.

We are left to prove Reisner's criterion for the reduced homology groups of the oriented chain complex associated to $\triangle_1 *_\lambda \triangle_2$. Assume by way of contradiction that there exist an $i < n-1$ and a linear combination $\underline{c} = \sum_{j,k} \alpha_{j,k} \cdot [\sigma_j \cup \tau_k] \in C_i(\triangle_1 *_\lambda \triangle_2)$ that belongs to $\mathrm{Ker}(\partial_i)$ but not to $\mathrm{Im}(\partial_{i+1})$. Here we may assume that each σ_j belongs to $\triangle_1$, each τ_k belongs to $\triangle_2$, and that no two of these sets are the same. Furthermore, we agree that in the oriented simplices we always list the elements of the face from $\triangle_1$ before the elements of the face from $\triangle_2$, hence "putting square brackets around a union of such faces" will not cause any confusion. Finally, for each fixed j at least one scalar $\alpha_{j,k}$ is not zero (otherwise σ_j is superfluous) and for each fixed k at least one $\alpha_{j,k}$ is not zero (otherwise τ_k is superfluous). Assume that our counterexample is smallest in the sense that the maximum of $|\tau_k|$ is as small as possible.

W.l.o.g. we may assume that τ_1 is of maximum size and thus it not contained in any other τ_k. Applying ∂_i to $\underline{c}$ yields a linear combination of oriented simplices, whose underlying simplices are of the form $\sigma_j \cup \tau_1 \setminus \{u\}$ where $u \in \sigma_j \cup \tau_1$. Consider among these oriented simplices the ones whose underlying simplex contains τ_1. Because of the maximality of τ_1, none of these may arise by removing some element from a $\sigma_j \cup \tau_k$ with $k \neq 1$. Hence the projection of $\partial_i(\underline{c})$ onto the vector space generated by the oriented simplices containing τ_1 may be obtained from $\partial_{i-|\tau_1|}\left(\sum_j \alpha_{j,1} \cdot [\sigma_j]\right) \in C_{i-|\tau_1|-1}(\triangle_1)$ by sending each $[\sigma'] \in C_{i-|\tau_1|-1}(\triangle_1)$ into $[\sigma' \cup \tau_1] \in C_{i-1}(\triangle_1 *_\lambda \triangle_2)$. Since $\partial_i(\underline{c}) = 0$, we obtain that $\sum_j \alpha_{j,1} \cdot [\sigma_j] \in \mathrm{Ker}(\partial_{i-|\tau_1|})$ in the oriented chain complex associated to $\triangle_1$. Here $i - |\tau_1| \le i < n-1$, hence applying Reisner's criterion to $\triangle_1$ yields $\sum_j \alpha_{j,1} \cdot [\sigma_j] \in \mathrm{Im}(\partial_{i+1-|\tau_1|})$. Assume

$$\sum_j \alpha_{j,1} \cdot [\sigma_j] = \partial_{i+1-|\tau_1|} \left(\sum_t \beta_t \cdot [\sigma'_t] \right)$$

holds in $C(\triangle_1)$, and consider

$$\underline{c'} = \sum_t \beta_t \cdot [\sigma'_t \cup \tau_1] \in C_{i+1}\left(\triangle_1 *_\lambda \triangle_2\right).$$

Subtracting $\partial_{i+1}(\underline{c'})$ from $\underline{c}$ removes all terms of the form $\alpha_{j,1} \cdot [\sigma_j \cup \tau_1]$ and introduces only new terms of the form $\alpha \cdot [\sigma' \cup \tau']$, where $\sigma' \in \triangle_1$, $\tau' \in \triangle_2$, and τ' is a proper subset of τ_1. Hence we reduced the number of τ_k's of maximum size in our counterexample. Repeating the same argument finitely many times we arrive at a counterexample in which the maximum size of all τ_k's is smaller than in the original one. We obtain a contradiction unless there is only one τ_k in $\underline{c}$, namely $\tau_1 = \emptyset$. However, for elements of the form $\sum_{j,1} \alpha_{j,1}[\sigma_j \cup \emptyset]$ where $\sigma_j \in \triangle_1$, the effect of the boundary map is described with the same formulas in the oriented chain complex associated to $\triangle_1$ and in the oriented chain complex associated to $\triangle_1 *_\lambda \triangle_2$. Applying Reisner's theorem to $\triangle_1$ yields a contradiction. $\square$

5. Jacobi posets and balanced Cohen-Macaulay complexes

Theorem 4.3 is applicable to $\triangle\left(P_n^{m-n} \setminus \{\widehat{0}, \widehat{1}\}\right) *_\lambda \triangle^{n-1}$ because of the following statement.

PROPOSITION 5.1. *The order complex $\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})$ associated to the Jacobi poset P_n^β is Cohen-Macaulay.*

PROOF. Let us label each cover relation $(p, q) \prec (p', q+1)$ with $n + \beta - p'$ and each cover relation $\widehat{0} \prec (p, 1)$ with $n + \beta - p$. Finally, let us label each cover relation $(p, n) \prec \widehat{1}$ with 0.

The resulting labeling is an *EL-labeling*, as defined by Björner and Wachs [**4**, Definition 2.1] (these labelings were first introduced in [**3**]). (We omit the proof that we get an EL-labeling, for brevity's sake.)

If a graded poset has an EL-labeling then its order complex is *shellable* by the result of Björner and Wachs [**4**, Proposition 2.3]. Shellable simplicial complexes are Cohen-Macaulay, see Stanley [**15**, Chapter III, Theorem 2.5]. $\square$

Since $\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})$ is Cohen-Macaulay and balanced, we may apply Theorem 1.3 to observe that its flag h-vector is the flag f-vector of some balanced simplicial complex. For a reader familiar with EL-labelings it is not difficult to construct such a simplicial complex, by inspecting the "descent sets" arising in the proof of Proposition 5.1. To ease the burden of the reader who is not familiar with EL-labelings, we provide an explicit description of such a balanced simplicial complex, and we verify the equality of the appropriate invariants by explicitly computing them. The simplicial complex to be constructed arises as the order complex of a partially ordered set.

DEFINITION 5.1. Given two partially ordered sets P and Q, we define their *strict direct product* $P \bowtie Q$ as the set $P \times Q$ ordered by the relation $(p, q) < (p', q')$ if $p < p'$ and $q < q'$.

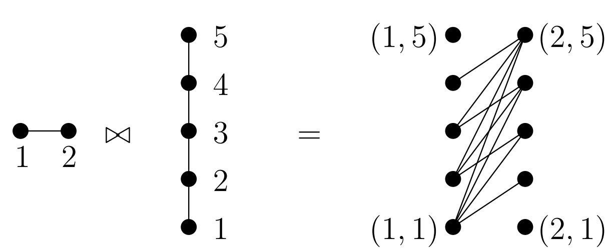


FIGURE 4. The strict direct product $C_1 \bowtie C_4$

Fig. 4 represents the strict direct product of a chain C_1 of length 1 with a chain C_4 of length 4. We obtain a partially ordered set that is not graded. However, the following statement is obviously true in general.

Lemma 5.2. *Given any pair of posets (P, Q), every admissible coloring of $\triangle(P)$ may be extended to an admissible coloring to $\triangle(P \bowtie Q)$ by coloring each $(p, q) \in P \times Q$ with the color of its first coordinate. The analogous statement is true for the second coordinates.*

Here we call a coloring admissible if the vertices of any face are colored with all different colors. For example, the order complex of the poset shown in Fig. 4 may be colored with 2 colors, by extending the coloring of the chain C_1, or with 5 colors, by extending the coloring of the chain C_5. Using the notion of the direct product, the flag h-vector of the order complex associated to a Jacobi poset may be described as follows.

Proposition 5.2. *The flag h-vector of $\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})$, colored by the rank function, equals the flag f-vector of $\triangle(C_{n+\beta-1} \bowtie C_{n-1})$, with respect to the coloring induced by the rank function of the second coordinate.*

Proof. For any $S \subseteq \{1, \ldots, n\}$, choosing a facet of $\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})_S$ is equivalent to choosing an $|S|$-element multiset on $\{0, 1, \ldots, n+\beta\}$. Hence we have

$$f_S\left(\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})\right) = \binom{n+\beta+|S|}{|S|}. \tag{5.1}$$

Using the identity

$$\binom{n+\beta+s}{s} = \sum_{t=0}^{s} \binom{s}{t} \cdot \binom{n+\beta}{t}$$

it is easy to deduce that the flag h-vector of $\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})$ must satisfy

$$h_S\left(\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})\right) = \binom{n+\beta}{|S|}. \tag{5.2}$$

Introducing ρ for the rank function $\rho : C_{n-1} \to \{1, \ldots n\}$ (note that the least element has rank one!), consider $\triangle(C_{n+\beta-1} \bowtie C_{n-1})$ with the coloring $\lambda(p, q) = \rho(q)$. For any $S \subseteq \{1, 2, \ldots, n\}$, choosing a saturated chain in $\triangle(C_{n+\beta-1} \bowtie C_{n-1})_S$ involves fixing the second coordinates, and choosing an $|S|$-element subset of a set with $n + \beta$ elements. Hence we have

$$f_S\left(C_{n+\beta-1} \bowtie C_{n-1}\right) = \binom{n+\beta}{|S|} \tag{5.3}$$

as stated. □

It should be noted that the strict direct product associated by Proposition 5.2 to P_5^{-3} (shown in Fig. 2) is $C_1 \bowtie C_4$ (shown in Fig. 4). To summarize our findings: we obtained that the asymmetric Delannoy number $\widetilde{d}_{m,n}$ counts the facets of the balanced Cohen-Macaulay complex $\triangle\left(P_n^{m-n} \setminus \{\widehat{0}, \widehat{1}\}\right) *_\lambda \triangle^{n-1}$. The flag h-vector of this complex is the flag f-vector of the order complex $\triangle(P_n^\beta \setminus \{\widehat{0}, \widehat{1}\})$. This complex is still balanced and Cohen-Macaulay, and its flag h-vector equals the flag f-vector of the colored complex described in Proposition 5.2. No further similar reduction is possible, since the order complex of the strict direct product of two chains is usually not Cohen-Macaulay. For example, the order complex of $C_1 \bowtie C_4$, shown in Fig. 4, is not even connected. The number of colors used also exceeds the size of the largest face. Thus, it appears, this is how far we may get using Theorem 1.3 in reducing the question of enumerating flags in the simplicial complex associated to the asymmetric Delannoy numbers.

6. The right-strict direct product of posets

The connection between the Jacobi posets and the strict direct product of two chains exposed in Proposition 5.2 suggests considering the following definition.

Definition 6.1. Given two partially ordered sets (P, Q) we define their *right-strict direct product* $P \rtimes Q$ to be the set $P \times Q$ partially ordered by the relation $(p, q) < (p', q')$ if $p \leq p'$ and $q < q'$.

The definition of the right-strict direct product is "halfway between" the usual definition of the direct product of posets and the strict direct product. Our interest is motivated by the following observation.

PROPOSITION 6.1. *The partially ordered set $P_n^\beta \setminus \{\widehat{0},\widehat{1}\}$ is isomorphic to $C_{n+\beta} \rtimes C_{n-1}$.*

The statement is an immediate consequence of the definitions. The fact that we obtain a graded poset (with the $\widehat{0}$ and the $\widehat{1}$ removed) may be generalized as follows.

PROPOSITION 6.2. *Assume P is an arbitrary poset and Q is a graded poset of rank $n+1$. Then $P \rtimes (Q \setminus \{\widehat{0},\widehat{1}\})$ may be turned into a graded poset of rank $n+1$ by adding a unique minimum element $\widehat{0}$ and a unique maximum element $\widehat{1}$. The rank function may be taken to be the rank function of Q applied to the second coordinate.*

The question naturally arises: how far can Proposition 5.1 be generalized, under what circumstances can we guarantee that a right-strict direct product of posets has a Cohen-Macaulay order complex?

CONJECTURE 6.2. *If P is a poset with a Cohen-Macaulay order complex and Q is a graded Cohen-Macaulay poset then $P \rtimes (Q \setminus \{\widehat{0},\widehat{1}\}) \cup \{\widehat{0},\widehat{1}\}$ is a graded Cohen-Macaulay poset.*

This Conjecture, inspired by Proposition 5.1, is also supported by the following.

THEOREM 6.3. *Assume that P is any poset whose order complex has a non-negative h-vector and that Q is a graded posets with a non-negative flag h-vector. Then the flag h-vector of the graded poset $P \rtimes (Q \setminus \{\widehat{0},\widehat{1}\}) \cup \{\widehat{0},\widehat{1}\}$ is non-negative.*

PROOF. Assume that the rank of Q is $n+1$ and that the dimension of $\triangle(P)$ is $(d-1)$. Then, by Proposition 6.2, the rank of $\widetilde{Q} := P \rtimes (Q \setminus \{\widehat{0},\widehat{1}\}) \cup \{\widehat{0},\widehat{1}\}$ is also $n+1$. For any $S \subseteq \{1,\ldots,n\}$, the saturated chains in $\widetilde{Q}_S$ are all sets of the form $\{(p_1,q_1),\ldots,(p_{|S|},q_{|S|})\}$, where $q_1 < \cdots < q_{|S|}$ is a saturated chain in Q_S and $p_1 \leq \cdots \leq p_{|S|}$ is any multichain in P. Thus we obtain

$$f_S(\widetilde{Q}) = f_S(Q) \sum_{j=1}^{\min(d,|S|)} f_{j-1}(P)\binom{j+|S|-j-1}{|S|-j} = f_S(Q) \sum_{j=1}^{\min(d,|S|)} f_{j-1}(P)\binom{|S|-1}{|S|-j} \quad \text{for } S \neq \emptyset.$$

Here, by abuse of notation, we write $f_{j-1}(P)$ as a shorthand for $f_{j-1}(\triangle(P))$. After some straightforward manipulation, which we omit for brevity's sake, we may rewrite these equations as

$$h_S(\widetilde{Q}) = \sum_{T\subseteq S} h_T(Q) \sum_{i=0}^{\min(d,|R|)} h_i(P)\binom{d+|T|-i-1}{|S|-i} \tag{6.1}$$

expressing the $h_S(\widetilde{Q})$'s as non-negative combinations of products of the (flag) h-entries of the original posets. □

7. Concluding remarks

There are two seemingly plausible generalizations that will *not* work.

REMARK 7.1. *It is not possible to generalize the definition of Jacobi posets in such a way that the polynomial on the left hand side of (3.2) became $P^{(\alpha,\beta)}(x)$ for some nonzero α. For any graded poset P of rank $n+1$, substituting $x=1$ into*

$$\sum_{j=0}^{n} f_{j-1}\left(\triangle\left(P\setminus\{\widehat{0},\widehat{1}\}\right)\right)\cdot\left(\frac{x-1}{2}\right)^j$$

*yields 1, while $P^{(\alpha,\beta)}(1) = \binom{n+\alpha}{\alpha}$ (see Chihara [**5**, Chapter V, (2.9)]) is 1 if and only if $\alpha = 0$.*

REMARK 7.2. *Sequences of symmetric orthogonal polynomials represented in the form*

$$\phi_P(x) := \sum_{j=0}^{n} f_{j-1}\left(\triangle\left(P\setminus\{\widehat{0},\widehat{1}\}\right)\right)\cdot\left(\frac{x-1}{2}\right)^j$$

*associated to certain Eulerian graded posets P appear in the paper [**9**] of the present author. A polynomial $p(x)$ of degree n is symmetric if it satisfies $P_n(x) = (-1)^n P_n(-x)$. A graded partially ordered set P is Eulerian* *if it satisfies* $\sum_{x\leq z\leq y}(-1)^{\rho(x,z)}=0$ *for all* $[x,y]\subseteq P$ *of rank at least* 1*. If a graded poset* P *is Eulerian then* $\phi_P(x)$ *is symmetric. A symmetric Jacobi polynomial* $P_n^{(\alpha,\beta)}(x)$ *satisfies* $\alpha=\beta$*. In fact, by* $P_n^{(\alpha,\beta)}(-x)=(-1)^n\cdot P_n^{(\beta,\alpha)}(-x)$ *(see Chihara* [**5**, Chapter V, (2.8)]*), it must satisfy* $P_n^{(\alpha,\beta)}(x)=P_n^{(\beta,\alpha)}(x)$*, so* $\alpha=\beta$ *follows from* $P^{(\alpha,\beta)}(1)=\binom{n+\alpha}{\alpha}$ *cited in Remark 7.1. By Remark 7.1 we may represent a Jacobi polynomial* $P_n^{(\alpha,\beta)}(x)$ *as* $\phi_P(x)$ *associated to some poset* P *only if* $\alpha=0$*. Therefore the only Jacobi polynomials that could be represented as* $\phi_P(x)$ *for some Eulerian graded poset* P *are the Legendre polynomials. It is not difficult to construct such posets for* $P_n^{(0,0)}(x)$ *for* $n\leq 2$*. However, for higher values of* n *we would need graded Eulerian poset of rank* $n+1$ *with* $f_{n-1}=\binom{2n}{n}$ *saturated chains, which is not an integer multiple of* $2^{\lfloor n/2\rfloor}$ *for* $n\geq 3$*. This makes constructing Eulerian "Legendre posets" of rank higher than* 3 *impossible, since the number of saturated chains of an Eulerian poset of rank* $n+1$ *is*

$$f_{\{1,\ldots,n\}}=2^{\lfloor n/2\rfloor}\cdot f_{2\cdot\mathbb{Z}\cap\{1,\ldots,n\}}.$$

This follows from the fact that, in an Eulerian poset, every interval of rank 2 *has* 4 *elements.*

The two areas, where the most interesting generalizations seem to be found, are the following. The right-strict direct product, introduced in Section 6, deserves further study. If Conjecture 6.2 turns out to be too hard or false, the proof of Proposition 5.1 may be a hint that preservation of EL-shellability (or a similar property) could be or should be shown instead. The other challenge is to find an algebraic generalization of Theorem 4.3. When $\triangle$ is balanced and colored with n-colors, its face ring is $\mathbb{Z}^n$-graded. The balanced join operation takes the tensor product of two $\mathbb{Z}^n$-graded rings and factors it by all terms of the form $u\otimes v$, where u and v are homogeneous terms of the same multi-degree. It is natural to ask whether such a factor of the tensor product of two $\mathbb{Z}^n$-graded Cohen-Macaulay modules would always have the Cohen Macaulay property.

References

[1] C. Banderier and S. Schwer, Why Delannoy numbers ?, preprint 2004, `http://xxx.lanl.gov/abs/math/0411128`, to appear in *J. Statist. Plan. Inference.*

[2] A. Björner, P. Frankl, and R. Stanley, The number of faces of balanced Cohen-Macaulay complexes and a generalized Macaulay theorem, *Combinatorica* **7** (1) (1987), 23–34.

[3] A. Björner and M. Wachs, Bruhat order of Coxeter groups and shellability, *Adv. in Math.* **43** (1982), 87–100.

[4] A. Björner and M. Wachs, On lexicographically shellable posets, *Trans. AMS* **277** (1983), 323–341.

[5] T.S. Chihara, "An Introduction to Orthogonal Polynomials," Gordon and Breach Science Publishers, New York-London-Paris, 1978.

[6] L. Comtet, "Advanced Combinatorics," D. Reidel Publishing Co., 1974.

[7] H. Delannoy, Employ d'échiquier pour la résolution de certains problèmes de probabilités, *Assoc. Franc. Bordeaux* **24** (1895), 70–90.

[8] I.J. Good, Legendre polynomials and trinomial random walks, *Proc. Cambridge Philos. Soc.* **54** (1958), 39–42.

[9] G. Hetyei, Orthogonal polynomials represented by CW-spheres, *Electron. J. Combin.* **11(2)** (2004), #R4, 28 pp.

[10] D.F. Lawden, On the solution of a linear difference equation, *Math. Gaz.* **36** (1952), 193–196.

[11] L. Moser and W. Zayachkowski, Lattice paths with diagonal steps, *Scripta Math.* **26** (1963), 223–229.

[12] M. Skandera, Dumont's statistic on words, *Electron. J. Combin* **8** (2001), #R11.

[13] N.J.A. Sloane, "On-Line Encyclopedia of Integer Sequences," `http://www.research.att.com/~njas/sequences`

[14] R.P. Stanley, Balanced Cohen–Macaulay complexes, *Transactions of the AMS* **Vol. 249** (1979), No. 1, 139-157.

[15] R.P. Stanley, "Combinatorics and Commutative Algebra," Second Edition, Birkhäuser, Boston, 1996.

[16] R.P. Stanley, "Enumerative Combinatorics, Volume II," Cambridge University Press, Cambridge, 1999.

[17] R.A. Sulanke, Objects counted by the central Delannoy numbers, *J. Integer Seq.* **6** (2003), Article 03.1.5, 19 pp.

[18] H.S. Wilf, "generatingfunctionology," Academic Press, 1994.

Department of Mathematics and Statistics, UNC Charlotte, Charlotte, NC 28223

E-mail address: `ghetyei@uncc.edu`

URL: `http://www.math.uncc.edu/ ghetyei`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Representation theories of some towers of algebras related to the symmetric groups and their Hecke algebras

Florent Hivert and Nicolas M. Thiéry

Abstract. We study the representation theory of three towers of algebras which are related to the symmetric groups and their Hecke algebras. The first one is constructed as the algebras generated simultaneously by the elementary transpositions and the elementary sorting operators acting on permutations. The two others are the monoid algebras of nondecreasing functions and nondecreasing parking functions. For these three towers, we describe the structure of simple and indecomposable projective modules, together with the Cartan map. The Grothendieck algebras and coalgebras given respectively by the induction product and the restriction coproduct are also given explicitly. This yields some new interpretations of the classical bases of quasi-symmetric and noncommutative symmetric functions as well as some new bases.

Résumé. Nous étudions la théorie des représentations de trois tours d'algèbres liées aux groupes symétriques et à leurs algèbres de Hecke. La première est formée des algèbres engendrées par les transpositions élémentaires ainsi que les opérateurs de tris élémentaires agissant sur les permutations. Les deux autres sont formées des algèbres des monoïdes des fonctions croissantes et des fonctions de parking croissantes. Pour ces trois tours, nous donnons la structure des modules simples et projectifs indécomposables ainsi que l'application de Cartan. Nous calculons également explicitement les algèbres et cogèbres de Grothendieck pour le produit d'induction et le coproduit de restriction. Il en découle de nouvelles interprétations de bases connues des fonctions quasi-symétriques et symétriques noncommutatives ainsi que des nouvelles bases.

Contents

1. Introduction

Given an *inductive tower of algebras*, that is a sequence of algebras

(1) $$A_0 \hookrightarrow A_1 \hookrightarrow \cdots \hookrightarrow A_n \hookrightarrow \cdots ,$$

with embeddings $A_m \otimes A_n \hookrightarrow A_{m+n}$ satisfying an appropriate associativity condition, one can introduce two *Grothendieck rings*

(2) $$\mathcal{G}(A) := \bigoplus_{n\geq 0} G_0(A_n) \qquad \text{and} \qquad \mathcal{K}(A) := \bigoplus_{n\geq 0} K_0(A_n)\,,$$

1991 *Mathematics Subject Classification.* Primary 16G99; Secondary 05E05.
Key words and phrases. Representation theory, towers of algebras, Grothendieck groups, symmetric groups, Hecke algebras, Quasi-symmetric and Noncommutative symmetric functions.

where $G_0(A)$ and $K_0(A)$ are the (complexified) Grothendieck groups of the categories of finite-dimensional A-modules and projective A-modules respectively, with the multiplication of the classes of an A_m-module M and an A_n-module N defined by the induction product

$$[M] \cdot [N] = [M \widehat{\otimes} N] = [M \otimes N \uparrow_{A_m \otimes A_n}^{A_{m+n}}] \,. \tag{3}$$

If A_{m+n} is a projective $A_m \otimes A_n$ modules, one can define a coproduct on these rings by means of restriction of representations, turning these into coalgebras. Under favorable circumstances the product and the coproduct are compatible turning these into mutually dual Hopf algebras.

The basic example of this situation is the character ring of the symmetric groups (over $\mathbb{C}$), due to Frobenius. Here the $A_n := \mathbb{C}[\mathfrak{S}_n]$ are semi-simple algebras, so that

$$G_0(A_n) = K_0(A_n) = R(A_n) \,, \tag{4}$$

where $R(A_n)$ denotes the vector space spanned by isomorphism classes of indecomposable modules which, in this case, are all simple and projective. The irreducible representations $[\lambda]$ of A_n are parametrized by partitions λ of n, and the Grothendieck ring is isomorphic to the algebra Sym of symmetric functions under the correspondence $[\lambda] \leftrightarrow s_\lambda$, where s_λ denotes the Schur function associated with λ. Other known examples with towers of group algebras over the complex numbers $A_n := \mathbb{C}[G_n]$ include the cases of wreath products $G_n := \Gamma \wr \mathfrak{S}_n$ (Specht), finite linear groups $G_n := GL(n, \mathbb{F}_q)$ (Green), *etc.*, all related to symmetric functions (see [**11, 16**]).

Examples involving non-semisimple specializations of Hecke algebras have also been worked out. Finite Hecke algebras of type A at roots of unity ($A_n = H_n(\zeta)$, $\zeta^r = 1$) yield quotients and subalgebras of Sym [**10**]. The Ariki-Koike algebras at roots of unity give rise to level r Fock spaces of affine Lie algebras of type A [**2**]. The 0-Hecke algebras $A_n = \mathrm{H}_n(0)$ correspond to the pair Quasi-symmetric functions / Noncommutative symmetric functions, $\mathcal{G} = \mathrm{QSym}$, $\mathcal{K} = \mathrm{NCSF}$ [**9**]. Affine Hecke algebras at roots of unity lead to $U(\widehat{sl}_r)$ and $U(\widehat{sl}_r)^*$ [**1**], and the case of affine Hecke generic algebras can be reduced to a subcategory admitting as Grothendieck rings $U(\widehat{gl}_\infty)$ and $U(\widehat{gl}_\infty)^*$ [**1**]. Further interesting examples are the tower of 0-Hecke-Clifford algebras [**13, 3**] giving rise to the peak algebras [**15**], and a degenerated version of the Ariki-Koike algebras [**7**] giving rise to a colored version of QSym and NCSF.

The goal of this article is to study the representation theories of several towers of algebras which are related to the symmetric groups and their Hecke algebras $\mathrm{H}_n(q)$. We describe their representation theory and the Grothendieck algebras and coalgebras arising from them. Here is the structure of the paper together with the main results.

In Section 3, we introduce the main object of this paper, namely a new tower of algebras denoted $\mathrm{H}\mathfrak{S}_n$. Each $\mathrm{H}\mathfrak{S}_n$ is constructed as the algebra generated by both elementary transpositions and elementary sorting operators acting on permutations of $\{1, \ldots, n\}$. We show that this algebra is better understood as the algebra of antisymmetry preserving operators; this allows us to compute its dimension and give an explicit basis. Then, we construct the projective and simple modules and compute their restrictions and inductions. This gives rise to a new interpretation of some bases of quasi-symmetric and noncommutative symmetric functions in representation theory. The Cartan matrix suggests a link between $\mathrm{H}\mathfrak{S}_n$ and the incidence algebra of the boolean lattice. We actually show that these algebra are Morita equivalent. We conclude this section by discussing some links with a certain central specialization of the affine Hecke algebra.

In Sections 4 and 5 we turn to the study of two other towers, namely the towers of the monoids algebras of nondecreasing functions and of nondecreasing parking functions. In both cases, we give the structure of projective and simple modules, the cartan matrices, and the induction and restrictions rules. We also show that the algebra of nondecreasing parking functions is isomorphic to the incidence algebra of some lattice. Finally, we prove that those two algebras are the respective quotients of $\mathrm{H}\mathfrak{S}_n$ and $\mathrm{H}_n(0)$, through their representations on exterior powers. The following diagram summarizes the relations between all the mentioned towers of algebras:

$$\begin{tikzcd}
\mathrm{H}_n(-1) \arrow[rrrr, hook, bend left] \arrow[d] & \mathrm{H}_n(0) \arrow[rrr, hook, bend left] \arrow[d] & \mathrm{H}_n(1) = \mathbb{C}[\mathfrak{S}_n] \arrow[rr, hook, bend left] \arrow[d] & \mathrm{H}_n(q) \arrow[r, hook] \arrow[d] & H\mathfrak{S}_n \arrow[d] \\
\text{Temperley-Lieb}_n \arrow[rrrr, hook, bend right] & \mathbb{C}[\mathrm{NDPF}_n] \arrow[rrr, hook, bend right] & \mathbb{C}[\mathfrak{S}_n] \hookrightarrow \bigwedge^{\cdot} \mathbb{C}^n \arrow[rr, hook, bend right] & \mathrm{H}_n(q) \hookrightarrow \bigwedge^{\cdot} \mathbb{C}^n \arrow[r, hook] & \mathbb{C}[\mathrm{NDF}_n]
\end{tikzcd} \tag{5}$$

This paper mostly reports on a computation driven research using the package `MuPAD-Combinat` by the authors of the present paper [**8**]. This package is designed for the computer algebra system `MuPAD` and is freely available from `http://mupad-combinat.sf.net/`. Among other things, it allows to automatically compute the dimensions of simple and indecomposable projective modules together with the Cartan invariants matrix of a finite dimensional algebra, knowing its multiplication table.

2. Background

2.1. Compositions and sets. Let n be a fixed integer. Recall that each subset S of $\{1,\ldots,n-1\}$ can be uniquely identified with a p-tuple $K := (k_1,\ldots,k_p)$ of positive integers of sum n:

$$S = \{i_1 < i_2 < \cdots < i_p\} \longmapsto \mathrm{C}(S) := (i_1, i_2 - i_1, i_3 - i_2, \ldots, n - i_p)\,. \tag{6}$$

We say that K is a *composition of* n and we write it by $K \vDash n$. The converse bijection, sending a composition to its *descent set*, is given by:

$$K = (k_1,\ldots,k_p) \longmapsto \mathrm{Des}(K) = \{k_1 + \cdots + k_j,\ j = 1,\ldots,p-1\}\,. \tag{7}$$

The number p is called the *length* of K and is denoted by $\ell(K)$.

The notions of complementary of a set S^c and of inclusion of sets can be transfered to compositions, leading to the complementary of a composition K^c and to the refinement order on compositions: we say that I is *finer* than J, and write $I \succeq J$, if and only if $\mathrm{Des}(I) \supseteq \mathrm{Des}(J)$.

2.2. Symmetric groups and Hecke algebras. Take $n \in \mathbb{N}$ and let $\mathfrak{S}_n$ be the n-th symmetric group. It is well known that it is generated by the $n-1$ elementary transpositions σ_i which exchange i and $i+1$, with the relations

$$\begin{aligned} \sigma_i^2 &= 1 && (1 \le i \le n-1)\,,\\ \sigma_i\sigma_j &= \sigma_j\sigma_i && (|i-j| \ge 2)\,,\\ \sigma_i\sigma_{i+1}\sigma_i &= \sigma_{i+1}\sigma_i\sigma_{i+1} && (1 \le i \le n-2)\,. \end{aligned} \tag{8}$$

The last two relations are called the *braids relations*. A *reduced word* for a permutation μ is a decomposition $\mu = \sigma_{i_1}\cdots\sigma_{i_k}$ of minimal length. When denoting permutations we also use the *word notation*, where μ is denoted by the word $\mu_1\mu_2\cdots\mu_n := \mu(1)\mu(2)\cdots\mu(n)$. For a permutation μ, the set $\{i,\, \mu_i > \mu_{i+1}\}$ of its *descents* is denoted $\mathrm{Des}(\mu)$. The descents of the inverse of μ are called the *recoils of* μ and their set is denoted $\mathrm{Rec}(\mu)$. For a composition I, we denote by $\mathfrak{S}_I := \mathfrak{S}_{i_1} \times \cdots \times \mathfrak{S}_{i_p}$ the *standard Young subgroup* of $\mathfrak{S}_n$, which is generated by the elementary transpositions σ_i where $i \notin \mathrm{Des}(I)$.

Recall that the (Iwahori-) *Hecke algebra* $\mathrm{H}_n(q)$ of type A_{n-1} is the $\mathbb{C}$-algebra generated by elements T_i for $i < n$ with the braids relations together with the quadratic relations:

$$T_i^2 = (q-1)T_i + q\,, \tag{9}$$

where q is a complex number.

The 0-Hecke algebra is obtained by setting $q = 0$ in these relations. Then, the first relation becomes $T_i^2 = -T_i$ [**12, 9**]. In this paper, we prefer to use another set of generators $(\pi_i)_{i=1\ldots n-1}$ defined by $\pi_i := T_i + 1$. They also satisfy the braids relations together with the quadratic relations $\pi_i^2 = \pi_i$.

Let $\sigma =: \sigma_{i_1}\cdots\sigma_{i_p}$ be a reduced word for a permutation $\sigma \in \mathfrak{S}_n$. The defining relations of $\mathrm{H}_n(q)$ ensures that the element $T_\sigma := T_{i_1}\cdots T_{i_p}$ (resp.: $\pi_\sigma := \pi_{i_1}\cdots\pi_{i_p}$) is independent of the chosen reduced word for σ. Moreover, the well-defined family $(T_\sigma)_{\sigma\in\mathfrak{S}_n}$ (resp.: $(\pi_\sigma)_{\sigma\in\mathfrak{S}_n}$) is a basis of the Hecke algebra, which is consequently of dimension $n!$.

2.3. Representation theory. In this paper, we mostly consider *right* modules over algebras. Consequently the composition of two endomorphisms f and g is denoted by $fg = g \circ f$ and their action on a vector v is written $v \cdot f$. Thus $g \circ f(v) = g(f(v))$ is denoted $v \cdot fg = (v \cdot f) \cdot g$.

It is known that $\mathrm{H}_n(0)$ has 2^{n-1} simple modules, all one-dimensional, and naturally labelled by compositions I of n [**12**]: following the notation of [**9**], let η_I be the generator of the simple $\mathrm{H}_n(0)$-module S_I associated with I in the left regular representation. It satisfies

$$\eta_I \cdot T_i := \begin{cases} -\eta_I & \text{if } i \in \mathrm{Des}(I),\\ 0 & \text{otherwise,}\end{cases} \qquad \text{or equivalently} \qquad \eta_I \cdot \pi_i := \begin{cases} 0 & \text{if } i \in \mathrm{Des}(I),\\ \eta_I & \text{otherwise.}\end{cases} \tag{10}$$

The bases of the indecomposable projective modules P_I associated to the simple module S_I of $\mathrm{H}_n(0)$ are indexed by the permutations σ whose descents composition is I.

The Grothendieck rings of $\mathrm{H}_n(0)$ are naturally isomorphic to the dual pair of Hopf algebras of quasi-symmetric functions QSym of Gessel [**6**] and of noncommutative symmetric functions NCSF [**5**] (see [**9**]). The reader who is not familiar with those should refer to these papers, as we will only recall the required notations here.

The Hopf algebra QSym of quasi-symmetric functions has two remarkable bases, namely the *monomial basis* $(M_I)_I$ and the *fundamental basis* (also called *quasi-ribbon*) $(F_I)_I$. They are related by

$$F_I = \sum_{I \succeq J} M_J \qquad \text{or equivalently} \qquad M_I = \sum_{I \succeq J} (-1)^{\ell(I)-\ell(J)} F_J\,. \tag{11}$$

The characteristic map $S_I \mapsto F_I$ which sends the simple $\mathrm{H}_n(0)$ module S_I to its corresponding fundamental function F_I also sends the induction product to the product of QSym and the restriction coproduct to the coproduct of QSym.

The Hopf algebra NCSF of noncommutative symmetric functions [**5**] is a noncommutative analogue of the algebra of symmetric functions [**11**]. It has for multiplicative bases the analogues $(\Lambda^I)_I$ of the elementary symmetric functions $(e_\lambda)_\lambda$ and as well as the analogues $(S^I)_I$ of the complete symmetric functions $(h_\lambda)_\lambda$. The relevant basis in the representation theory of $\mathrm{H}_n(0)$ is the basis of so called *ribbon Schur functions* $(R_I)_I$ which is an analogue of skew Schur functions of ribbon shape. It is related to $(\Lambda_I)_I$ and $(S_I)_I$ by

$$S_I = \sum_{I \succeq J} R_J \qquad \text{and} \qquad \Lambda_I = \sum_{I \succeq J} R_{J^c}\,. \tag{12}$$

Their interpretation in representation theory goes as follows. The complete function S^n is the characteristic of the trivial module $S_n \approx P_n$, the elementary function Λ^n being the characteristic of the sign module $S_{1^n} \approx P_{1^n}$. An arbitrary indecomposable projective module P_I has R_I for characteristic. Once again the map $P_I \mapsto R_I$ is an isomorphism of Hopf algebras.

Recall that S_J is the semi-simple module associated to P_I, giving rise to the duality between $\mathcal{G}$ and $\mathcal{K}$:

$$S_I = P_J/\mathrm{rad}(P_J) \qquad \text{and} \qquad \langle P_I\ ,\ S_J\rangle = \delta_{I,J} \tag{13}$$

This translates into QSym and NCSF by setting that $(F_I)_I$ and $(R_I)_I$ are dual bases, or equivalently that $(M_I)_I$ and $(S^I)_I$ are dual bases.

3. The algebra $\mathrm{H}\mathfrak{S}_n$

The algebra of the symmetric group $\mathbb{C}[\mathfrak{S}_n]$ and the 0-Hecke algebra $\mathrm{H}_n(0)$ can be realized simultaneously as operator algebras by identifying the underlying vector spaces of their right regular representations.

Namely, consider the plain *vector space* $\mathbb{C}\mathfrak{S}_n$ (distinguished from the *group algebra* which is denoted by $\mathbb{C}[\mathfrak{S}_n]$). On the first hand, the algebra $\mathbb{C}[\mathfrak{S}_n]$ acts naturally on $\mathbb{C}\mathfrak{S}_n$ by multiplication on the right (action on positions). That is, a transposition σ_i acts on a permutation $\mu := (\mu_1, \dots, \mu_n)$ by permuting μ_i and μ_{i+1}: $\mu \cdot \sigma_i = \mu\sigma_i$.

On the other hand, the 0-Hecke algebra $\mathrm{H}_n(0)$ acts on the right on $\mathbb{C}\mathfrak{S}_n$ by decreasing sort. That is, π_i acts on the right on μ by:

$$\mu \cdot \pi_i = \begin{cases} \mu & \text{if } \mu_i > \mu_{i+1}, \\ \mu\sigma_i & \text{otherwise.} \end{cases} \tag{14}$$

DEFINITION 1. For each n, the algebra $\mathrm{H}\mathfrak{S}_n$ is the subalgebra of $\mathrm{End}(\mathbb{C}\mathfrak{S}_n)$ generated by both sets of operators $\sigma_1, \dots, \sigma_{n-1}, \pi_1, \dots, \pi_{n-1}$.

By construction, the algebra $\mathrm{H}\mathfrak{S}_n$ contains both $\mathbb{C}[\mathfrak{S}_n]$ and $\mathrm{H}_n(0)$. In fact, it contains simultaneously all the Hecke algebras $\mathrm{H}_n(q)$ for all values of q; each one can be realized by taking the subalgebra generated by the operators:

$$T_i := (q-1)(1-\pi_i) + q\sigma_i, \qquad \text{for } i = 1, \dots, n-1\,. \tag{15}$$

The natural embedding of $\mathbb{C}\mathfrak{S}_n \otimes \mathbb{C}\mathfrak{S}_m$ in $\mathbb{C}\mathfrak{S}_{n+m}$ makes $(\mathrm{H}\mathfrak{S}_n)_{n\in\mathbb{N}}$ into a tower of algebras, which contains the similar towers of algebras $(\mathbb{C}[\mathfrak{S}_n])_{n\in\mathbb{N}}$ and $(\mathrm{H}_n(q))_{n\in\mathbb{N}}$.

3.1. Basic properties of $\mathrm{H}\mathfrak{S}_n$. Let $\overline{\pi}_i$ be the *increasing sort operator* on $\mathbb{C}\mathfrak{S}_n$. Namely: $\overline{\pi}_i$ acts on the right on μ by:

$$\mu \cdot \overline{\pi}_i = \begin{cases} \mu & \text{if } \mu_i < \mu_{i+1}, \\ \mu\sigma_i & \text{otherwise.} \end{cases} \tag{16}$$

Since $\pi_i + \overline{\pi}_i$ is a symmetrizing operator, we have the identity:

$$\pi_i + \overline{\pi}_i = 1 + \sigma_i \ . \tag{17}$$

It follows that the operator $\overline{\pi}_i$ also belongs to $\mathrm{H}\mathfrak{S}_n$.

The following identities are also easily checked:

$$\begin{aligned} \sigma_i\pi_i = \pi_i \ , \quad \sigma_i\overline{\pi}_i = \overline{\pi}_i \ , \\ \overline{\pi}_i\pi_i = \pi_i \ , \quad \pi_i\overline{\pi}_i = \overline{\pi}_i \ , \\ \pi_i\sigma_i = \overline{\pi}_i \ , \quad \overline{\pi}_i\sigma_i = \pi_i \ . \end{aligned} \tag{18}$$

A computer exploration suggests that the dimension of $\mathrm{H}\mathfrak{S}_n$ is given by the following sequence (sequence A000275 of the encyclopedia of integer sequences [**14**]):

$$1, 1, 3, 19, 211, 3651, 90921, 3081513, 136407699, 7642177651, 528579161353, 44237263696473, \dots$$

These are the numbers h_n of pairs (σ, τ) of permutations such that $\mathrm{Des}(\sigma) \cap \mathrm{Des}(\tau) = \emptyset$. Together with Equation (18), this leads to state the following

THEOREM 3.1. *A vector space basis of* $\mathrm{H}\mathfrak{S}_n$ *is given by the family of operators*

$$B_n := \left\{ \sigma\pi_\tau \mid \mathrm{Des}(\sigma) \cap \mathrm{Des}(\tau^{-1}) = \emptyset \right\} \ . \tag{19}$$

One approach to prove this theorem would be to find a presentation of the algebra. The following relations are easily proved to hold in $\mathrm{H}\mathfrak{S}_n$:

$$\begin{aligned} \pi_{i+1}\sigma_i &= \pi_{i+1}\pi_i + \sigma_i\sigma_{i+1}\pi_i\pi_{i+1} - \pi_i\pi_{i+1}\pi_i \ , \\ \pi_i\sigma_{i+1} &= \pi_i\pi_{i+1} + \sigma_{i+1}\sigma_i\pi_{i+1}\pi_i - \pi_i\pi_{i+1}\pi_i \ , \\ \sigma_1\pi_2\sigma_1 &= \sigma_2\pi_1\sigma_2 \ , \end{aligned} \tag{20}$$

and we conjecture that they generate all relations.

CONJECTURE 1. A presentation of $\mathrm{H}\mathfrak{S}_n$ is given by the defining relations of $\mathbb{C}[\mathfrak{S}_n]$ and $\mathrm{H}_n(0)$ together with the relations $\sigma_i\pi_i = \pi_i$ and of Equations (20).

Using those relations as rewriting rules yields a straightening algorithm which rewrites any expression in the σ_i's and π_i's into a linear combination of the $\sigma\pi_\tau$. This algorithm seems, in practice and with an appropriate strategy, to always terminate. However we have no proof of this fact; moreover this algorithm is not efficient, due to the explosion of the number and length of words in intermediate results.

This is a standard phenomenon with such algebras. Their properties often become clearer when considering their concrete representations (typically as operator algebras) rather than their abstract presentation. Here, theorem 3.1 as well as the representation theory of $\mathrm{H}\mathfrak{S}_n$ follow from its upcoming structural characterization as the algebra of operators preserving certain anti-symmetries.

3.2. $\mathrm{H}\mathfrak{S}_n$ as algebra of antisymmetry-preserving operators. Let $\overline{\sigma}_i$ be the *right operator* in $\mathrm{End}(\mathbb{C}\mathfrak{S}_n)$ describing the action of s_i by multiplication *on the left* (action on values), namely $\overline{\sigma}_i$ is defined by

$$\sigma \cdot \overline{\sigma}_i := \sigma_i\sigma \ . \tag{21}$$

A vector v in $\mathbb{C}\mathfrak{S}_n$ is *left i-symmetric* (resp. *antisymmetric*) if $v \cdot \overline{\sigma}_i = v$ (resp. $v \cdot \overline{\sigma}_i = -v$). The subspace of left i-symmetric (resp. antisymmetric) vectors can be alternatively described as the image (resp. kernel) of the idempotent operator $\frac{1}{2}(1 + \overline{\sigma}_i)$, or as the kernel (resp. image) of the idempotent operator $\frac{1}{2}(1 - \overline{\sigma}_i)$.

THEOREM 3.2. $\mathrm{H}\mathfrak{S}_n$ *is the subspace of* $\mathrm{End}(\mathbb{C}\mathfrak{S}_n)$ *defined by the* $n-1$ idempotent sandwich equations*:*

$$\frac{1}{2}(1 - \overline{\sigma}_i) f \frac{1}{2}(1 + \overline{\sigma}_i) = 0, \qquad \text{for } i = 1, \dots, n-1 \ . \tag{22}$$

In other words, $\mathrm{H}\mathfrak{S}_n$ *is the subalgebra of those operators in* $\mathrm{End}(\mathbb{C}\mathfrak{S}_n)$ *which preserve left anti-symmetries.*

Note that, $\overline{\sigma}_i$ being self-adjoint, the adjoint algebra of $\mathrm{H}\mathfrak{S}_n$ satisfies the equations:

$$\frac{1}{2}(1+\overline{\sigma}_i)f\frac{1}{2}(1-\overline{\sigma}_i) = 0; \tag{23}$$

thus, it is the subalgebra of those operators in $\mathrm{End}(\mathbb{C}\mathfrak{S}_n)$ which preserve left symmetries. The symmetric group algebra has a similar description as the subalgebra of those operators in $\mathrm{End}(\mathbb{C}\mathfrak{S}_n)$ which preserve both left symmetries and antisymmetries.

Proof. The proof of theorem 3.2 proceeds as follow. We first exhibit a triangularity property of the operators in B_n; this proves that they are linearly independent, so that $\dim \mathrm{H}\mathfrak{S}_n \geq h_n$. Let $<$ be any linear extension of the right permutahedron order. Given an endomorphism f of $\mathbb{C}\mathfrak{S}_n$, we order the rows and columns of its matrix $M := [f_{\mu\nu}]$ accordingly to $<$, and denote by $\mathrm{init}(f) := \min\{\mu, \exists\nu, f_{\mu\nu} \neq 0\}$ the index of the first non zero row of M.

Lemma 3.1. *(a) Let $f := \sigma\pi_\tau$ in B_n. Then,* $\mathrm{init}(f) = \tau$, *and*

$$f_{\tau\nu} = \begin{cases} 1 & \textit{if } \nu \in \mathfrak{S}_{\mathrm{Des}(\tau^{-1})}\sigma^{-1} \\ 0 & \textit{otherwise} \end{cases} \tag{24}$$

(b) The family B_n is free.

Then, we note that $\mathrm{H}\mathfrak{S}_n$ preserves all antisymmetries, because its generators σ_i and π_i do. It follows that $\mathrm{H}\mathfrak{S}_n$ satisfies the sandwich equations. We conclude by giving an explicit description of the sandwich equations. Given an endomorphism f of $\mathbb{C}\mathfrak{S}_n$, denote by $(f_{\mu,\nu})_{\mu,\nu}$ the coefficients of its matrix in the natural permutation basis. Given two permutations μ, ν, and an integer i in $\{1, \dots, n-1\}$, let $R_{\mu,\nu,i}$ be the linear form:

$$R_{\mu,\nu,i} : \begin{cases} \mathrm{End}(\mathbb{C}\mathfrak{S}_n) & \mapsto \mathbb{C} \\ f & \mapsto f_{\mu,\nu} + f_{s_i\mu,\nu} - f_{\mu,s_i\nu} + f_{s_i\mu,s_i\nu} \end{cases} \tag{25}$$

Given a pair of permutations μ, ν having at least one descent in common, set $R_{\mu,\nu} = R_{\mu,\nu,i}$, where i is the smallest common descent of μ and ν (the choice of the common descent i is, in fact, irrelevant). Finally, let $R_n := \{R_{\mu,\nu},\ \mathrm{Des}(\mu) \cap \mathrm{Des}(\nu) \neq \emptyset\}$.

Lemma 3.2. *(a) If an operator f in $\mathrm{End}\mathbb{C}\mathfrak{S}_n$ preserves i-antisymmetries, then $R_{\mu,\nu,i}(f) = 0$ for any permutations μ and ν.*
(b) The $n!^2 - h_n$ linear relations in R_n are linearly independent.

Theorems 3.1 and 3.2 follow. □

3.3. The representation theory of $\mathrm{H}\mathfrak{S}_n$.

3.3.1. *Projective modules of* $\mathrm{H}\mathfrak{S}_n$. Recall that $\mathrm{H}\mathfrak{S}_n$ is the algebra of operators preserving left antisymmetries. Thus, given $S \subset \{1, \dots, n-1\}$, it is natural to introduce the $\mathrm{H}\mathfrak{S}_n$-submodule $\bigcap_{i\in S}\ker(1+\overline{\sigma}_i)$ of the vectors in $\mathbb{C}\mathfrak{S}_n$ which are i-antisymmetric for all $i \in S$. For the ease of notations, it turns out to be better to index this module by the composition associated to the *complementary set*; thus we define

$$P_I := \bigcap_{i \notin \mathrm{Des}(I)} \ker(1+\overline{\sigma}_i)\,. \tag{26}$$

The goal of this section is to prove that the family of modules $(P_I)_{I\models n}$ forms a complete set of representatives of the indecomposable projective modules of $\mathrm{H}\mathfrak{S}_n$.

The simplest element of P_I is:

$$v_I := \sum_{\nu\in\mathfrak{S}_I} (-1)^{l(\nu)}\nu, \tag{27}$$

One easily shows that

Lemma 3.3. *v_I generates P_I as an $\mathrm{H}\mathfrak{S}_n$-module.*

Given a permutation σ, let $v_\sigma := v_{\mathrm{Rec}(\sigma)}\sigma$ (recall that $\mathrm{Rec}(\sigma) = \mathrm{Des}(\sigma^{-1})$). Note that σ is the permutation of minimal length appearing in v_σ. By triangularity, it follows that the family $(v_\sigma)_{\sigma\in\mathfrak{S}_n}$ forms a vector space basis of $\mathbb{C}\mathfrak{S}_n$. The usefulness of this basis comes from the fact that

PROPOSITION 1. *For any composition $I := (i_1, \dots, i_k)$ of sum n, the families*

$$\{v_I \cdot \sigma \mid \sigma \in \mathfrak{S}_n,\ \mathrm{Rec}(\sigma) \cap \mathrm{Des}(I) = \emptyset\} \qquad \text{and} \qquad \{v_\sigma \mid \sigma \in \mathfrak{S}_n,\ \mathrm{Rec}(\sigma) \cap \mathrm{Des}(I) = \emptyset\} \tag{28}$$

are both vector space bases of P_I; in particular, P_I is of dimension $\frac{n!}{i_1! i_2! \dots i_k!}$.

Since $\mathfrak{S}_n$ and $\mathrm{H}_n(0)$ are both sub-algebras of $\mathrm{H}\mathfrak{S}_n$, the space P_I is naturally a module over them. The following proposition elucidates its structure.

PROPOSITION 2. *Let (-1) denote the sign representation of the symmetric group as well as the corresponding representation of the Hecke algebra $\mathrm{H}_n(0)$ (sending T_i to -1, or equivalently π_i to 0).*

(a) *As a $\mathfrak{S}_n$ module, $P_I \approx (-1)\uparrow_{\mathfrak{S}_I}^{\mathfrak{S}_n}$; its character is the symmetric function $e_I := e_{i_1} \cdots e_{i_k}$.*
(b) *As a $\mathrm{H}_n(0)$ module, $P_I \approx (-1)\uparrow_{\mathrm{H}_I(0)}^{\mathrm{H}_n(0)}$; it is a projective module whose character is the noncommutative symmetric function $\Lambda^I := \Lambda_{i_1} \cdots \Lambda_{i_k}$.*
(c) *In particular the P_I's are non isomorphic as $\mathrm{H}_n(0)$-modules and thus as $\mathrm{H}\mathfrak{S}_n$-modules.*

We are now in position to state the main theorem of this section.

THEOREM 3.3. *For $\sigma \in \mathfrak{S}_n$, let $p_\sigma \in \mathrm{End}(\mathbb{C}\mathfrak{S}_n)$ denote the projector on $\mathbb{C}v_\sigma$ parallel to $\oplus_{\tau\neq\sigma}\mathbb{C}v_\tau$. Then,*

(a) *The ideal $p_\sigma \mathrm{H}\mathfrak{S}_n$ is isomorphic to $P_{\mathrm{Rec}(\sigma)} = P_{\mathrm{Des}(\sigma^{-1})}$ as an $\mathrm{H}\mathfrak{S}_n$ module;*
(b) *The idempotents p_σ all belong to $\mathrm{H}\mathfrak{S}_n$; they give a maximal decomposition of the identity into orthogonal idempotents in $\mathrm{H}\mathfrak{S}_n$;*
(c) *The family of modules $(P_I)_{I\models n}$ forms a complete set of representatives of the indecomposable projective modules of $\mathrm{H}\mathfrak{S}_n$.*

PROOF. Item (a) is an easy consequence of Proposition 1. To prove (b) one needs to check that p_σ belongs to $\mathrm{H}\mathfrak{S}_n$. This is done by showing that it preserves left antisymmetries. Then, since the p_σ's give a maximal decomposition of the identity in $\mathrm{End}(\mathbb{C}\mathfrak{S}_n)$, they are as well a maximal decomposition of the identity in $\mathrm{H}\mathfrak{S}_n$. Finally, Item (c) follows from (a) and (b) and Item (c) of Proposition 2. □

3.3.2. *Simple modules.* The simple modules are obtained as quotients of the projective modules by their radical:

THEOREM 3.4. *The modules $S_I := P_I / \sum_{J\subsetneq I} P_J$ form a complete set of representatives of the simple modules of $\mathrm{H}\mathfrak{S}_n$. Moreover, the projection of the family $\{v_\sigma, \mathrm{Rec}(\sigma) = I\}$ in S_I forms a vector space basis of S_I.*

The modules S_I are closely related to the projective modules of the 0-Hecke algebra:

PROPOSITION 3. *The restriction of the simple module S_I to $\mathrm{H}_n(0)$ is an indecomposable projective module whose characteristic is the noncommutative symmetric function R_{I^c}.*

3.3.3. *Cartan's invariants matrix and the boolean lattice.* We now turn to the description of the Cartan matrix. Let $p_I := p_\alpha$ where α is the shortest permutation such that $\mathrm{Rec}(\alpha) = I$ (this choice is in fact irrelevant).

PROPOSITION 4. *Let I and J be two subsets of $\{1, \dots, n\}$. Then,*

$$\dim Hom(P_I, P_J) = \dim p_I \mathrm{H}\mathfrak{S}_n p_J = \begin{cases} 1 & \text{if } I \subset J, \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

In other words, the Cartan matrix of $\mathrm{H}\mathfrak{S}_n$ is the incidence matrix of the boolean lattice. This suggests that there is a close relation between $\mathrm{H}\mathfrak{S}_n$ and the incidence algebra of the boolean lattice. Recall that the *incidence algebra* $\mathbb{C}[P]$ of a partially ordered set $(P, \leq_P)$ is the algebra whose basis elements are indexed by the couples $(u, v) \in P^2$ such that $u \leq_P v$ with the multiplication rule

$$(u, v) \cdot (u', v') = \begin{cases} (u, v') & \text{if } v = u', \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

An algebra is called *elementary* (or sometimes reduced) if its simple modules are all one dimensional. Starting from an algebra A, it is possible to get a canonical elementary algebra by the following process. Start with a maximal decomposition of the identity $1 = \sum_i e_i$ into orthogonal idempotents. Two idempotents e_i and e_j are *conjugate* if e_i can be written as ae_jb where a and b belongs to A, or equivalently, if the projective modules e_iA and e_jA are isomorphic. Select an idempotent e_c in each conjugacy classes c and put $e := \sum e_c$. Then, it is well known [**4**] that the algebra eAe is elementary and that the functor $M \mapsto Me$ which sends a right A module to a eAe module is an equivalence of category. Recall finally that two algebra A and B such that the category of A-modules and B-modules are equivalent are said *Morita equivalent*. Thus A and eAe are Morita-equivalent.

Applying this to $\mathrm{H}\mathfrak{S}_n$, one gets

Theorem 3.5. *Let e be the idempotent defined by $e := \sum_{I \vDash n} p_I$. Then the algebra $e\mathrm{H}\mathfrak{S}_n e$ is isomorphic to the incidence algebra $\mathbb{C}[B_{n-1}]$ of the boolean lattice B_{n-1} of subsets of $\{1, \ldots, n-1\}$. Consequently, $\mathrm{H}\mathfrak{S}_n$ and $\mathbb{C}[B_{n-1}]$ are Morita equivalent.*

3.3.4. *Induction, restriction, and Grothendieck rings.* Let $\mathcal{G} := \mathcal{G}\left((\mathrm{H}\mathfrak{S}_n)_n\right)$ and $\mathcal{K} := \mathcal{K}\left((\mathrm{H}\mathfrak{S}_n)_n\right)$ be respectively the Grothendieck rings of the characters of the simple and projective modules of the tower of algebras $(\mathrm{H}\mathfrak{S}_n)_n$. Let furthermore C be the cartan map from $\mathcal{K}$ to $\mathcal{G}$. It is the algebra and coalgebra morphism which gives the projection of a module onto the direct sum of its composition factors. It is given by

$$C(P_I) = \sum_{I \succeq J} S_J\,. \tag{31}$$

Since the indecomposable projective modules are indexed by compositions, it comes out as no surprise that the structure of algebras and coalgebras of $\mathcal{G}$ and $\mathcal{K}$ are each isomorphic to QSym and NCSF. However, we do not get Hopf algebras, because the structures of algebras and coalgebras are not compatible.

Proposition 5. *The following diagram gives a complete description of the structures of algebras and of coalgebras on $\mathcal{G}$ and $\mathcal{K}$.*

$$\begin{tikzcd}
(\mathrm{QSym}, .) & (\mathcal{G}, .) \arrow[l, hook', two heads, "\chi(S_I)\mapsto M_{I^c}"'] & (\mathcal{K}, .) \arrow[l, hook', two heads, "C"'] \arrow[r, hook, two heads, "\chi(P_I)\mapsto F_{I^c}"] & (\mathrm{QSym}, .) \\
(\mathrm{NCSF}, \Delta) & (\mathcal{G}, \Delta) \arrow[l, hook', two heads, "\chi(S_I)\mapsto R_{I^c}"] & (\mathcal{K}, \Delta) \arrow[l, hook', two heads, "C"] \arrow[r, hook, two heads, "\chi(P_I)\mapsto \Lambda^I"'] & (\mathrm{NCSF}, \Delta)
\end{tikzcd} \tag{32}$$

Proof. The bottom line is already known from Proposition 2 and the fact that, for all m and n, the following diagram commutes

$$\begin{tikzcd}
\mathrm{H}_m(0) \otimes \mathrm{H}_n(0) \arrow[r, hook] \arrow[d, hook] & \mathrm{H}_{m+n}(0) \arrow[d, hook] \\
\mathrm{H}\mathfrak{S}_m \otimes \mathrm{H}\mathfrak{S}_n \arrow[r, hook] & \mathrm{H}\mathfrak{S}_{m+n}
\end{tikzcd} \tag{33}$$

Thus the map which sends a module to the characteristic of its restriction to $\mathrm{H}_n(0)$ is a coalgebra morphism. The isomorphism from $(\mathcal{K}, .)$ to $QSym$ is then obtained by Frobenius duality between induction of projective modules and restriction of simple modules. And the last case is obtained by applying the Cartan map C. $\square$

It is important to note that the algebra $(\mathcal{G}, .)$ is not the dual of the coalgebra $(\mathcal{K}, \Delta)$ because the dual of the restriction of projective modules is the so called *co-induction* of simple modules which is, in general, not the same as the induction for non self-injective algebras.

Finally the same process applied to the adjoint algebra which preserve symmetries would have given the following diagram

$$\begin{tikzcd}
(\mathrm{QSym}, .) & (\mathcal{G}, .) \arrow[l, hook', two heads, "\chi(S_I)\mapsto X_{I^c}"'] & (\mathcal{K}, .) \arrow[l, hook', two heads, "C"'] \arrow[r, hook, two heads, "\chi(P_I)\mapsto F_I"] & (\mathrm{QSym}, .) \\
(\mathrm{NCSF}, \Delta) & (\mathcal{G}, \Delta) \arrow[l, hook', two heads, "\chi(S_I)\mapsto R_I"] & (\mathcal{K}, \Delta) \arrow[l, hook', two heads, "C"] \arrow[r, hook, two heads, "\chi(P_I)\mapsto S^I"'] & (\mathrm{NCSF}, \Delta)
\end{tikzcd} \tag{34}$$

where $(X_I)_I$ is the dual basis of the elementary basis $(\Lambda_I)_I$ of NCSF. Thus we have a representation theoretical interpretation of many bases of NCSF and QSym.

3.4. Links with the affine Hecke algebra. Recall that, for any complex number q, the extended affine Hecke algebra $\hat{\mathrm{H}}_n(q)$ of type A_{n-1} is the $\mathbb{C}$-algebra generated by $(T_i)_{i=1\cdots n-1}$ together with an extra generator Ω verifying the defining relations of the Hecke algebra and the relation:

$$\Omega T_i = T_{i-1}\Omega \qquad \text{for } 1 \leq i \leq n. \tag{35}$$

The center of the affine Hecke algebra is isomorphic to the ring of symmetric polynomials in some variables $\xi_1, \dots, \xi_n$ and it can thus be specialized. Let us denote $\mathcal{H}_n(q)$ the specialization of the center $\hat{\mathrm{H}}_n(q)$ to the alphabet $1, q, \dots q^{n-1}$. That is

$$\mathcal{H}_n(q) := \hat{\mathrm{H}}_n(q)/\langle e_i(\xi_1, \dots, \xi_n) - e_i(1, q, \dots q^{n-1}) \mid i = 1 \dots n\rangle\,. \tag{36}$$

It is well known that the simple modules S_I of $\mathcal{H}_n(q)$ are indexed by compositions I and that their bases are indexed by descent classes of permutations. Thus one expects a strong link between $\mathrm{H}\mathfrak{S}_n$ and $\mathcal{H}_n(q)$. It comes out as follows. Let q be a generic complex number (i.e.: not 0 nor a root of the unity). Sending Ω to $\sigma_1\sigma_2\cdots\sigma_{n-1}$ and T_i to itself yields a surjective morphism from $\mathcal{H}_n(q)$ to $\mathrm{H}\mathfrak{S}_n$. Thus, the simple modules of $\mathcal{H}_n(q)$ are the simple modules of $\mathrm{H}\mathfrak{S}_n$ lifted back through this morphism. This also explains the link between the projective modules of $\mathrm{H}_n(0)$ and the simple modules of $\mathcal{H}_n(q)$, thanks to Proposition 2.

4. The algebra of non-decreasing functions

Definition 2. Let NDF_n be the set of *non-decreasing functions* from $\{1, \dots, n\}$ to itself. The composition and the neutral element id_n make NDF_n into a monoid. Its cardinal is $\binom{2n-1}{n-1}$, and we denote by $\mathbb{C}[\mathrm{NDF}_n]$ its monoid algebra.

The monoid $\mathrm{NDF}_n \times \mathrm{NDF}_m$ can be identified as the submonoid of NDF_{n+m} whose elements stabilize both $\{1, \dots, n\}$ and $\{n+1, \dots, n+m\}$. This makes $(\mathbb{C}[\mathrm{NDF}_n])_n$ into a tower of algebras.

One can take as generators for NDF_n and A_n the functions π_i et $\overline{\pi}_i$, such that $\pi_i(i+1) = i$, $\pi_i(j) = j$ for $j \neq i+1$, $\overline{\pi}_i(i) = i+1$, and $\pi_i(j) = j$ for $j \neq i$. The functions π_i are idempotents, and satisfy the braid relations, together with a new relation:

$$\pi_i^2 = \pi_i \qquad \text{and} \qquad \pi_{i+1}\pi_i\pi_{i+1} = \pi_i\pi_{i+1}\pi_i = \pi_{i+1}\pi_i\,. \tag{37}$$

This readily defines a morphism $\phi : \pi_{\mathrm{H}_n(0)} \mapsto \pi_{\mathbb{C}[\mathrm{NDF}_n]}$ of $\mathrm{H}_n(0)$ into $\mathbb{C}[\mathrm{NDF}_n]$. Its image is the monoid algebra of *non-decreasing parking functions* which will be discussed in Section 5. The same properties hold for the operators $\overline{\pi}_i$'s. Although this is not a priori obvious, it will turn out that the two morphisms $\phi : \pi_{\mathrm{H}_n(0)} \mapsto \pi_{\mathbb{C}[\mathrm{NDF}_n]}$ and $\overline{\phi} : \overline{\pi}_{\mathrm{H}_n(0)} \mapsto \overline{\pi}_{\mathbb{C}[\mathrm{NDF}_n]}$ are compatible, making $\mathbb{C}[\mathrm{NDF}_n]$ into a quotient of $\mathrm{H}\mathfrak{S}_n$.

4.1. Representation on exterior powers. We now want to construct a suitable representation of $\mathbb{C}[\mathrm{NDF}_n]$ where the existence of the epimorphism from $\mathrm{H}\mathfrak{S}_n$ onto $\mathbb{C}[\mathrm{NDF}_n]$, and the representation theory of $\mathbb{C}[\mathrm{NDF}_n]$ become clear.

The *natural representation* of $\mathbb{C}[\mathrm{NDF}_n]$ is obtained by taking the vector space $\mathbb{C}^n$ with canonical basis $e_1, \dots, e_n$, and letting a function f act on it by $e_i.f = e_{f(i)}$. For $n > 2$, this representation is a faithful representation of the monoid NDF_n but not of the algebra, as $\dim \mathbb{C}[\mathrm{NDF}_n] = \binom{2n-1}{n-1} \gg n^2$. However, since NDF_n is a monoid, the diagonal action on *exterior powers*

$$(x_1 \wedge \cdots \wedge x_k) \cdot f := (x_1 \cdot f) \wedge \cdots \wedge (x_k \cdot f) \tag{38}$$

still define an action. Taking the *exterior powers* $\bigwedge^k \mathbb{C}^n$ of the natural representation gives a new representation, whose basis $\{e_S := e_{s_1} \wedge \cdots \wedge e_{s_k}\}$ is indexed by subsets $S := \{s_1, \dots, s_k\}$ of $\{1, \dots, n\}$. The action of a function f in NDF_n is simply given by (note the absence of sign!):

$$e_S.f := \begin{cases} e_{f(S)} & \text{if } |f(S)| = |S|, \\ 0 & \text{otherwise.} \end{cases} \tag{39}$$

We call *representation of* $\mathbb{C}[\mathrm{NDF}_n]$ *on exterior powers* the representation of $\mathbb{C}[\mathrm{NDF}_n]$ on $\bigoplus_{k=1}^n \bigwedge^k \mathbb{C}^n$, which is of dimension $2^n - 1$ (it turns out that we do not need to include the component $\bigwedge^0 \mathbb{C}^n$ for our purposes).

Lemma 4.1. *The representation of* $\mathbb{C}[\mathrm{NDF}_n]$ *on* $\bigoplus_{k=1}^n \bigwedge^k \mathbb{C}^n \bigwedge \mathbb{C}^n$ *is faithful.*

We now want to realize the representation of $\mathbb{C}[\mathrm{NDF}_n]$ on the k-th exterior power as a representation of $\mathrm{H}\mathfrak{S}_n$. To this end, we use a variation on the standard construction of the Specht module $V_{k,1,\dots,1}$ of $\mathfrak{S}_n$ to make it a $\mathrm{H}\mathfrak{S}_n$-module. The trick is to use an appropriate quotient of $\mathbb{C}\mathfrak{S}_n$ to simulate the symmetries that we usually get by working with polynomials, while preserving the $\mathrm{H}\mathfrak{S}_n$-module structure. Namely, consider the following $\mathrm{H}\mathfrak{S}_n$-module:

$$P_n^k := P_{k,1,\dots,1}/\bigcup P_{k,1,\dots,1,2,1,\dots,1}. \tag{40}$$

An element in P_n^k is left-antisymmetric on the values $1,\dots,k-1$ and symmetric on the values $k+1,\dots,n-1$, the effect of the quotient being to identify two permutations which differ by a permutation of the values $\{k+1,\dots,n\}$. A basis of P_n^k indexed by subsets of size k of $\{1,\dots,n\}$ is obtained by taking for each such subset S the image in the quotient P_n^k of

$$e_S := \sum_{\sigma,\sigma(S)=\{1,\dots,k\},\sigma(i)<\sigma(j) \text{ for } i<j\notin S} (-1)^{\mathrm{sign}\sigma}\sigma\,. \tag{41}$$

It is straightforward to check that the actions of π_i and $\overline{\pi}_i$ of $\mathrm{H}\mathfrak{S}_n$ on e_S of P_k coincide with the actions of π_i and $\overline{\pi}_i$ of $\mathbb{C}[\mathrm{NDF}_n]$ on e_S of $\bigwedge^k \mathbb{C}^n$ (justifying a posteriori the identical notations). In the sequel, we identify the modules P_n^k and $\bigwedge^k \mathbb{C}^n$ of $\mathrm{H}\mathfrak{S}_n$ and $\mathbb{C}[\mathrm{NDF}_n]$, and we call *representation on exterior powers of* $\mathrm{H}\mathfrak{S}_n$ its representation on $\bigoplus_{k=1}^n \bigwedge^k \mathbb{C}^n$. Using Lemma 4.1 we are in position to state the following

PROPOSITION 6. *$\mathbb{C}[\mathrm{NDF}_n]$ is the quotient of $\mathrm{H}\mathfrak{S}_n$ obtained by considering its representation on exterior powers. The restriction of this representation of $\mathrm{H}\mathfrak{S}_n$ to $\mathbb{C}[\mathfrak{S}_n]$, $\mathrm{H}_n(0)$, and $\mathrm{H}_n(-1)$ yield respectively the usual representation of $\mathfrak{S}_n$ on exterior powers, the algebra of non-decreasing parking functions (see Section 5), and the Temperley-Lieb algebra.*

4.2. Representation theory.

4.2.1. *Projective modules, simple modules, and Cartan's invariant matrix.* Let δ be the usual homology border map:

$$\delta : \begin{cases} P_n^k & \to P_n^{k-1} \\ S := \{s_1,\dots,s_k\} & \mapsto \sum_{i\in\{1,\dots,k\}} (-1)^{k-i} S\backslash\{s_i\} \end{cases}. \tag{42}$$

This map is naturally a morphism of $\mathbb{C}[\mathrm{NDF}_n]$-module. For each k in $1,\dots,n$, let $S_k := P_k/\ker\delta$. It turns out that together with the identity, δ is essentially the only $\mathbb{C}[\mathrm{NDF}_n]$-morphism. We are now in position to describe the projective and simple modules, as well as the Cartan matrix of $\mathbb{C}[\mathrm{NDF}_n]$.

PROPOSITION 7. *The modules $(P_n^k)_{k=1,\dots,n}$ form a complete set of representatives of the indecomposable projective modules of $\mathbb{C}[\mathrm{NDF}_n]$.*

The modules $(S_n^k)_{k=1,\dots,n}$ form a complete set of representatives of the simple modules of $\mathbb{C}[\mathrm{NDF}_n]$.

Let k and l be two integers in $\{1,\dots,n\}$. Then,

$$\dim Hom(P_n^k, P_n^l) = \begin{cases} 1 & \textit{if } l\in\{k,k-1\}, \\ 0 & \textit{otherwise.} \end{cases} \tag{43}$$

The proof relies essentially on the following lemma:

LEMMA 4.2. *There exists a minimal decomposition of the identity of $\mathbb{C}[\mathrm{NDF}_n]$ into 2^n-1 orthogonal idempotents. In particular, the representation on exterior powers is the smallest faithful representation of $\mathbb{C}[\mathrm{NDF}_n]$.*

4.2.2. *Induction, restriction, and Grothendieck groups.*

PROPOSITION 8. *The restriction and induction of indecomposable projective modules and simple modules are described by:*

$$P_{n_1+n_2}^k \downarrow_{\mathbb{C}[\mathrm{NDF}_{n_1}]\otimes\mathbb{C}[\mathrm{NDF}_{n_2}]}^{\mathbb{C}[\mathrm{NDF}_{n_1+n_2}]} \approx \bigoplus_{\substack{n_1+n_2=n \\ k_1+k_2=k \\ 1\le k_i\le n_i \textit{ or } k_i=n_i=0}} P_{n_1}^{k_1}\otimes P_{n_2}^{k_2} \tag{44}$$

$$P_{n_1}^{k_1}\otimes P_{n_2}^{k_2} \uparrow_{\mathbb{C}[\mathrm{NDF}_{n_1}]\otimes\mathbb{C}[\mathrm{NDF}_{n_2}]}^{\mathbb{C}[\mathrm{NDF}_{n_1+n_2}]} \approx P_{n_1+n_2}^{k_1+k_2} \oplus P_{n_1+n_2}^{k_1+k_2-1} \tag{45}$$

$$S^{k}_{n_1+n_2} \downarrow^{\mathbb{C}[\mathrm{NDF}_{n_1+n_2}]}_{\mathbb{C}[\mathrm{NDF}_{n_1}]\otimes\mathbb{C}[\mathrm{NDF}_{n_2}]} = \bigoplus_{\substack{n_1+n_2=n\\ k_1+k_2\in\{k,k+1\}\\ 1\leq k_i\leq n_i \ or\ k_i=n_i=0}} S^{k_1}_{n_1}\otimes S^{k_2}_{n_2} \tag{46}$$

$$S^{k_1}_{n_1}\otimes S^{k_2}_{n_2} \uparrow^{\mathbb{C}[\mathrm{NDF}_{n_1+n_2}]}_{\mathbb{C}[\mathrm{NDF}_{n_1}]\otimes\mathbb{C}[\mathrm{NDF}_{n_2}]} \approx S^{k_1+k_2}_{n_1+n_2} \tag{47}$$

Those rules yield structures of commutative algebras and cocommutative coalgebras on $\mathcal{G}$ and $\mathcal{K}$ which can be realized as quotients or sub(co)algebras of Sym, QSym, and NCSF. However, we do not get Hopf algebras, because the structures of algebras and coalgebras are not compatible (compute for example $\Delta(\chi(P^1_1)\chi(P^1_1))$ in the two ways, and check that the coefficients of $\chi(P^1_1)\otimes\chi(P^1_1)$ differ).

5. The algebra of non-decreasing parking functions

Definition 3. A *nondecreasing parking function* of size n is a nondecreasing function f from $\{1,2,\dots n\}$ to $\{1,2,\dots n\}$ such that $f(i)\leq i$, for all $i\leq n$.

The composition of maps and the neutral element id_n make the set of nondecreasing parking function of size n into a monoid denoted NDPF_n.

It is well known that the nondecreasing parking functions are counted by the Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$. It is also clear that NDPF_n is the sub-monoid of NDF_n generated by the π_i's.

5.1. Simple modules.

The goal of the sequel is to study the representation theory of NDPF_n, or equivalently of its algebra $\mathbb{C}[\mathrm{NDPF}_n]$. The following remark allows us to deduce the representations of $\mathbb{C}[\mathrm{NDPF}_n]$ from the representations of $\mathrm{H}_n(0)$.

Proposition 9. *The kernel of the algebra epi-morphism $\phi:\mathrm{H}_n(0)\to\mathbb{C}[\mathrm{NDPF}_n]$ defined by $\phi(\pi_i)=\pi_i$ is a sub-ideal of the radical of $\mathrm{H}_n(0)$.*

Proof. It is well known (see [**12**]) that the quotient of $\mathrm{H}_n(0)$ by its radical is a commutative algebra. Consequently, $\pi_i\pi_{i+1}\pi_i - \pi_i\pi_{i+1} = [\pi_i\pi_{i+1},\pi_i]$ belongs to the radical of $\mathrm{H}_n(0)$. □

As a consequence, taking the quotient by their respective radical shows that the projection ϕ is an isomorphism from $\mathbb{C}[\mathrm{NDPF}_n]/\mathrm{rad}(\mathbb{C}[\mathrm{NDPF}_n])$ to $\mathrm{H}_n(0)/\mathrm{rad}(\mathrm{H}_n(0))$. Moreover $\mathbb{C}[\mathrm{NDPF}_n]/\mathrm{rad}(\mathbb{C}[\mathrm{NDPF}_n])$ is isomorphic to the commutative algebra generated by the π_i such that $\pi_i^2=\pi_i$. As a consequence, $\mathrm{H}_n(0)$ and $\mathrm{H}\mathfrak{S}_n$ share, roughly speaking, the same simple modules:

Corollary 1. *There are 2^{n-1} simple $\mathbb{C}[\mathrm{NDPF}_n]$-modules S_I, and they are all one dimensional. The structure of the module S_I, generated by η_I, is given by*

$$\begin{cases} \eta_I\cdot\pi_i = 0 & \text{if } i\in\mathrm{Des}(I),\\ \eta_I\cdot\pi_i=\eta_I & \text{otherwise.}\end{cases} \tag{48}$$

5.2. Projective modules.

The projective modules of NDPF_n can be deduced from the ones of NDF_n.

Theorem 5.1. *Let I be a composition of n, and $S:=\mathrm{Des}(I)=\{s_1,\dots,s_k\}$ be its associated set. Then, the principal sub-module*

$$P_I := (e_1\wedge e_{s_1+1}\wedge\cdots\wedge e_{s_1+1})\cdot\mathbb{C}[\mathrm{NDPF}_n] \quad\subset\quad \bigwedge^{k+1}\mathbb{C}^n \tag{49}$$

is an indecomposable projective module. Moreover, the set $(P_I)_{I\vDash n}$ is a complete set of representatives of indecomposable projective modules of $\mathbb{C}[\mathrm{NDPF}_n]$.

This suggests an alternative description of the algebra $\mathbb{C}[\mathrm{NDPF}_n]$. Let $G_{n,k}$ be the lattice of subsets of $\{1,\dots,n\}$ of size k for the *product order* defined as follows. Let $S:=\{s_1<s_2<\cdots<s_k\}$ and $T:=\{t_1<t_2<\cdots<t_k\}$ be two subsets. Then,

$$S\leq_G T \qquad \text{if and only if} \qquad s_i\leq t_i,\ \text{for } i=1,\dots,k. \tag{50}$$

One easily sees that $S\leq_G T$ if and only if there exists a nondecreasing parking function f such that $e_S=e_T\cdot f$. This lattice appears as the Bruhat order associated to the Grassman manifold G^n_k of k-dimensional subspaces in $\mathbb{C}^n$.

Theorem 5.2. *There is a natural algebra isomorphism*

$$\mathbb{C}[\mathrm{NDPF}_n] \;\approx\; \bigoplus_{k=0}^{n-1} \mathbb{C}[G_{n-1,k}]\ . \tag{51}$$

In particular the Cartan map $C : \mathcal{K} \rightarrow \mathcal{G}$ is given by the lattice $\leq_G$:

$$C(P_I) \;=\; \sum_{J,\ \mathrm{Des}(J) \leq_G \mathrm{Des}(I)} S_J \tag{52}$$

On the other hand, due to the commutative diagram

$$\begin{tikzcd} \mathrm{H}_m(0) \otimes \mathrm{H}_n(0) \arrow[r, hook] \arrow[d, two heads] & \mathrm{H}_{m+n}(0) \arrow[d, two heads] \\ \mathrm{NDPF}_m \otimes \mathrm{NDPF}_n \arrow[r, hook] & \mathrm{NDPF}_{m+n} \end{tikzcd} \tag{53}$$

it is clear that the restriction of simple modules and the induction of indecomposable projective modules follow the same rule as for $\mathrm{H}_n(0)$. The induction of simple modules can be deduced via the Cartan map, giving rise to a new basis G_I of NCSF. It remains finally to compute the restrictions of indecomposable projective modules. It can be obtained by a not yet completely explicit algorithm. All of this is summarized by the following diagram:

$$\begin{tikzcd} (\mathrm{NCSF}, .) & (\mathcal{G}, .) \arrow[l, hook, two heads, "\chi(S_I) \mapsto G_I"'] & (\mathcal{K}, .) \arrow[l, hook, two heads, "C"'] \arrow[r, hook, two heads, "\chi(P_I) \mapsto R_I"] & (\mathrm{NCSF}, .) \\ (\mathrm{QSym}, \Delta) & (\mathcal{G}, \Delta) \arrow[l, hook, two heads, "\chi(S_I) \mapsto F_I"] & (\mathcal{K}, \Delta) \arrow[l, hook, two heads, "C"] \arrow[r, hook, two heads, "\chi(P_I) \mapsto ???"'] & ??? \end{tikzcd} \tag{54}$$

References

[1] S. Ariki, *On the decomposition numbers of the Hecke algebra of $G(m, 1, n)$*, J. Math. Kyoto Univ. **36** (1996), 789–808.

[2] S. Ariki and K. Koike, *A Hecke Algebra of $\mathbb{Z}/r\mathbb{Z} \wr \mathfrak{S}_n$ and construction of its irreducible representations*, Adv. Math. **106** (1994), 216–243.

[3] N. Bergeron, F. Hivert, and J.-Y. Thibon, *The peak algebra and the Hecke-Clifford algebras at $q = 0$*, J. Comb. Theory A **117** (2004), 1–19.

[4] C. W. Curtis and I. Reiner, *Methods of Representation Theory*, Wiley Classics Library, New York, 1981.

[5] I. M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, and J.-Y. Thibon, *Noncommutative symmetric functions*, Adv. in Math. **112** (1995), 218–348.

[6] I. Gessel, *Multipartite P-partitions and inner products of skew Schur functions*, [in "Combinatorics and algebra", C. Greene, Ed.], Contemporary Mathematics, **34** (1984), 289–301.

[7] F. Hivert, J.-C. Novelli, and J.-Y. Thibon, *Yang-Baxter bases of* 0*-Hecke algebras and representation theory of* 0*-Ariki-Koike-Shoji algebras*, Advances in Maths, to appear, `arxiv:math.CO/0506546`.

[8] F. Hivert and N. Thiéry, *Mupad-combinat, an open-source package for research in algebraic combinatorics,* Séminaire Lotharingien de Combinatoire, **51** (2003), 70 p. electronic. See also `http://mupad-combinat.sourceforge.net/`

[9] D. Krob and J.-Y. Thibon, *Noncommutative symmetric functions IV: Quantum linear groups and Hecke algebras at $q = 0$*, J. Alg. Comb. **6** (1997), 339–376.

[10] A. Lascoux, B. Leclerc, and J.-Y. Thibon, *Hecke algebras at roots of unity and crystal bases of quantum affine algebras*, Communications in Mathematical Physics **181** (1996), 205–263.

[11] I.G. Macdonald, *Symmetric functions and Hall polynomials*, 2nd ed., Oxford University Press, 1995.

[12] P. N. Norton, 0*-Hecke algebras*, J. Austral. Math. Soc., A, **27** (1979) 337–357.

[13] G. Olshanski, *Quantized universal enveloping superalgebra of type Q and a super-extension of the Hecke algebra*, Lett. Math. Phys. **24** (1992), 93–102.

[14] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://www.research.att.com/~njas/sequences/`.

[15] J. Stembridge, *Enriched P-partitions*, Trans. Amer. Math. Soc. **349** (1997), 763–788.

[16] A. V. Zelevinsky, *Representations of finite classical groups. A Hopf algebra approach*, Springer LNM **869**, Springer-Verlag, Heidelberg, 1981.

LIFAR, University of Rouen, Avenue de l'Université, Technopôle du Madrillet, 76801 Saint Etienne du Rouvray Cedex, FRANCE
E-mail address: `florent.hivert@univ-rouen.fr`

Laboratoire de Mathématiques, Université Paris Sud, Bât 425, 91405 Orsay Cedex, FRANCE
E-mail address: `nthiery@users.sf.net`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

On the Number of Factorizations of a Full Cycle

John Irving

Abstract. We give a new expression for the number of factorizations of a full cycle into an ordered product of permutations of specified cycle types. This is done through purely algebraic means, extending recent work of Biane [*Nombre de factorisations d'un grand cycle*, Sém. Lothar. de Combinatoire **51** (2004)]. We deduce from our result a remarkable formula of Poulalhon and Schaeffer [*Factorizations of large cycles in the symmetric group*, Discrete Math. **254** (2002), 433–458] that was previously derived through an intricate combinatorial argument.

Résumé. Nous proposons une nouvelle formule pour le nombre de factorisations d'un grand cycle en un produit ordonné de permutations de types cycliques donnés. Nous utilisons des arguments purement algébriques, étendant un travail récent de Biane [*Nombre de factorisations d'un grand cycle.*, Sém. Lothar. de Combinatoire **51** (2004)]. Nous déduisons de notre résultat une formule remarquable de Poulalhon et Schaeffer [*Factorizations of large cycles in the symmetric group*, Discrete Math. **254** (2002), 433–458] obtenue précédemment à l'aide d'arguments combinatoires complexes.

1. Notation

Our notation is generally consistent with Macdonald [**5**]. We write $\lambda \vdash n$ (or $|\lambda| = n$) and $\ell(\lambda) = k$ to indicate that λ is a partition of n into k parts; that is, $\lambda = (\lambda_1, \ldots, \lambda_k)$ with $\lambda_1 \geq \cdots \geq \lambda_k \geq 1$ and $\lambda_1 + \ldots + \lambda_k = n$. If λ has exactly m_i parts equal to i then we write $\lambda = [1^{m_1} 2^{m_2} \cdots]$, suppressing terms with $m_i = 0$. We also define $z_\lambda = \prod_i i^{m_i} m_i!$ and $\mathrm{Aut}(\lambda) = \prod_i m_i!$. A *hook* is a partition of the form $[1^b,\, a+1]$ with $a, b \geq 0$. We use Frobenius notation for hooks, writing $(a|b)$ in place of $[1^b,\, a+1]$.

The conjugacy class of the symmetric group $\mathfrak{S}_n$ consisting of all $n!/z_\lambda$ permutations of cycle type $\lambda \vdash n$ will be denoted by $\mathcal{C}_\lambda$. The irreducible characters χ^λ of $\mathfrak{S}_n$ are naturally indexed by partitions λ of n, and we use the usual notation χ^λ_μ for the common value of χ^λ at any element of $\mathcal{C}_\mu$. We write f^λ for the degree $\chi^\lambda_{[1^n]}$ of χ^λ.

For vectors $\mathbf{j} = (j_1, \ldots, j_m)$ and $\mathbf{x} = (x_1, \ldots, x_m)$ we use the abbreviations $\mathbf{j}! = j_1! \cdots j_m!$ and $\mathbf{x}^{\mathbf{j}} = x_1^{j_1} \cdots x_m^{j_m}$. Finally, if $\alpha \in \mathbb{Q}$ and $f \in \mathbb{Q}[[\mathbf{x}]]$ is a formal power series, then we write $[\alpha \mathbf{x}^{\mathbf{j}}]\, f(\mathbf{x})$ for the coefficient of the monomial $\alpha \mathbf{x}^{\mathbf{j}}$ in $f(\mathbf{x})$.

2. Factorizations of Full Cycles

Given $\lambda, \alpha_1, \ldots, \alpha_m \vdash n$, let $c^\lambda_{\alpha_1,\ldots,\alpha_m}$ be the number of factorizations in $\mathfrak{S}_n$ of a given permutation $\pi \in \mathcal{C}_\lambda$ as an ordered product $\pi = \sigma_1 \ldots \sigma_m$, with $\sigma_i \in \mathcal{C}_{\alpha_i}$ for all i. The problem of evaluating $c^\lambda_{\alpha_1,\ldots,\alpha_m}$ for various λ and α_i has attracted a good deal of attention and is linked to various questions in algebra, geometry, and physics. For details on the history of this problem and its connections to other areas of

2000 *Mathematics Subject Classification.* Primary 05A05; Secondary 20B30.

Key words and phrases. permutations, factorizations, symmetric group, enumeration, combinatorics.

Thanks go to David Jackson for financial support during the preparation of this manuscript, and Ian Goulden for helpful discussions along the way.

mathematics, we direct the reader to [**4**] and the references therein. Here we focus on the particularly well-studied case $\lambda = (n)$, which corresponds to counting factorizations of the full cycle $(1\,2\,\cdots\,n) \in \mathfrak{S}_n$ into factors of specified cycle types.

While it is straightforward to express $c^{(n)}_{\alpha_1,\ldots,\alpha_m}$ as a character sum (see (3.1) below), the appearance of alternating signs in this sum — and resulting cancellations — preclude asymptotic analysis. Goupil and Schaeffer [**4**, FPSAC'98] overcame this difficulty in the case $m = 2$ by interpreting certain characters combinatorially (*viz.* the Murnaghan-Nakayama rule) and employing a sequence of bijections in which a sign-reversing involution accounts for cancellations. This leads to an expression for $c^{(n)}_{\alpha,\beta}$ as a sum of *positive* terms, which in turn permits nontrivial asymptotics. Poulalhon and Schaeffer [**6**] later extended this argument to arrive at a similar formula for $c^{(n)}_{\alpha_1,\ldots,\alpha_m}$.

Biane [**1**] has recently given a remarkably succinct algebraic derivation of Goupil and Schaeffer's formula for $c^{(n)}_{\alpha,\beta}$. Our purpose here is to extend his method to give a new expression for $c^{(n)}_{\alpha_1,\ldots,\alpha_m}$ as a sum of positive contributions. In particular, if for $\gamma = (\gamma_1, \gamma_2, \ldots) \vdash n$ we define the polynomial $R_\gamma(x, y)$ and the *nonnegative* constants $r^\gamma_{j,k}$ by

$$R_\gamma(x,y) := \frac{1}{2y}\prod_{i\geq 1}\left((x+y)^{\gamma_i} - (x-y)^{\gamma_i}\right) = \sum_{j+k=n-1} r^\gamma_{j,k} x^j y^k, \tag{2.1}$$

then our main result is the following:

THEOREM 2.1. *Let $\alpha_1, \ldots, \alpha_m \vdash n$ and, for $\lambda = [1^{m_1}2^{m_2}3^{m_3}\cdots]$, let $2\lambda - 1 = [1^{m_1}3^{m_2}5^{m_3}\cdots]$. Set $\mathbf{x} = (x_1, \ldots, x_m)$ and let $e_\lambda(\mathbf{x})$ denote the elementary symmetric function in $x_1, \ldots, x_m$ indexed by λ. Then*

$$c^{(n)}_{\alpha_1,\ldots,\alpha_m} = \frac{n^{m-1}}{2^{(n-1)(m-1)}\prod_i z_{\alpha_i}} \sum_{\mathbf{j}+\mathbf{k}=\mathbf{n}-\mathbf{1}} [\mathbf{x}^{\mathbf{j}}] \sum_{\ell(\lambda)=n-1} \frac{e_{2\lambda-1}(\mathbf{x})}{\mathrm{Aut}(\lambda)} \cdot \prod_{i=1}^m j_i!\, k_i!\, r^{\alpha_i}_{j_i,k_i},$$

where the outer sum extends over all vectors $\mathbf{j} = (j_1, \ldots, j_m)$ and $\mathbf{k} = (k_1, \ldots, k_m)$ of nonnegative integers such that $j_i + k_i = n - 1$ for all i, and the inner sum over all partitions λ with $n - 1$ parts.

A proof of Theorem 2.1 is given in the next section. In §4, we use this result to deduce Poulalhon and Schaeffer's formula for $c^{(n)}_{\alpha_1,\ldots,\alpha_m}$ (listed here as Theorem 4.1), thereby giving a purely algebraic derivation that avoids the detailed combinatorial constructions in [**6**].

3. Proof of the Main Result

It is well known that the class sums $\mathsf{K}_\lambda = \sum_{\sigma\in\mathcal{C}_\lambda}\sigma$ (for $\lambda \vdash n$) form a basis of the centre of the group algebra $\mathbb{C}\mathfrak{S}_n$. Indeed, the linearization relations $\mathsf{K}_{\alpha_1}\cdots\mathsf{K}_{\alpha_m} = \sum_{\lambda\vdash n} c^\lambda_{\alpha_1,\ldots,\alpha_m}\mathsf{K}_\lambda$ identify the constants $c^\lambda_{\alpha_1,\ldots,\alpha_m}$ as the connection coefficients of $\mathbb{C}\mathfrak{S}_n$. By using character theory to express K_λ in terms of central idempotents of $\mathbb{C}\mathfrak{S}_n$ (see [**7**], Problem 7.67b) one finds that

$$c^\lambda_{\alpha_1,\ldots,\alpha_m} = \frac{n!^{m-1}}{z_{\alpha_1}\cdots z_{\alpha_m}} \sum_{\beta\vdash n} \frac{\chi^\beta_{\alpha_1}\cdots\chi^\beta_{\alpha_m}}{(f^\beta)^{m-1}}\chi^\beta_\lambda.$$

This sum is generally intractable but simplifies considerably in the case $\lambda = (n)$, since there χ^β_λ vanishes when β is not a hook; in particular, the Murnaghan-Nakayama rule [**7**] implies $\chi^\beta_{(n)} = (-1)^b$ if $\beta = (a|b)$, while $\chi^\beta_{(n)} = 0$ otherwise. Moreover, the hook-length formula gives $f^{(a|b)} = \binom{a+b}{b}$, so

$$c^{(n)}_{\alpha_1,\ldots,\alpha_m} = \frac{n^{m-1}}{z_{\alpha_1}\cdots z_{\alpha_m}} \sum_{a+b=n-1} (a!\,b!)^{m-1}\chi^{(a|b)}_{\alpha_1}\cdots\chi^{(a|b)}_{\alpha_m}(-1)^b. \tag{3.1}$$

Let μ be the measure on $\mathbb{C}$ defined by the density $d\mu(z) = \frac{1}{\pi}e^{-|z|^2}dz$, where dz is the standard Lebesgue density (*i.e.* $dz = ds\,dt$ for $z = s + t\sqrt{-1}$). Following Biane [**1**], we shall make use of the formula

$$\int_{\mathbb{C}} z^j\bar{z}^k d\mu(z) = j!\,\delta_{jk}, \tag{3.2}$$

which is easily verified by changing to polar form.

Proof of Theorem 2.1: For $\gamma \vdash n$, let $F_\gamma(u,v) = \sum \chi_\gamma^{(a|b)} u^a v^b$ be the generating series for hook characters, where the sum extends over all pairs (a,b) of nonnegative integers with $a+b=n-1$. Then

$$\frac{1}{(n-1)!}(u_1\cdots u_m - v_1\cdots v_m)^{n-1}\prod_{i=1}^m F_{\alpha_i}(\bar u_i,\bar v_i)$$
$$= \sum_{a+b=n-1}\frac{u_1^a\cdots u_m^a\cdot v_1^b\cdots v_m^b}{a!\,b!}(-1)^b\prod_{i=1}^m\sum_{a_i+b_i=n-1}\chi_{\alpha_i}^{(a_i|b_i)}\bar u_i^{a_i}\bar v_i^{b_i}.$$

Consider the effect of integrating the RHS with respect to $d\mu(\mathbf{u},\mathbf{v}) := \prod_{i=1}^m d\mu(u_i)d\mu(v_i)$. Using (3.2), note that all monomials $\frac{(-1)^b}{a!b!}\prod_i \chi_{a_i}^{(a_i|b_i)} u_i^a \bar u_i^{a_i} v_i^b \bar v_i^{b_i}$ vanish except those with $a_i = a$ and $b_i = b$ for all i, and each monomial of this special form is replaced by $(-1)^b(a!\,b!)^{m-1}\prod_i \chi_{\alpha_i}^{(a|b)}$. Thus we obtain

$$\int_{\mathbb{C}^{2m}}(u_1\cdots u_m - v_1\cdots v_m)^{n-1}\prod_{i=1}^m F_{\alpha_i}(\bar u_i,\bar v_i)\,d\mu(\mathbf{u},\mathbf{v})$$
$$= (n-1)!\sum_{a+b=n-1}(a!\,b!)^{m-1}\chi_{\alpha_1}^{(a|b)}\cdots\chi_{\alpha_m}^{(a|b)}(-1)^b.$$

Let I be the integral on the LHS, and change variables by letting $u_i = (y_i+x_i)/\sqrt{2}$, $v_i = (y_i - x_i)/\sqrt{2}$. As an immediate consequence of the Murnaghan-Nakayama rule we have

$$F_\gamma(u,v) = \frac{1}{u+v}\prod_{i\geq 1}(u^{\gamma_i} - (-v)^{\gamma_i})$$

for a partition $\gamma = (\gamma_1,\gamma_2,\ldots)$. Thus (2.1) gives $F_\gamma(y+x,y-x) = R_\gamma(x,y)$, and since F_{α_i} is homogeneous of degree $n-1$ the change of variables yields $F_{\alpha_i}(\bar u_i,\bar v_i) = 2^{-(n-1)/2}R_{\alpha_i}(\bar x_i,\bar y_i)$ for all i. Furthermore, it is easy to check that $d\mu(\mathbf{u},\mathbf{v}) = d\mu(\mathbf{x},\mathbf{y})$ and

$$u_1\cdots u_m - v_1\cdots v_m = \frac{1}{\sqrt{2^m}}\left(\prod_{i=1}^m(y_i+x_i) - \prod_{i=1}^m(y_i-x_i)\right) = \frac{2y_1\cdots y_m}{\sqrt{2^m}}\sum_{s\geq 1}e_{2s-1}(\mathbf{x}/\mathbf{y}),$$

where $\mathbf{x}/\mathbf{y} = (\frac{x_1}{y_1},\ldots,\frac{x_m}{y_m})$. Thus, with the aid of (3.2), we get

$$\begin{aligned}
I &= \frac{1}{2^{(n-1)(m-1)}}\int_{\mathbb{C}^{2m}}\left(y_1\cdots y_m\sum_{s\geq 1}e_{2s-1}(\mathbf{x}/\mathbf{y})\right)^{n-1}\prod_{i=1}^m R_{\alpha_i}(\bar x_i,\bar y_i)\,d\mu(\mathbf{x},\mathbf{y})\\
&= \frac{1}{2^{(n-1)(m-1)}}\sum_{\mathbf{j},\mathbf{k}}\mathbf{j}!\,\mathbf{k}!\,[\mathbf{x}^{\mathbf{j}}\mathbf{y}^{\mathbf{k}}]\left(y_1\cdots y_m\sum_{s\geq 1}e_{2s-1}(\mathbf{x}/\mathbf{y})\right)^{n-1}\cdot[\bar{\mathbf{x}}^{\mathbf{j}}\bar{\mathbf{y}}^{\mathbf{k}}]\prod_{i=1}^m R_{\alpha_i}(\bar x_i,\bar y_i)\\
&= \frac{1}{2^{(n-1)(m-1)}}\sum_{\mathbf{j}+\mathbf{k}=\mathbf{n-1}}\mathbf{j}!\,\mathbf{k}!\,[\mathbf{x}^{\mathbf{j}}]\left(\sum_{s\geq 1}e_{2s-1}(\mathbf{x})\right)^{n-1}\prod_{i=1}^m r_{j_i,k_i}^{\alpha_i}\\
&= \frac{(n-1)!}{2^{(n-1)(m-1)}}\sum_{\mathbf{j}+\mathbf{k}=\mathbf{n-1}}\mathbf{j}!\,\mathbf{k}!\,[\mathbf{x}^{\mathbf{j}}]\sum_{\ell(\lambda)=n-1}\frac{e_{2\lambda-1}(\mathbf{x})}{\mathrm{Aut}(\lambda)}\prod_{i=1}^m r_{j_i,k_i}^{\alpha_i}.
\end{aligned}$$

The result now follows from (3.1). $\square$

4. Recovery of Poulalhon & Schaeffer's Formula

We require some extra notation to state the Poulalhon-Schaeffer formula for $c_{\alpha_1,\ldots,\alpha_m}^{(n)}$. First, we define symmetric polynomials $S_p(x_1,\ldots,x_l)$ by setting $S_0(x_1,\ldots,x_l) = 1$ and

$$S_p(x_1,\ldots,x_l) = \sum_{p_1+\cdots+p_l=p}\prod_{i=1}^l\frac{1}{x_i}\binom{x_i}{2p_i+1}$$

for $p > 0$. Note that these have the simple generating series

$$\sum_{p\geq 0} S_p(x_1,\ldots,x_l)t^{2p} = \prod_{i=1}^{l} \frac{(1+t)^{x_i}-(1-t)^{x_i}}{2x_i t}, \tag{4.1}$$

which is obviously closely related to our series $R_\gamma(x,y)$ (see (2.1)). We also introduce an operator $\mathfrak{D}$ on $\mathbb{Q}[[x_1,\ldots,x_m]]$ defined as follows: For each i and all $j\geq 0$ set $\mathfrak{D}(x_i^j) = x_i(x_i-1)\cdots(x_i-j+1)$, and extend the action of $\mathfrak{D}$ multiplicatively to monomials $x_1^{j_1}\cdots x_m^{j_m}$ and then linearly to all of $\mathbb{Q}[[x_1,\ldots,x_m]]$. Finally, we define polynomials $P_a^b(x_1,\ldots,x_m)$ by setting $P_0^b(x_1,\ldots,x_m)=1$ for all $b\geq 1$ and letting

$$P_a^b(x_1,\ldots,x_m) \;= \sum_{\substack{\lambda\vdash a\\ \ell(\lambda)\leq b}} \mathfrak{D}\Big(\frac{e_{2\lambda+1}(\mathbf{x})}{\mathrm{Aut}(\lambda)}\Big) \tag{4.2}$$

for $a,b\geq 1$, where $2\lambda+1 = [3^{m_1}5^{m_2}\cdots]$ when $\lambda=[1^{m_1}2^{m_2}\cdots]$. Then the main result of [**6**] is the following intriguing formula for $c^{(n)}_{\alpha_1,\ldots,\alpha_m}$.

Theorem 4.1 (Poulalhon-Schaeffer). *Let $\alpha_1,\ldots,\alpha_m\vdash n$ and set $r_i = n-\ell(\alpha_i)$ for all i. Let $g = \frac{1}{2}(\sum_i r_i - n+1)$. If g is a nonnegative integer, then*

$$c^{(n)}_{\alpha_1,\ldots,\alpha_m} = \frac{n^{m-1}}{2^{2g}\prod_i \mathrm{Aut}(\alpha_i)} \sum P_q^{n-1}(\mathbf{r}-\mathbf{2p}) \prod_{i=1}^{m} (\ell(\alpha_i)+2p_i-1)!\, S_{p_i}(\alpha_i),$$

where $\mathbf{r}-\mathbf{2p} = (r_1-2p_1,\ldots,r_m-2p_m)$ and the sum extends over all tuples $(q,p_1,\ldots,p_m)$ of nonnegative integers with $q+p_1+\cdots+p_m = g$.

Before proceeding to deduce this result from Theorem 2.1, we pause for a few remarks. First, the integer g identified in Theorem 4.1 is called the *genus* of the associated factorizations of $(1\,2\,\cdots\,n)$, and it has well-understood geometric meaning; see [**2**], for example. The primary benefit of the Poulalhon-Schaeffer formula (over Theorem 2.1) is that the dependence on genus is explicit. For instance, when $g=0$ it is immediately clear that Theorem 4.1 reduces to the very simple

$$c^{(n)}_{\alpha_1,\ldots,\alpha_m} = n^{m-1}\prod_{i=1}^{m}\frac{(\ell(\alpha_i)-1)!}{\mathrm{Aut}(\alpha_i)}.$$

The $c^{(n)}_{\alpha_1,\ldots,\alpha_m}$ in this case are known as *top connection coefficients*, and the above formula was originally given by Goulden and Jackson [**3**].

Secondly, we note that Poulalhon and Schaeffer actually define $P_a(\mathbf{x}) = \sum_{\lambda\vdash a}\mathfrak{D}(e_{2\lambda+1}(\mathbf{x})/\mathrm{Aut}(\lambda))$, and ignore the condition $\ell(\lambda)\leq b$ in our definition of P_a^b. However, replacing P_q^{n-1} with P_q in Theorem 4.1 has nil effect, since for $\mathfrak{D}(e_{2\lambda+1}(\mathbf{x}))|_{\mathbf{x}=\mathbf{r}-\mathbf{2p}}$ to be nonzero some monomial in $e_{2\lambda+1}(\mathbf{x})$ must be of the form $x_1^{j_1}\cdots x_m^{j_m}$ with $j_i\leq r_i-2p_i$ for all i. This implies $|2\lambda+1| = \sum_i j_i \leq \sum_i (r_i-2p_i) = 2g+n-1-\sum_i 2p_i$, while the conditions $\lambda\vdash q$ and $q+\sum_i p_i = g$ give $|2\lambda+1| = 2q+\ell(\lambda) = 2g-\sum_i 2p_i+\ell(\lambda)$. Thus we require $\ell(\lambda)\leq n-1$ for nonzero contributions to $P_q(\mathbf{r}-\mathbf{2p})$.

Lemma 4.2. *Let $s,t_1,\ldots,t_m$ be nonnegative integers. Set $\mathbf{x}=(x_1,\ldots,x_m)$ and $\mathbf{t}=(t_1,\ldots,t_m)$, and let $f(\mathbf{x})$ be a homogeneous polynomial of total degree $t_1+\cdots+t_m-s$. Then*

$$\left[\frac{\mathbf{x}^{\mathbf{t}}}{\mathbf{t}!}\right](x_1+\cdots+x_m)^s f(\mathbf{x}) = s!\,\mathfrak{D}\big(f(\mathbf{x})\big)\big|_{x_1=t_1,\ldots,x_m=t_m}.$$

Proof. Consider the case where $f(\mathbf{x}) = x_1^{j_1}\cdots x_m^{j_m}$ with $\sum_i j_i = \sum_i t_i - s$. Here

$$\begin{aligned}
[\mathbf{x}^{\mathbf{t}}](x_1+\cdots+x_m)^s f(\mathbf{x}) &= [\mathbf{x}^{\mathbf{t}}]\sum_{i_1+\cdots+i_m=s}\frac{s!}{i_1!\cdots i_m!}x_1^{i_1+j_1}\cdots x_m^{i_m+j_m}\\
&= \begin{cases} \dfrac{s!}{\mathbf{t}!}\displaystyle\prod_{i=1}^{m}\frac{t_i!}{(t_i-j_i)!} & \text{if } j_i\leq t_i \text{ for all } i,\\ 0 & \text{otherwise}\end{cases}\\
&= \frac{s!}{\mathbf{t}!}\,\mathfrak{D}(f(\mathbf{x}))|_{\mathbf{x}=\mathbf{t}}.
\end{aligned}$$

The general result now follows by linearity. □

Proof of Theorem 4.1: Comparing (2.1) and (4.1) we find that, for $\gamma = (\gamma_1, \dots, \gamma_l) \vdash n$,

$$R_\gamma(x, y) = 2^{l-1} \prod_{i=1}^{l} \gamma_i \cdot \sum_{p \geq 0} S_p(\gamma) x^{n-l-2p} y^{2p+l-1}.$$

Thus

$$r^\gamma_{j,k} = \begin{cases} \dfrac{2^{\ell(\gamma)-1} z_\gamma}{\mathrm{Aut}(\gamma)} S_p(\gamma) & \text{if } (j, k) = (n - \ell(\gamma) - 2p, \ell(\gamma) + 2p - 1), \\ 0 & \text{otherwise.} \end{cases}$$

From this and Theorem 2.1 we immediately have

$$c^{(n)}_{\alpha_1,\dots,\alpha_m} = \frac{n^{m-1}}{2^{2g} \prod_i \mathrm{Aut}(\alpha_i)} \sum_{\mathbf{p}} \left[\frac{\mathbf{x}^{\mathbf{r}-\mathbf{2p}}}{(\mathbf{r}-\mathbf{2p})!} \right] \sum_{\ell(\lambda)=n-1} \frac{e_{2\lambda-1}(\mathbf{x})}{\mathrm{Aut}(\lambda)} \prod_{i=1}^{m} (\ell(\alpha_i) + 2p_i - 1)!\, S_{p_i}(\alpha_i),$$

where the outer sum extends over all tuples $\mathbf{p} = (p_1, \dots, p_m)$ of nonnegative integers. Now

$$\left[\frac{\mathbf{x}^{\mathbf{r}-\mathbf{2p}}}{(\mathbf{r}-\mathbf{2p})!} \right] \sum_{\ell(\lambda)=n-1} \frac{e_{2\lambda-1}(\mathbf{x})}{\mathrm{Aut}(\lambda)} = \sum_{s=0}^{n-1} \sum_{\substack{\lambda \vdash q \\ \ell(\lambda)=n-1-s}} \left[\frac{\mathbf{x}^{\mathbf{r}-\mathbf{2p}}}{(\mathbf{r}-\mathbf{2p})!} \right] \frac{e_1(\mathbf{x})^s}{s!} \frac{e_{2\lambda+1}(\mathbf{x})}{\mathrm{Aut}(\lambda)}, \tag{4.3}$$

where q is chosen to make $e_1(\mathbf{x})^s e_{2\lambda+1}(\mathbf{x})$ of total degree $\sum_i (r_i - 2p_i)$. In particular, if $\lambda \vdash q$ and $\ell(\lambda) = n - 1 - s$, then $e_1(\mathbf{x})^s e_{2\lambda+1}(\mathbf{x})$ is of degree $|2\lambda + 1| + s = 2|\lambda| + \ell(\lambda) + s = 2q + n - 1$, so we require

$$2q + n - 1 = \textstyle\sum_i (r_i - 2p_i) = (2g + n - 1) - \sum_i 2p_i,$$

or simply $q + p_1 + \cdots + p_m = g$. Finally, applying the lemma to the RHS of (4.3) results in

$$\sum_{s=0}^{n-1} \sum_{\substack{\lambda \vdash q \\ \ell(\lambda)=n-1-s}} \mathfrak{D}\left(\frac{e_{2\lambda+1}(\mathbf{x})}{\mathrm{Aut}(\lambda)} \right) \bigg|_{\mathbf{x}=\mathbf{r}-\mathbf{2p}} = P^{n-1}_q(\mathbf{r} - \mathbf{2p})$$

and this completes the proof. □

References

1. P. Biane, *Nombre de factorisations d'un grand cycle*, Sém. Lothar. de Combinatoire **51** (2004).
2. M. Bousquet-Mélou and G. Schaeffer, *Enumeration of planar constellations*, Adv. in Appl. Math. **24** (2000), 337–368.
3. I.P. Goulden and D.M. Jackson, *The combinatorial relationship between trees, cacti and certain connection coefficients for the symmetric group*, European J. Combin. **13** (1992), 357–365.
4. A. Goupil and G. Schaeffer, *Factoring n-cycles and counting maps of given genus*, European J. Combin. **19** (1998), 819–834.
5. I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press (Clarendon), Oxford, 1979.
6. D. Poulalhon and G. Schaeffer, *Factorizations of large cycles in the symmetric group*, Discrete Math. **254** (2002), 433–458.
7. R.P. Stanley, *Enumerative Combinatorics: Volume 2*, Cambridge University Press, 1999.

E-mail address: `john.irving@smu.ca`

Department of Mathematics and Computer Science, St. Mary's University, Halifax, Canada

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Shellable complexes and topology of diagonal arrangements

Sangwook Kim

Abstract. We prove that if a simplicial complex Δ is shellable, then the intersection lattice L_Δ for the corresponding diagonal arrangement $\mathcal{A}_\Delta$ is homotopy equivalent to a wedge of spheres. Furthermore, we describe precisely the spheres in the wedge, based on the data of shelling.

Résumé. Nous prouvons que si un complexe simplicial Δ est shellable, alors le treillis d'intersection L_Δ pour le correspondre l'arrangement diagonal $\mathcal{A}_\Delta$ est l'équivalent de homotopy à un bouquet de sphères. De plus, nous décrivons précisément les sphères dans le bouquet, basé sur les données d'écaler.

1. Introduction

Consider $\mathbb{R}^n$ with coordinates $u_1, \ldots, u_n$. A *diagonal subspace* $U_{i_1 \cdots i_r}$ is a linear subspace of the form $u_{i_1} = \cdots = u_{i_r}$ with $r \geq 2$. A *diagonal arrangement* (or a *hypergraph arrangement*) $\mathcal{A}$ is a finite set of diagonal subspaces of $\mathbb{R}^n$.

For a simplicial complex Δ on $[n] = \{1, \ldots, n\}$ such that $\dim \Delta \leq n-3$, one can associate a diagonal arrangement $\mathcal{A}_\Delta$ as follows. For a facet F of Δ, let $U_{\overline{F}}$ be the diagonal subspace $u_{i_1} = \cdots = u_{i_r}$ where $\overline{F} = [n] - F = \{i_1, \ldots, i_r\}$. Define

$$\mathcal{A}_\Delta = \{U_{\overline{F}} | F \text{ is a facet of } \Delta\}.$$

For each diagonal arrangement $\mathcal{A}$, one can find a simplicial complex Δ such that $\mathcal{A} = \mathcal{A}_\Delta$.

Two important spaces associated with an arrangement $\mathcal{A}$ of linear subspaces in $\mathbb{R}^n$ are

$$\mathcal{M}_\mathcal{A} = \mathbb{R}^n - \bigcup_{H \in \mathcal{A}} H \quad \text{and} \quad \mathcal{V}_\mathcal{A}^\circ = \mathbb{S}^{n-1} \cap \bigcup_{H \in \mathcal{A}} H,$$

called the *complement* and the *singularity link* of $\mathcal{A}$.

We are interested in the topology of $\mathcal{M}_\mathcal{A}$ and $\mathcal{V}_\mathcal{A}^\circ$ for a diagonal arrangement $\mathcal{A}$. We mention here some applications. In computer science, Björner, Lovász and Yao [**3**] find lower bounds on complexity of k-equal problems using the topology of diagonal arrangements (see also [**2**]). In group cohomology, it is well-known that $\mathcal{M}_{\mathcal{B}_n}$ for the braid arrangement $\mathcal{B}_n$ in $\mathbb{C}^n$ is a $K(\pi, 1)$ space with the fundamental group isomorphic to the pure braid group ([**6**]). Khovanov [**9**] shows that $\mathcal{M}_{\mathcal{A}_{n,3}}$ for the 3-equal arrangement $\mathcal{A}_{n,3}$ in $\mathbb{R}^n$ is also a $K(\pi, 1)$ space.

Note that $\mathcal{M}_\mathcal{A}$ and $\mathcal{V}_\mathcal{A}^\circ$ are related by Alexander duality as follows:

$$H^i(\mathcal{M}_\mathcal{A}; \mathbb{F}) = H_{n-2-i}(\mathcal{V}_\mathcal{A}^\circ; \mathbb{F}) \qquad (\mathbb{F} \text{ is any field}) \tag{1.1}$$

In the mid 1980's Goresky and MacPherson [**7**] found a formula for the Betti numbers of $\mathcal{M}_\mathcal{A}$, while the homotopy type of $\mathcal{V}_\mathcal{A}^\circ$ was computed by Ziegler and Živaljević [**14**] (see Section 4). The answers are phrased in terms of the lower intervals in the *intersection lattice* $L_\mathcal{A}$ of the subspace arrangement $\mathcal{A}$, that is the collection of all nonempty intersections of subspaces of $\mathcal{A}$ ordered by reverse inclusion. For general subspace arrangements, these lower intervals in $L_\mathcal{A}$ can have arbitrary homotopy type (see [**14**, Corollary 3.1]).

2000 *Mathematics Subject Classification.* Primary 06A07; Secondary 05E99.

Key words and phrases. shellable, hypergraph complex, diagonal subspace arrangement, Golod ring.

This research forms part of the author's doctoral dissertation at the University of Minnesota, under the supervision of Victor Reiner, and partially supported by NSF grant DMS-0245379.

Our goal is to find a general sufficient condition for the intersection lattice $L_{\mathcal{A}}$ of a diagonal arrangement $\mathcal{A}$ to be well-behaved. Björner and Welker [**4**] show that $L_{\mathcal{A}_{n,k}}$ is shellable, and hence has the homotopy type of a wedge of spheres, where $\mathcal{A}_{n,k}$ is the k-equal arrangement consisting of all $U_{i_1\cdots i_k}$ for all $1 \leq i_1 < \cdots < i_k \leq n$ (see Section 2). Kozlov [**11**] shows that $L_{\mathcal{A}}$ is shellable if $\mathcal{A}$ satisfies some conditions (see Section 2). Suggested by a homological calculation (Theorem 4.4 below), we will prove the following main result, capturing the homotopy type assertion from [**11**] (see Section 3).

THEOREM 1.1. *Let Δ be a shellable simplicial complex. Then the intersection lattice L_Δ for the diagonal arrangement $\mathcal{A}_\Delta$ is homotopy equivalent to a wedge of spheres.*

Furthermore, one can describe precisely the spheres in the wedge, based on the shelling data. Let Δ have vertex set $[n]$ with a shelling order $F_1, \ldots, F_q$ on its facets. Let σ be the intersection of all facets, and $\bar{\sigma}$ its complement. For each i, let G_i be the face of F_i obtained by intersecting the walls of F_i that lie in the subcomplex generated by $F_1, \ldots, F_{i-1}$, where a *wall of* F_i is a codimension 1 face of F_i. An (*unordered*) *shelling-trapped decomposition* (*of* $\bar{\sigma}$ *over* Δ) is defined to be a family $\{(\bar{\sigma}_1, F_{i_1}), \ldots, (\bar{\sigma}_p, F_{i_p})\}$ such that $\{\bar{\sigma}_1, \ldots, \bar{\sigma}_p\}$ is a decomposition of $\bar{\sigma}$ as a disjoint union

$$\bar{\sigma} = \bigsqcup_{j=1}^{p} \bar{\sigma}_j$$

and $F_{i_1} < \cdots < F_{i_p}$ are facets of Δ such that $G_{i_j} \subseteq \sigma_j \subseteq F_{i_j}$ for all j. Then the wedge of spheres in Theorem 1.1 consists of $(p-1)!$ copies of spheres of dimension

$$p(2-n) + \sum_{j=1}^{p} |F_{i_j}| + |\bar{\sigma}| - 3$$

for each shelling-trapped decomposition $D = \{(\bar{\sigma}_1, F_{i_1}), \ldots, (\bar{\sigma}_p, F_{i_p})\}$ of $\bar{\sigma}$. Moreover, for each shelling-trapped decomposition D of $\bar{\sigma}$ and a permutation ω of $[p-1]$, there exists a saturated chain $\overline{C}_{D,\omega}$ (see Section 3) such that removing the simplices corresponding to these chains in $\overline{L}_\Delta$ leaves a contractible simplicial complex.

The following example shows that the intersection lattice in Theorem 1.1 is not shellable in general, even though it has the homotopy type of a wedge of spheres.

EXAMPLE 1.2. Let Δ be a simplicial complex on $\{1,2,3,4,5,6,7,8\}$ with the shelling $123456, 127, 237, 137, 458, 568, 468$. Then $\Delta(U_{78}, \hat{1})$ is a disjoint union of two circles, hence is not shellable. Therefore, the intersection lattice L_Δ for the diagonal arrangement $\mathcal{A}_\Delta$ is also not shellable. The intersection lattice L_Δ is shown in Figure 1 (thick lines represent the open interval $(U_{78}, \hat{1})$).

The next example shows that there is a nonshellable simplicial complex whose intersection lattice is shellable.

EXAMPLE 1.3. Let Δ be a simplicial complex on $\{1,2,3,4\}$ whose facets are 12 and 34. Then Δ is not shellable. But the order complex of $\overline{L}_\Delta$ consists of two vertices, hence is shellable.

2. Some known special cases

In this section, we give Kozlov's theorem and show how its consequence for homotopy type follows from Theorem 1.1. Also, we give Björner and Welker's theorem about the intersection lattice of the k-equal arrangements which can be recovered using Theorem 1.1.

Kozlov [**11**] shows that $\mathcal{A}_\Delta$ has shellable intersection lattice if Δ satisfies some conditions. This class includes k-equal arrangements and all other diagonal arrangements for which the intersection lattice was proved shellable up to now.

THEOREM 2.1. ([**11**, Corollary 3.2]) *Consider a partition of*

$$[n] = E_1 \cup \cdots \cup E_r$$

such that $\max E_i < \min E_{i+1}$ *for* $i = 1, \ldots, r-1$. *Let*

$$f : \{1, 2, \ldots, r\} \to \{2, 3, \ldots\}$$

be a nondecreasing map. Let Δ be a simplicial complex on $[n]$ such that F is a facet of Δ if and only if

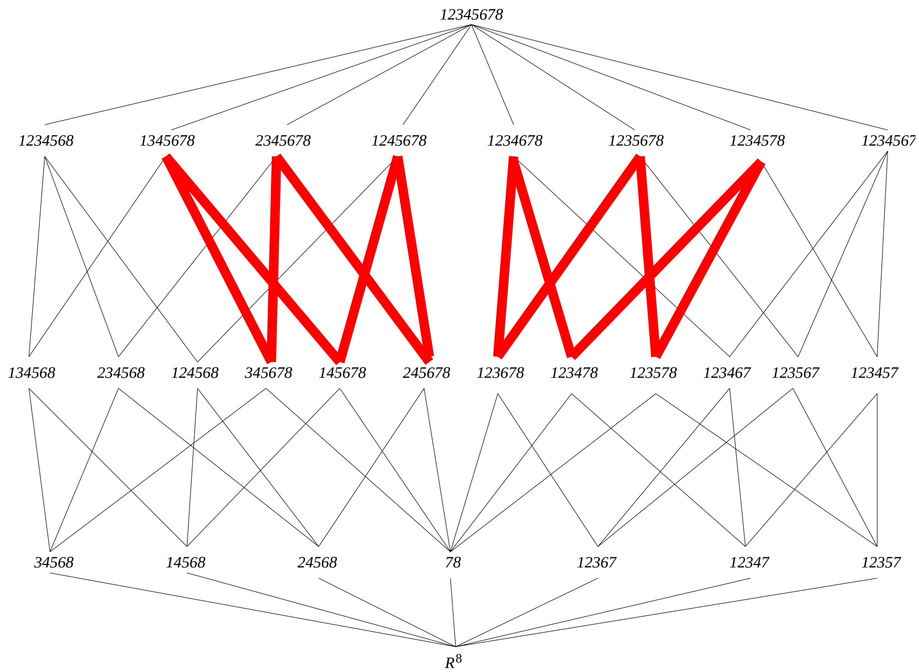


FIGURE 1. The intersection lattice for $\mathcal{A}_\Delta$

$\min \overline{F}$	F	w	$\min \overline{F}$	F	w	$\min \overline{F}$	F	w
1	23456	17	2	1356	247	3	1256	347
	23457	16		1357	246		1257	346
	23467	15		1367	245		1267	345
	23567	14						
	24567	13						
	34567	12						

TABLE 1. Table for Example 2.2

(1) $|E_i - F| \le 1$ *for* $i = 1, \ldots, r$;
(2) if $\min \overline{F} \in E_i$ *then* $|F| = n - f(i)$.

Then the intersection lattice for $\mathcal{A}_\Delta$ *is shellable.*

In particular, this intersection lattice has the homotopy type of a wedge of spheres.

PROPOSITION 2.1. Δ *in Theorem 2.1 is shellable.*

PROOF SKETCH. One checks that a shelling order is $F_1, F_2, \ldots, F_q$ such that the words $w_1, w_2, \ldots, w_q$ are in lexicographic order, where w_i is the increasing array of elements in $\overline{F}_i$. □

EXAMPLE 2.2. Consider the partition of

$$[7] = \{1\} \cup \{2,3\} \cup \{4\} \cup \{5,6,7\}$$

and the function f given by $f(1) = 2, f(2) = 3, f(3) = 4$, and $f(4) = 5$. Then the facets of the simplicial complex that satisfy the conditions of Theorem 2.1 and the corresponding words can be found in Table 1. Thus the ordering 34567, 24567, 23567, 23467, 23457, 23456, 1367, 1357, 1356, 1267, 1257 and 1256 is a shelling for this simplicial complex.

One can also use Theorem 1.1 to recover the following theorem of Björner and Welker [**4**].

Theorem 2.3. *The order complex of the intersection lattice $L_{\mathcal{A}_{n,k}}$ has the homotopy type of a wedge of spheres consisting of*

$$(p-1)! \sum_{0=i_0\leq i_1\leq\cdots\leq i_p=n-pk} \prod_{j=0}^{p-1} \binom{n-jk-i_j-1}{k-1}(j+1)^{i_{j+1}-i_j}$$

copies of $(n-3-p(k-2))$-dimensional spheres for $1\leq p\leq \lfloor\frac{n}{k}\rfloor$.

3. Proof of main theorem

Theorem 1.1 will be deduced from a more general statement about homotopy types of lower intervals $\Delta(\hat{0},H)$ in $L_{\mathcal{A}}$, Theorem 3.1 below.

Theorem 3.1. *Let Δ be a shellable simplicial complex on $[n]$ with a shelling $F_1,\ldots,F_q$ and $\dim\Delta\leq n-3$. Let $U_{\bar\sigma}$ be a subspace in L_Δ for some subset $\bar\sigma$ of $[n]$. Then $\Delta(\hat{0},U_{\bar\sigma})$ is homotopy equivalent to a wedge of spheres, consisting of $(p-1)!$ copies of spheres of dimension*

$$\delta(D):=p(2-n)+\sum_{j=1}^{p}|F_{i_j}|+|\bar\sigma|-3$$

for each shelling-trapped decomposition $D=\{(\bar\sigma_1,F_{i_1}),\ldots,(\bar\sigma_p,F_{i_p})\}$ of $\bar\sigma$.

Moreover, for each such shelling-trapped decomposition D and each permutation ω of $[p-1]$, one can construct a saturated chain $\overline{C}_{D,\omega}$ (see Section 3.1 below), such that if one removes the corresponding $\delta(D)$-dimensional simplices for all pairs (D,ω), the remaining simplicial complex $\widehat{\Delta}(\hat{0},U_{\bar\sigma})$ is contractible.

To prove this result, we begin with some preparatory lemmas.

First of all, one can characterize exactly which subspaces lie in L_Δ when Δ is shellable. Recall that for $\bar\sigma=\{i_1,\ldots,i_r\}\subseteq[n]$, we denote by $U_{\bar\sigma}$ the linear subspace of the form $u_{i_1}=\cdots=u_{i_r}$.

Lemma 3.2. *Let Δ be a simplicial complex on $[n]$ with $\dim\Delta\leq n-3$.*

(1) Every subspace H in L_Δ has the form

$$H=U_{\bar\sigma_1}\cap\cdots\cap U_{\bar\sigma_p}$$

for pairwise disjoint subsets $\bar\sigma_1,\ldots,\bar\sigma_p$ of $[n]$ such that σ_i can be expressed as an intersection of facets of Δ for $i=1,2,\ldots,p$.

(2) Conversely, when Δ is shellable, every subspace H of $\mathbb{R}^n$ that has the above form lies in L_Δ.

The next example shows that Lemma 3.2(2) can fail when Δ is not assumed to be shellable.

Example 3.3. Let Δ be a simplicial complex with two facets 123 and 345. Then Δ is not shellable. Since L_Δ has only three subspaces U_{12},U_{45} and $U_{12}\cap U_{45}$, it does not have the subspace U_{1245}, even though $\overline{1245}=3$ is an intersection of facets 123 and 345 of Δ. Thus Lemma 3.2(2) fails for Δ.

In fact, Lemma 3.2(2) is true for a more general class of simplicial complexes. A simplicial complex is called *locally gallery-connected* if any pair F,F' of facets are connected by a path

$$F=F_0,F_1,\ldots,F_{r-1},F_r=F'$$

of facets in which $F_i\cap F_{i-1}$ share a $(\min\{\dim F_i,\dim F_{i-1}\}-1)$-dimensional face for each i. It is not hard to show that sequentially Cohen-Macaulay simplicial complexes (and hence shellable simplicial complexes) are locally gallery-connected. One can show that Lemma 3.2(2) is true when Δ is locally gallery-connected. Although Lemma 3.2(2) is true for locally gallery-connected simplicial complexes, Theorem 3.1 can fail when Δ is locally gallery-connected. E.g., any triangulation of $\mathbb{RP}^2$ gives a counterexample.

The following lemma shows that every lower interval $[\hat{0},H]$ can be written as a product of lower intervals of the form $[\hat{0},U_{\bar\sigma}]$.

Lemma 3.4. *Let Δ be a simplicial complex on $[n]$ with $\dim\Delta\leq n-3$ and let $H\in L_\Delta$ be a subspace of the form*

$$H=U_{\bar\sigma_1}\cap\cdots\cap U_{\bar\sigma_p}$$

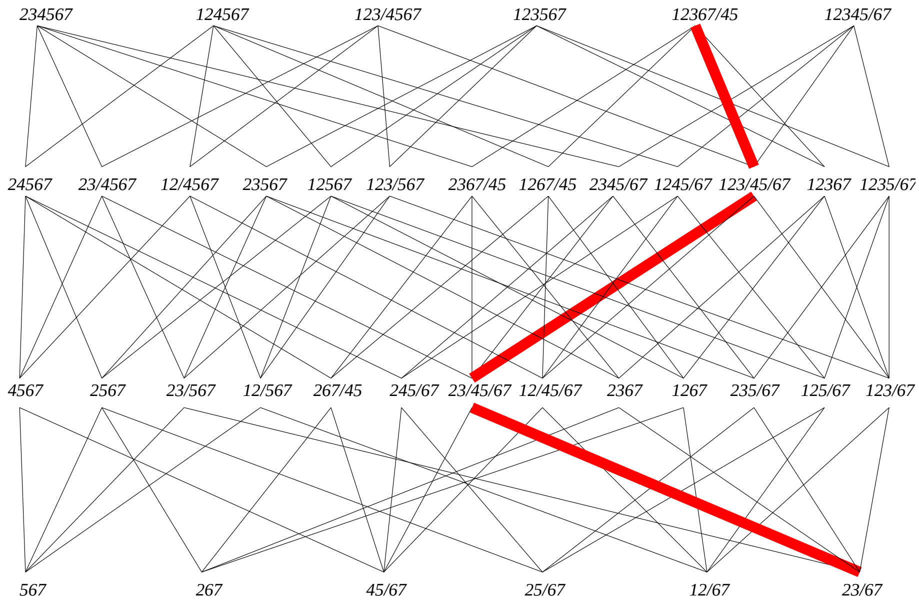


FIGURE 2. The upper interval $(U_{67}, \hat{1})$ in L_Δ

for pairwise disjoint subsets $\bar{\sigma}_1, \ldots, \bar{\sigma}_p$ *of* $[n]$*. Then*

$$[\hat{0}, H] = [\hat{0}, U_{\bar{\sigma}_1}] \times \cdots \times [\hat{0}, U_{\bar{\sigma}_p}].$$

In particular,

$$\Delta(\hat{0}, H) = \Delta(\hat{0}, U_{\bar{\sigma}_1}) * \cdots * \Delta(\hat{0}, U_{\bar{\sigma}_p}) * \mathbb{S}^{p-2},$$

where $*$ *denotes the join of topological spaces.*

PROOF. The first assertion is straightforward, and the second then follows from [**13**, Theorem 4.3]. □

The next lemma, whose proof is completely straightforward and omitted, shows that the lower interval $[\hat{0}, U_{\bar{\sigma}}]$ is *isomorphic* to the intersection lattice for the diagonal arrangement corresponding to $\mathrm{link}_\Delta \sigma$.

LEMMA 3.5. *Let* Δ *be a simplicial complex on* $[n]$ *with* $\dim \Delta \leq n-3$ *and let* $U_{\bar{\sigma}}$ *be a subspace in* L_Δ *for some face* σ *of* Δ*. Then the lower interval* $[\hat{0}, U_{\bar{\sigma}}]$ *is isomorphic to the intersection lattice of the diagonal arrangement* $\mathcal{A}_{link_\Delta(\sigma)}$ *corresponding to* $link_\Delta(\sigma)$ *on the vertex set* $\bar{\sigma}$*.*

The following lemma shows that upper intervals in L_Δ are at least still *homotopy equivalent* to the intersection lattice of a diagonal arrangement.

LEMMA 3.6. *Let* Δ *be a simplicial complex on* $[n]$ *with* $\dim \Delta \leq n-3$ *and let* $U_{\bar{\sigma}}$ *be a subspace in* L_Δ *for some face* $\sigma = \{v_1, \ldots, v_t\}$ *of* Δ*. Then the upper interval* $[U_{\bar{\sigma}}, \hat{1}]$ *is homotopy equivalent to the intersection lattice of the diagonal arrangement* $\mathcal{A}_{\Delta_\sigma}$ *corresponding to the simplicial complex* Δ_σ *on the vertex set* $\{v_1, \ldots, v_t, v\}$ *whose facets are obtained in the following way:*

(A) If $F \cap \sigma$ *is maximal among*

$$\{F \cap \sigma \mid F \text{ is a facet of } \Delta \text{ such that } \sigma \nsubseteq F \text{ and } F \cup \sigma \neq [n]\},$$

then $\widetilde{F} = F \cap \sigma$ *is a facet of* Δ_σ*.*

(B) If a facet F *of* Δ *satisfies* $F \cup \sigma = [n]$*, then* $\widetilde{F} = (F \cap \sigma) \cup \{v\}$ *is a facet of* Δ_σ*.*

EXAMPLE 3.7. Let Δ be a simplicial complex on $\{1,2,3,4,5,6,7\}$ with facets $12367, 12346, 13467, 34567, 13457, 14567, 12345$ and let $\sigma = \{1,2,3,4,5\}$. The open interval $(U_{67}, \hat{1})$ is shown in Figure 2. Then Δ_F is a simplicial complex on $\{1,2,3,4,5,v\}$ and its facets are $123v, 1234, 134v, 345v, 1345, 145v$. The proper part of the intersection lattice L_{Δ_F} is shown in Figure 3 and it is easy to see that its order complex is homotopy equivalent to $(U_{67}, \hat{1})$.

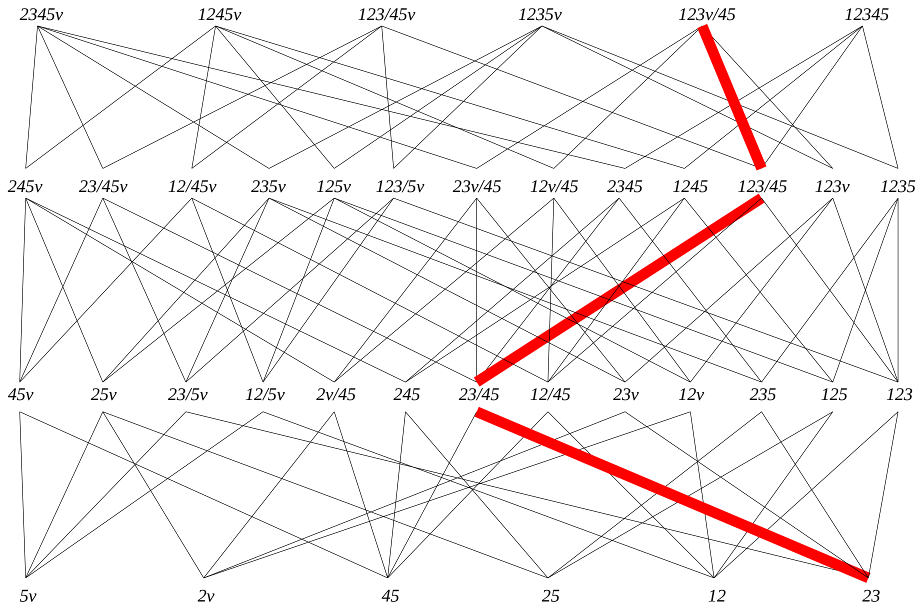


FIGURE 3. The interval $(\hat{0}, \hat{1})$ in L_{Δ_F}

In general, the simplicial complex Δ_σ of Lemma 3.6 is not shellable, even though Δ is shellable (see Example 1.2). However, the next lemma shows that Δ_F is shellable if F is the *last* facet in the shelling order.

LEMMA 3.8. *Let Δ be a shellable simplicial complex on $[n]$ such that $\dim \Delta \leq n-3$ and let F be the last facet in a shelling order of Δ. Then Δ_F is shellable.*

PROOF. Using the notation of Lemma 3.6, a shelling order for Δ_F is the ordering of facets of type (A) in any order, followed by the facets of type (B) according to the order of the corresponding facets of Δ. $\square$

EXAMPLE 3.9. The simplicial complex Δ in Example 3.7 is shellable with a shelling $12367, 12346, 13467, 34567, 13457, 14567, 12345$. Since $1234, 1345$ are facets of Δ_F of type (A) and $123v, 134v, 345v, 145v$ are facets of Δ_F of type (B), $1234, 1345, 123v, 134v, 345v, 145v$ is a shelling of Δ_F.

We next construct the saturated chains appearing in the statement of Theorem 3.1.

3.1. Constructing the chains $C_{D,\omega}$. Let Δ be a shellable simplicial complex on $[n]$ with $\dim \Delta \leq n-3$ and let $U_{\bar{\sigma}}$ is a subspace in L_Δ. Let $D = \{(\bar{\sigma}_1, F_{i_1}), \ldots, (\bar{\sigma}_p, F_{i_p})\}$ be a shelling-trapped decomposition of $\bar{\sigma}$ and let ω be a permutation on $[p-1]$. We define a chain $C_{D,\omega}$ in $[\hat{0}, U_{\bar{\sigma}}]$ as follows:

(1) By Lemma 3.2, the interval $[\hat{0}, U_{\bar{\sigma}}]$ contains $U_{\bar{\sigma}_1} \cap \cdots \cap U_{\bar{\sigma}_p}$ and the interval $[U_{\bar{\sigma}_1} \cap \cdots \cap U_{\bar{\sigma}_p}, U_{\bar{\sigma}}]$ is isomorphic to the set partition lattice Π_p. It is well known that the order complex of $\overline{\Pi}_p = \Pi_p - \{\hat{0}, \hat{1}\}$ is homotopy equivalent to a wedge of $(p-1)!$ spheres of dimension $p-3$ and there is a saturated chain C_ω in Π_p for each permutation ω of $[p-1]$ such that removing $\{\overline{C}_\omega = C_\omega - \{\hat{0}, \hat{1}\} | \omega \in \mathfrak{S}_{p-1}\}$ from the order complex of $\overline{\Pi}_p$ gives a contractible subcomplex (see [**1**, Example 2.9]). Identify $U_{\bar{\sigma}_1}, \cdots, U_{\bar{\sigma}_p}$ with $1, \ldots, p$ in this order and take the saturated chain $\widetilde{C}_\omega$ in $[U_{\bar{\sigma}_1} \cap \cdots \cap U_{\bar{\sigma}_p}, U_{\bar{\sigma}}]$ which corresponds to the chain C_ω in Π_p.

(2) By Lemma 3.4,
$$[\hat{0}, U_{\bar{\sigma}_1} \cap \cdots \cap U_{\bar{\sigma}_p}] \cong [\hat{0}, U_{\bar{\sigma}_1}] \times \cdots \times [\hat{0}, U_{\bar{\sigma}_p}].$$
Since Δ is shellable and $G_{i_j} \subseteq \sigma_j \subseteq F_{i_j}$ for all j, one can see that $[\hat{0}, U_{\bar{\sigma}_j}]$ has a subinterval $[U_{\overline{F}_{i_j}}, U_{\bar{\sigma}_j}]$ which is isomorphic to the boolean algebra of the set of order $|\bar{\sigma}_j| - |\overline{F}_{i_j}|$. Thus
$$[U_{\overline{F}_{i_1}} \cap \cdots \cap U_{\overline{F}_{i_p}}, U_{\bar{\sigma}_1} \cap \cdots \cap U_{\bar{\sigma}_p}]$$

is isomorphic to

$$[U_{\overline{F}_{i_1}}, U_{\bar{\sigma}_1}] \times \cdots \times [U_{\overline{F}_{i_p}}, U_{\bar{\sigma}_p}]$$

and hence is isomorphic to the boolean algebra of the set of order $\sum_{j=1}^{p} \left(|\bar{\sigma}_j| - |\overline{F}_{i_j}|\right)$. Take any saturated chain $\widetilde{C}$ in

$$[U_{\overline{F}_{i_1}} \cap \cdots \cap U_{\overline{F}_{i_p}}, U_{\bar{\sigma}_1} \cap \cdots \cap U_{\bar{\sigma}_p}].$$

(3) Define a saturated chain $C_{D,\omega}$ by

$$\hat{0} \prec U_{\overline{F}_{i_p}} \prec U_{\overline{F}_{i_p}} \cap U_{\overline{F}_{i_{p-1}}} \prec \cdots \prec U_{\overline{F}_{i_p}} \cap \cdots \cap U_{\overline{F}_{i_1}}$$

followed by the chains $\widetilde{C}$ and $\widetilde{C}_\omega$ (where $\prec$ means the covering relation in L_Δ).

Let

$$\overline{C}_{D,\omega} = C_{D,\omega} - \{\hat{0}, U_{\bar{\sigma}}\}.$$

Then $\overline{C}_{D,\omega} \in \Delta(\hat{0}, U_{\bar{\sigma}})$.

Note that the length of this chain $\overline{C}_{D,\omega}$ is

$$\begin{aligned} l(\overline{C}_{D,\omega}) &= p + \sum_{j=1}^{p} \left(|\bar{\sigma}_j| - |\overline{F}_{i_j}|\right) + (p-1) - 2 \\ &= p(2-n) + \sum_{j=1}^{p} |F_{i_j}| + |\bar{\sigma}| - 3. \end{aligned}$$

Example 3.10. Let Δ be the shellable simplicial complex in Example 3.7. Then one can see that

$$D = \{(45, F_1 = 12367), (123, F_6 = 14567), (67, F_7 = 12345)\}$$

is a shelling-trapped decomposition of $\{1,2,3,4,5,6,7\}$. Let ω be a permutation in $\mathfrak{S}_2$ with $\omega(1) = 2$ and $\omega(2) = 1$. Then the maximal chain C_ω in Π_3 corresponding to ω is $(1 \mid 2 \mid 3) - (1 \mid 23) - (123)$. By identifying U_{45}, U_{123}, U_{67} with $1, 2, 3$ in this order, one can get

$$\widetilde{C}_\omega = U_{45} \cap U_{123} \cap U_{67} \prec U_{45} \cap U_{12367} \prec U_{1234567}.$$

Since $[U_{45} \cap U_{23} \cap U_{67}, U_{45} \cap U_{123} \cap U_{67}]$ is isomorphic to a boolean algebra of the set of order 1, one can take

$$\widetilde{C} = U_{45} \cap U_{23} \cap U_{67} \prec U_{45} \cap U_{123} \cap U_{67}.$$

Thus $C_{D,\omega}$ is the chain

$$\begin{gathered} \hat{0} \prec U_{67} \prec U_{23} \cap U_{67} \prec U_{45} \cap U_{23} \cap U_{67} \\ \prec U_{45} \cap U_{123} \cap U_{67} \prec U_{45} \cap U_{12367} \prec U_{1234567}. \end{gathered}$$

The chain $\overline{C}_{D,\omega}$ is represented by thick lines in Figure 2.

The following lemma gives the relationship between the shelling-trapped decompositions of $[n]$ containing F and the shelling-trapped decompositions of $F \cup \{v\}$.

Lemma 3.11. *Let Δ be a shellable simplicial complex on $[n]$ such that $\dim \Delta \leq n-3$ and let F be the last facet in the shelling order of Δ.*

Then there is a one-to-one correspondence between the set of all pairs (D, ω) of shelling-trapped decompositions D of $[n]$ over Δ containing F and $\omega \in \mathfrak{S}_{|D|-1}$, and the set of all pairs $(\widetilde{D}, \tilde{\omega})$ of shelling-trapped decompositions $\widetilde{D}$ of $F \cup \{v\}$ over Δ_F and $\tilde{\omega} \in \mathfrak{S}_{|\widetilde{D}|-1}$. Moreover, one can choose $\overline{C}_{D,\omega}$ and $\overline{C}_{\widetilde{D},\tilde{\omega}}$ so that $\overline{C}_{D,\omega} - U_{\overline{F}}$ corresponds to $\overline{C}_{\widetilde{D},\tilde{\omega}}$ under the homotopy equivalence in Theorem 3.6.

Example 3.12. Let Δ be the shellable simplicial complex in Example 3.7. In Example 3.10, we had

$$\begin{aligned} C_{D,\omega} = &\hat{0} \prec U_{67} \prec U_{23} \cap U_{67} \prec U_{45} \cap U_{23} \cap U_{67} \\ &\prec U_{45} \cap U_{123} \cap U_{67} \prec U_{45} \cap U_{12367} \prec U_{1234567} \end{aligned}$$

for a shelling-trapped decomposition

$$D = \{(45, F_1 = 12367), (123, F_6 = 14567), (67, F_7 = 12345)\}$$

of $\{1,2,3,4,5,6,7\}$ and a permutation ω in $\mathfrak{S}_2$ with $\omega(1) = 2$ and $\omega(2) = 1$.

Decomposition	Facets
1234	$G_3 = \emptyset \subseteq \overline{1234} = \emptyset \subseteq F_3 = 23$
1234	$G_5 = \emptyset \subseteq \overline{1234} = \emptyset \subseteq F_5 = 34$
$24 \cup 13$	$G_2 = 1 \subseteq \overline{24} = 13 \subseteq F_2 = 13$, $G_4 = 2 \subseteq \overline{13} = 24 \subseteq F_4 = 24$
$34 \cup 12$	$G_1 = 12 \subseteq \overline{34} = 12 \subseteq F_1 = 12$, $G_5 = \emptyset \subseteq \overline{12} = 34 \subseteq F_5 = 34$

Table 2. Shelling-trapped decompositions of $\bar{\sigma} = 1234$

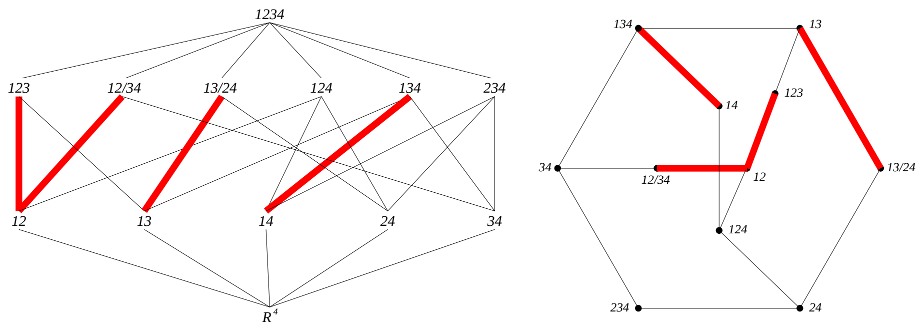


Figure 4. The intersection lattice for Δ and the order complex for its proper part

Since $67 = \overline{F}_7$, the corresponding shelling-trapped decomposition $\widetilde{D}$ of $\{1, 2, 3, 4, 5, v\}$ is

$$\widetilde{D} = \{(45, \widetilde{F}_1 = 123v), (123v, \widetilde{F}_6 = 145v)\}$$

and the corresponding permutation $\tilde{\omega} \in \mathfrak{S}_1$ is the identity.

The corresponding chain $\overline{C}_{\widetilde{D},\tilde{\omega}}$ is

$$\hat{0} \prec U_{23} \prec U_{45} \cap U_{23} \prec U_{45} \cap U_{123} \prec U_{45} \cap U_{123v}.$$

Proof Sketch of Theorem 3.1. One can consider the following decomposition of $\widehat{\Delta}(\overline{L})$:

$$\widehat{\Delta}(\overline{L}) = \widehat{\Delta}(\overline{L} - \{H\}) \cup \widehat{\Delta}(\overline{L}_{\geq H}),$$

where $\widehat{\Delta}(\overline{L} - \{H\})$ is obtained by removing all chains $\overline{C}_{D,\omega}$ not containing H from $\overline{L} - \{H\}$ and $\widehat{\Delta}(\overline{L}_{\geq H})$ is obtained by removing $\overline{C}_{D,\omega}$ and $\overline{C}_{D,\omega} - H$ from $\overline{L}_{\geq H}$ for all $\overline{C}_{D,\omega}$ containing H. Then one can show that all three spaces $\widehat{\Delta}(\overline{L} - \{H\})$, $\widehat{\Delta}(\overline{L}_{\geq H})$ and their intersection are contractible, and hence $\widehat{\Delta}(\overline{L})$ is also contractible. □

Example 3.13. Let Δ be a simplicial complex with a shelling

$$F_1 = 12, \quad F_2 = 13, \quad F_3 = 23, \quad F_4 = 24, \quad F_5 = 34.$$

Then

$$G_1 = 12, \quad G_2 = 1, \quad G_3 = \emptyset, \quad G_4 = 2, \quad G_5 = \emptyset.$$

Let $\bar{\sigma} = 1234$. Then there are four possible shelling-trapped decompositions of $\bar{\sigma}$ (see Table 2). Thus $\Delta(\hat{0}, U_{1234})$ is homotopy equivalent to a wedge of four circles. The intersection lattice and the order complex for its proper part are shown in Figure 4. Note that the chains and the simplices corresponding to each shelling-trapped decomposition are represented by thick lines.

4. The homology of the singularity link of $\mathcal{A}_\Delta$

In this section, we give the corollary about the homotopy type of the singularity link of $\mathcal{A}_\Delta$ when Δ is shellable. Also we give the homology version of the corollary.

Ziegler and Živaljević [**14**] show the following theorem about the homotopy type of $\mathcal{V}^\circ_{\mathcal{A}}$.

THEOREM 4.1. *For every subspace arrangement $\mathcal{A}$ in $\mathbb{R}^n$,*

$$\mathcal{V}^\circ_{\mathcal{A}} \simeq \bigvee_{x \in L_{\mathcal{A}} - \{\hat{0}\}} (\Delta(\hat{0}, x) * \mathbb{S}^{\dim(x)-1}).$$

From this and our results in Section 3, one can deduce the following.

COROLLARY 4.2. *Let Δ be a shellable simplicial complex on $[n]$ with $\dim \Delta \leq n-3$. The singularity link of $\mathcal{A}_\Delta$ has the homotopy type of a wedge of spheres, consisting of $p!$ spheres of dimension*

$$n + p(2-n) + \sum_{j=1}^{p} |F_{i_j}| - 2$$

for each shelling-trapped decomposition $\{(\bar{\sigma}_1, F_{i_1}), \ldots, (\bar{\sigma}_p, F_{i_p})\}$.

REMARK 4.3. The following theorem is a homology version of this corollary.

THEOREM 4.4. *Let Δ be a shellable simplicial complex on $[n]$ with $\dim \Delta \leq n-3$ and $F_1, \ldots, F_q$ be the shelling order on facets of Δ. Then $\dim_{\mathbb{F}} H_i(\mathcal{V}^\circ_{\mathcal{A}_\Delta}; \mathbb{F})$ is the number of ordered shelling-trapped decompositions $((\bar{\sigma}_1, F_{i_1}), \ldots, (\bar{\sigma}_p, F_{i_p}))$ with $i = n + p(2-n) + \sum_{j=1}^{p} |F_{i_j}| - 2$.*

This last result can be proven without Theorem 3.1 by combining

(1) a result of Peeva, Reiner and Welker [**12**, Theorem 1.3],
(2) results of Herzog, Reiner and Welker [**8**, Theorem 4, Theorem 9],
(3) the theory of *Golod rings*.

It is what motivated Corollary 4.2 and eventually Theorem 1.1.

5. $K(\pi, 1)$ examples from matroids

Davis, Januszkiewicz and Scott [**5**] show the following theorem.

THEOREM 5.1. *Let $\mathcal{H}$ be a simplicial real hyperplane arrangement in $\mathbb{R}^n$. Let $\mathcal{A}$ be any arrangement of codimension-2 intersection subspaces in $\mathcal{H}$ which intersects every chamber in a codimension-2 subcomplex. Then $\mathcal{M}_{\mathcal{A}}$ is $K(\pi, 1)$.*

REMARK 5.2. In order to apply this to diagonal arrangements, we need to consider hyperplane arrangements $\mathcal{H}$ which are subarrangements of the braid arrangement $\mathcal{B}_n$ and also simplicial. It turns out (and we omit the straightforward proof) that all such arrangements $\mathcal{H}$ are direct sums of smaller braid arrangements. So we only consider $\mathcal{H} = \mathcal{B}_n$ itself here.

COROLLARY 5.3. *Let $\mathcal{A}$ be diagonal arrangement of codimension 2 subspaces inside $\mathcal{H} = \mathcal{B}_n$, so that*

$$\mathcal{A} = \{U_{ijk} \mid \{i, j, k\} \in T_{\mathcal{A}}\},$$

for some collection $T_{\mathcal{A}}$ of 3-element subsets of $[n]$. Then $\mathcal{A}$ satisfies the hypothesis of Theorem 5.1 (and hence $\mathcal{M}_{\mathcal{A}}$ is $K(\pi, 1)$) if and only if every permutation w in $\mathfrak{S}_n$ has at least one triple in $T_{\mathcal{A}}$ consecutive.

PROOF. It is easy to see that there is a bijection with chambers of $\mathcal{B}_n$ and permutations $w = w_1 \cdots w_n$ in $\mathfrak{S}_n$. Moreover, each chamber has the form $x_{w_1} > \cdots > x_{w_n}$ with bounding hyperplanes $x_{w_1} = x_{w_2}, \ldots, x_{w_{n-1}} = x_{w_n}$ and intersects the 3-equal subspaces of the form $x_{w_i} = x_{w_{i+1}} = x_{w_{i+2}}$ for $i = 1, 2, \ldots, n-2$. $\square$

A rich source of shellable complexes are the *matroid complexes* $\mathcal{I}(M)$, that is the independent sets of a matroid M. If $\Delta = \mathcal{I}(M)$ for some matroid M, then facets of Δ are bases of M. Therefore

$$\begin{aligned} \mathcal{A}_\Delta &= \{U_{ijk} \mid \{i, j, k\} = [n] - B \text{ for some } B \in \mathcal{B}(M)\} \\ &= \{U_{ijk} \mid \{i, j, k\} \in \mathcal{B}(M^\perp)\}, \end{aligned}$$

where $M^\perp$ is the dual matroid of M.

DEFINITION 5.4. Say a rank 3 matroid M on $[n]$ is *DJS* if its bases $\mathcal{B}(M)$ satisfies the condition of Corollary 5.3.

Note that a matroid M which is DJS gives rise to a diagonal arrangement $\mathcal{A}_\Delta$ for $\Delta = \mathcal{I}(M^\perp)$ which has $\mathcal{M}_{\mathcal{A}_\Delta}$ $K(\pi,1)$ and with the homotopy type of $L_\Delta, \mathcal{V}^\circ_{\mathcal{A}_\Delta}$ all predicted by Theorem 3.1. Unfortunately, the following example shows that matroids are not always DJS in general.

Example 5.5. Let Δ be the boundary of an octahedron. Then it is a simplicial complex on $\{1,2,3,4,5,6\}$ whose facets are $123, 134, 145, 125, 236, 346, 456$ and 256. It is easy to see that it is vertex-decomposable, hence is shellable. Also note that Δ is the independent set complex $\mathcal{I}(M)$ of a matroid M of rank 3 which has three distinct parallel classes $\{1,6\}, \{2,4\}$ and $\{3,5\}$. But,

$$T_{\mathcal{A}_\Delta} = \{123, 134, 145, 125, 236, 346, 456, 256\}$$

and $w = 124356$ is a permutation that does not satisfy the condition of Corollary 5.3.

Thus we look for some subclasses of matroids which are DJS. The following two propositions give some rank 3 matroids which are DJS.

Proposition 5.1. *Let M be a rank 3 matroid on the ground set $[n]$ with no circuits of size 3. Let $P_1, \ldots, P_k$ be distinct parallel classes which have more than one element and let N be the set of all elements which are not parallel with anything else. Then, M is DJS if and only if $\lfloor \frac{|P_1|}{2} \rfloor + \cdots + \lfloor \frac{|P_k|}{2} \rfloor - k < |N| - 2$.*

A simplicial complex Δ on $[n]$ is *shifted* if, for any face of Δ, replacing any vertex i by a vertex $j(< i)$ gives another face in Δ. The *Gale ordering* on all k element subsets of $[n]$ is given by $\{x_1 < \cdots < x_k\}$ is less than $\{y_1 < \cdots < y_k\}$ if

$$x_i \leq y_i \text{ for all } i \text{ and } \{x_1, \ldots, x_k\} \neq \{y_1, \ldots, y_k\}.$$

Then it is known that shifted complexes are exactly the order ideals of Gale ordering. Klivans [**10**] shows the following theorem.

Theorem 5.6. *Let M be a rank 3 loop-coloop free matroid on the ground set $[n]$ such that $\mathcal{I}(M)$ is also shifted. Then its bases are the principal order ideal generated by $\{a, b, n\}$ in the Gale ordering such that $1 < a < b < n$. Moreover, M has the following form:*

(1) elements $b+1, b+2, \ldots, n$ form the unique non-trivial parallel class.
(2) elements $a+1, a+2, \ldots, n$ form a rank 2 flat, and this is the only rank 2 flat which can contain more than two parallelism classes.

From this, one can see the following.

Proposition 5.2. *Let M be the rank 3 matroid on the ground set $[n]$ corresponding to the principal order ideal generated by $\{a, b, n\}$. Then, M is DJS if and only if $\lfloor \frac{n-b}{2} \rfloor < a$.*

Problem: Characterize the rank 3 matroids which are DJS.

References

[1] A. Björner, Shellable and Cohen-Macaulay partially ordered sets, *Trans. Amer. Math. Soc.* **260** (1980), pp. 159-183.
[2] A. Björner and L. Lovász, Linear decision trees, subspace arrangements and Möbius functions, *J. Amer. Math. Soc.* **7** (1994), 677-706.
[3] A. Björner, L. Lovász and A. Yao, Linear decision trees: volume estimates and topological bounds, 24th ACM Symp. on Theory of Computing (May 1992), ACM Press, N.Y., 1992, 170-177.
[4] A. Björner and V. Welker, The homology of "k-equal" manifolds and related partition lattices, *Advances in Math.* **110** (1995), 277-313.
[5] M. Davis, T. Januszkiewicz and R. Scott, Nonpositive curvature of blow-ups, *Selecta Math. (N.S.)* **4** (1998), 491–547.
[6] E. Fadell and L. Neuwirth, Configuration spaces, *Math. Scand.* **10** (1962), 111-118.
[7] M. Goresky and R. D. MacPherson, *Stratified Morse theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas (3)], 14. Springer-Verlag, Berlin, 1988.
[8] J. Herzog, V. Reiner and V. Welker, Componentwise linear ideals and Golod rings, *Michigan Math. J.* **46** (1999), 211-223.
[9] M. Khovanov, Real $K(\pi,1)$ arrangements from finite root systems, *Math. Res. Lett.* **3** (1996), 261-274.
[10] C. Klivans, Shifted matroid complexes, preprint.
[11] D. Kozlov, A class of hypergraph arrangements with shellable intersection lattice, *J. Combin. Theory Ser. A* **86** (1999), 169-176.
[12] I. Peeva, V. Reiner and V. Welker, Cohomology of real diagonal subspace arrangements via resolutions, *Compositio Mathematica* **117** (1999), 99-115.
[13] J. W. Walker, Topology and combinatorics of ordered sets, Thesis, M.I.T., Cambridge, MA, 1981.

[14] G. M. Ziegler and R. T. Živaljević, Homotopy types of subspace arrangements via diagrams of spaces, *Math. Ann.* **295** (1993), 527-548.

School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
E-mail address: swkim@math.umn.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Braided differential calculus and quantum Schubert calculus

Anatol N. Kirillov and Toshiaki Maeno

Abstract. We provide a new realization of the quantum cohomology ring of a flag variety as a certain commutative subalgebra in the cross product of the Nichols-Woronowicz algebras associated to a certain Yetter-Drinfeld module over the Weyl group. We also give a generalization of some recent results by Y.Bazlov to the case of the Grothendieck ring of a flag variety of classical type.

Résumé. Nous fournissons une nouvelle réalisation de l'anneau de la cohomologie quantique d'une variété de drapeaux comme sous-algèbre commutative dans le produit croise des algèbres de Nichols-Woronowicz associées à un certain module de Yetter-Drinfeld sur le groupe de Weyl. Nous donnons aussi une généralisation de résultats récents par Y. Bazlov au cas de l'anneau de Grothendieck d'une variété de drapeaux de type classique.

1. Introduction

The main purpose of this work is

• to construct a model of the quantum cohomology ring of the flag variety G/B corresponding to a semisimple finite-dimensional Lie group G as a quantization of Bazlov's model of the coinvariant algebra of finite Coxeter groups,

• to construct a model for the Grothendieck ring of the flag varieties of classical type, in terms of a braided (and discrete) analogue of the differential calculus.

Such a construction of a model for the classical cohomology ring of a flag variety, and more generally for the coinvariant algebra of a finite Coxeter group, as a subalgebra in a braided Hopf algebra called the Nichols-Woronowicz algebra has been invented recently by Y. Bazlov [**2**]. In the present paper we provide a new realization of the quantum cohomology ring of a flag variety as a certain commutative subalgebra in the braided cross product of the corresponding Nichols-Woronowicz algebra and its dual. We also give a generalization of some results from [**2**] to the case of the Grothendieck ring of the flag variety of classical type.

The K-theoretic counterpart of the theory of the quantum cohomology ring has been invented by Givental and Lee. In their paper [**8**], they study the quantum K-theory for the flag variety in a connection with the difference Toda system. The author hopes to report on the Nichols-Woronowicz model of the quantum Grothendieck ring of the flag variety elsewhere in the near future. A description of the quantum K-ring of the flag variety of type A in terms of generators and relations will be given in our forthcoming paper [**13**].

2. Braided differential calculus

In order to formulate our construction, we will remind of the basic notion on the braided differential calculus and the Nichols-Woronowicz algebra in this section.

Definition 2.1. The category C equipped with a functor $\otimes : C \times C \to C$, a collection of isomorphisms

$$(\Phi_{U,V,W} : (U \otimes V) \otimes W \to U \otimes (V \otimes W))_{U,V,W \in \mathrm{Ob}(C)},$$

Key words and phrases. braided Hopf algebras, quantum cohomology rings, flag varieties.

The authors thank Yuri Bazlov for explaining his result. This work is supported by Grant-in-Aid for Scientific Research.

an object $\mathbf{1} \in \mathrm{Ob}(C)$ and isomorphisms of functors

$$\iota_{\text{left}} : \bullet \otimes \mathbf{1} \xrightarrow{\sim} \mathrm{id}, \quad \iota_{\text{right}} : \mathbf{1} \otimes \bullet \xrightarrow{\sim} \mathrm{id}$$

is called a monoidal category if the following diagrams commute:
(1) (pentagon condition)

$$\begin{array}{ccc} & (U \otimes V) \otimes (W \otimes X) & \\ \nearrow & & \searrow \\ ((U \times V) \otimes W) \otimes X & & U \otimes (V \otimes (W \otimes X)) \\ \downarrow & & \uparrow \\ (U \otimes (V \otimes W)) \otimes X & \longrightarrow & U \otimes ((V \otimes W) \otimes X), \end{array}$$

where all the arrows are induced by Φ,
(2) (triangle condition)

$$\begin{array}{ccc} (U \otimes \mathbf{1}) \otimes V & \xrightarrow{\Phi} & U \otimes (\mathbf{1} \otimes V) \\ \iota \otimes \mathrm{id} \searrow & & \swarrow \mathrm{id} \otimes \iota \\ & U \otimes V. & \end{array}$$

Definition 2.2. A monoidal category $C = (C, \otimes, \Phi, \mathbf{1}, \iota)$ is called a braided category if a collection of functorial isomorphisms

$$(\Psi_{U,V} : U \otimes V \to V \otimes U)_{U,V \in \mathrm{Ob}(C)}$$

is given so that the following hexagon conditions are satisfied:

$$(\Psi \otimes \mathrm{id}) \circ \Phi^{-1} \circ (\mathrm{id} \otimes \Psi) = \Phi^{-1} \circ \Psi \circ \Phi^{-1} : U \otimes (V \otimes W) \longrightarrow (W \otimes U) \otimes V,$$

$$(\mathrm{id} \otimes \Phi) \circ \Phi \circ (\Psi \otimes \mathrm{id}) = \Phi \circ \Psi \circ \Phi : (U \otimes V) \otimes W \longrightarrow V \otimes (W \otimes U).$$

Let us take a braided category C consisting of vector spaces over a fixed field k and a braided vector space $V \in \mathrm{Ob}(C)$. Then, the braiding $\psi_V : V \otimes V \to V \otimes V$ is naturally associated to V, and the pair (V, ψ_V) is used to designate V together with the braiding ψ_V. Note that the morphism ψ_V is not necessarily an involution. Denote by ψ_i the endomorphism on the tensor product $V^{\otimes n}$ obtained by applying ψ_V on the i-th and $(i+1)$-st components of $V^{\otimes n}$. Then the braid relation

$$\psi_i \psi_{i+1} \psi_i = \psi_{i+1} \psi_i \psi_{i+1}$$

is a consequence of the hexagon condition.

The Nichols-Woronowicz algebra provides a natural framework to discuss the braided differential calculus. When a finite-dimensional braided vector space (V, ψ) is given, we can attach naturally a braided Hopf algebra structure to the tensor algebra $T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n}$.

Definition 2.3. A k-algebra A in the braided category C is called a braided algebra if its multiplication $m : A \times A \to A$ commutes with the braiding $\psi = \psi_A$, i.e.

$$(m \otimes \mathrm{id}) \circ (\psi \otimes \mathrm{id}) \circ (\mathrm{id} \otimes \psi) = \psi \circ (\mathrm{id} \otimes m) : A \otimes A \otimes A \to A \otimes A.$$

The tensor algebra $T(V)$ is naturally braided by the braiding $\psi_{T(V)}$ which is uniquely characterized by the conditions:
(1) $T(V)$ is a braided algebra,
(2) $\psi_{T(V)}|_{T^1(V) \otimes T^1(V)} = \psi_V$.

Now we can discuss the braided Hopf algebra structure on the tensor algebra $T(V)$. Define the linear maps $\triangle : V \to V \otimes V$, $S : V \to V$ and $\varepsilon : V \to k$ by

$$\triangle(v) := v \otimes 1 + 1 \otimes v, \ S(v) := -v, \ \varepsilon(v) := 0.$$

Then, one can extend the maps $\triangle$, S and ε to endomorphisms on $T(V)$ so that they respectively define the coproduct, the antipode and the counit of the braided Hopf algebra. In particular, $\triangle$ is made to satisfy the condition

$$(m \otimes m) \circ (\mathrm{id} \otimes \psi \otimes \mathrm{id}) \circ (\triangle \otimes \triangle) = \triangle \circ m, \ \text{ on } T(V) \otimes T(V).$$

We call $T(V)$ the free braided Hopf algebra or the free braided group.

DEFINITION 2.4. Let H and K be braided Hopf algebras provided with a k-linear pairing $\langle\ ,\ \rangle : H \times K \to k$. We say that H and K are dually paired if the following conditions are satisfied:

$$\langle \gamma, \kappa\kappa' \rangle = \langle \gamma_{(1)}, \kappa' \rangle \langle \gamma_{(2)}, \kappa \rangle, \quad \langle \gamma\gamma', \kappa \rangle = \langle \gamma', \kappa_{(1)} \rangle \langle \gamma, \kappa_{(2)} \rangle,$$
$$\langle \gamma, 1 \rangle = \varepsilon_H(\gamma), \quad \langle 1, \kappa \rangle = \varepsilon_K(\kappa), \quad \langle S_H(\gamma), \kappa \rangle = \langle \gamma, S_K(\kappa) \rangle,$$

where we use Sweedler's notation $\triangle(a) = a_{(1)} \otimes a_{(2)}$. If the conditions above are satisfied, the pairing $\langle\ ,\ \rangle$ is called a duality pairing.

Let V^* be the dual vector space of V. Then it has the natural braiding ψ^* dual to ψ. It is nontrivial problem to extend the natural pairing $\langle\ ,\ \rangle : V^* \times V \to k$ to the duality pairing between the braided Hopf algebras $T(V^*)$ and $T(V)$. The construction due to Woronowicz [**22**] guarantees the possibility of such an extension of the pairing $\langle\ ,\ \rangle$. Note that from the braid relation one can define the endomorphism Ψ_w on $V^{\otimes n}$ associated to an element w in the symmetric group S_n with a reduced decomposition $w = s_{i_1} \cdots s_{i_l}$, $s_i = (i, i+1)$, as $\Psi_w := \psi_{i_1} \cdots \psi_{i_l}$. The *Woronowicz symmetrizer* is defined as $\sigma_n(\psi) := \sum_{w \in S_n} \Psi_w$. Then the pairing $\langle\ ,\ \rangle : V^* \times V \to k$ can be extended to the one between $(V^*)^{\otimes n}$ and $V^{\otimes n}$ for each $n \geq 2$ by the formula

$$\langle \alpha_1 \otimes \cdots \otimes \alpha_n, v_1 \otimes \cdots \otimes v_n \rangle := (\alpha_n \otimes \cdots \otimes \alpha_1)(\sigma_n(\psi)(v_1 \otimes \cdots \otimes v_n)), \quad \alpha_i \in V^*, \ v_j \in V.$$

PROPOSITION 2.1. *The free braided Hopf algebras $T(V^*)$ and $T(V)$ are dually paired with respect to the pairing $\langle\ ,\ \rangle : T(V^*) \times T(V) \to k$.*

The dually paired braided Hopf algebras $T(V^*)$ and $T(V)$ are not appropriate objects to perform the braided differential calculus on them, since the kernel of the duality pairing is big in general. The *Nichols-Woronowicz algebra* is a braided Hopf algebra which is obtained as a quotient of the free braided Hopf algebra by the kernel of the duality pairing. Such a construction is due to Majid [**17**]. Note that the kernels

$$I(V^*) := \{\xi \in T(V^*) \mid \langle \xi, x \rangle = 0, \ \forall x \in T(V)\},$$
$$I(V) := \{x \in T(V) \mid \langle \xi, x \rangle = 0, \ \forall \xi \in T(V^*)\}$$

are Hopf ideals.

DEFINITION 2.5. The Nichols-Woronowicz algebras $\mathbf{B}(V^*)$ and $\mathbf{B}(V)$ are the dually paired braided Hopf algebras defined to be the quotients of the free braided Hopf algebras by $I(V^*)$ and $I(V)$ respectively:

$$\mathbf{B}(V^*) := T(V^*)/I(V^*), \quad \mathbf{B}(V) := T(V)/I(V).$$

The following equivalent definition is due to Andruskiewitsch and Schneider [**1**]:

DEFINITION 2.6. The Nichols-Woronowicz algebra $\mathbf{B}(V)$ is the graded braided Hopf algebra characterized by the conditions:
(1) $\mathbf{B}^0(V) = k$,
(2) $V = \mathbf{B}^1(V) = \{x \in \mathbf{B}(V) \mid \triangle(x) = x \otimes 1 + 1 \otimes x\}$,
(3) $\mathbf{B}(V)$ is generated by $\mathbf{B}^1(V)$ as an algebra.

Each element $v \in \mathbf{B}^1(V)$ acts on $\mathbf{B}(V^*)$ as a *twisted derivation* $\overleftarrow{D}_v$ from the right:

$$\overleftarrow{D}_v : \mathbf{B}(V^*) \xrightarrow{\triangle} \mathbf{B}(V^*) \otimes \mathbf{B}(V^*) \xrightarrow{\mathrm{id} \otimes \langle\ ,v\rangle} \mathbf{B}(V^*) \otimes k = \mathbf{B}(V^*).$$

The twisted derivation $\overleftarrow{D}_v$ satisfies the twisted Leibniz rule

$$(fg)\overleftarrow{D}_v = f(g\overleftarrow{D}_v) + f \triangleleft \psi^{-1}(g \otimes \overleftarrow{D}_v),$$

where $f \triangleleft \psi^{-1}(g \otimes \overleftarrow{D}_v) = \sum_i (f\overleftarrow{D}_{v_i})g_i$ if $\psi^{-1}(g \otimes v) = \sum_i v_i \otimes g_i$. This action extends to the left action of the opposite algebra $\mathbf{B}(V)^{op}$ on $\mathbf{B}(V)$. The braided cross product $\mathbf{B}(V)^{op} \rtimes\!\!\!\!\ltimes \mathbf{B}(V^*)$ with respect to the action by the twisted derivations can be identified with the algebra of the braided differential operators acting on $\mathbf{B}(V^*)$. In other words, the algebra structure of $\mathbf{B}(V)^{op} \bowtie \mathbf{B}(V^*)$ is given by the multiplication rule

$$(u \otimes x) \cdot (v \otimes y) = u(\psi^{-1}(x \otimes v_{(1)}) \triangleleft v_{(2)})y$$

on $\mathbf{B}(V)^{op} \otimes \mathbf{B}(V^*)$, see [**18**] for details.

At the end of this section, we introduce an important example of the braided categories, which is called the category of the *Yetter-Drinfeld modules.* Let Γ be a finite group.

DEFINITION 2.7. A k-vector space V is called a Yetter-Drinfeld module over Γ, if
(1) V is a Γ-module,
(2) V is Γ-graded, i.e. $V = \bigoplus_{g\in\Gamma} V_g$, where V_g is a linear subspace of V,
(3) for $h \in \Gamma$ and $v \in V_g$, $h(v) \in V_{hgh^{-1}}$.

One of the importance of the category ${}^{\Gamma}_{\Gamma}YD$ of the Yetter-Drinfeld modules over a fixed group Γ is that it is naturally braided. The tensor product of V and W in ${}^{\Gamma}_{\Gamma}YD$ is again a Yetter-Drinfeld module with the Γ-action $g(v \otimes w) = g(v) \otimes g(w)$ and the Γ-grading $(V \otimes W)_g = \bigoplus_{h,h'\in\Gamma,\, hh'=g} V_h \otimes W_{h'}$. The braiding between V and W is defined by $\psi_{V,W}(v \otimes w) = g(w) \otimes v$, for $v \in V_g$ and $w \in W$.

3. Nichols-Woronowicz model of quantum Schubert calculus

Let G be a connected, simply-connected and semi-simple complex Lie group. Fix a Borel subgroup B of G. Denote by Δ the set of roots, which is decomposed into the disjoint union $\Delta = \Delta_+ \sqcup (-\Delta_+)$ by choosing the set of positive roots Δ_+ corresponding to B. Our main interest is a combinatorial structure of the (quantum) cohomology ring of the flag variety G/B. It is well-known that the cohomology ring of the flag variety is isomorphic to the quotient ring of the ring of polynomial functions on the Cartan subalgebra $\mathfrak{h}$ by the ideal generated by the fundamental invariants $f_1, \dots, f_r$, $r = \mathrm{rk}\mathfrak{h}$, of the Weyl group W, i.e.

$$H^*(G/B, \mathbf{Q}) \cong \mathrm{Sym}_{\mathbf{Q}}\mathfrak{h}^*/(f_1, \dots, f_r).$$

On the other hand, the Schubert classes Ω_w, $w \in W$, corresponding to the dual of the cycles $\overline{Bw_0wB/B}$ form a linear basis of $H^*(G/B, \mathbf{Q})$. Then the fundamental problems of the Schubert calculus are stated as follows:

PROBLEM 3.1. (1) Find the natural polynomial representative for the Schubert class Ω_w in the coinvariant algebra $\mathrm{Sym}_{\mathbf{Q}}\mathfrak{h}^*/(f_1, \dots, f_r)$.
(2) Determine the structure constants c_{uv}^w in the multiplication rule

$$\Omega_u\Omega_v = \sum_{w\in W} c_{uv}^w \Omega_w.$$

The answer to the first problem (1) is given for example by the polynomials due to Bernstein, Gelfand and Gelfand [**3**] for general root system, and Schubert polynomials defined by Lascoux and Schützenberger [**14**] for the root system of type A. The latter have nice combinatorial properties. As for the second problem (2), the structure constants c_{uv}^w are complicated in general. However, for special choices of the element $u, v, w \in W$, some combinatorial descriptions of the constants c_{uv}^w, such as Pieri's formula, are known.

The origin of the model of the cohomology ring of the flag variety in terms of a certain noncommutative algebra defined by the data of the root system is the work by Fomin and Kirillov [**5**]. They have introduced an associative $\mathbf{Q}$-algebra $\mathcal{E}_n$, for the root system of type A_{n-1}, generated by the symbols

$$[i, j] = -[j, i], \ \ 1 \le i, j \le n, i \ne j,$$

subject to the quadratic relations:
(1) $[i, j]^2 = 0$,
(2) $[i, j][k, l] = [k, l][i, j]$, if $\{i, j\} \cap \{k, l\} = \emptyset$,
(3) $[i, j][j, k] + [j, k][k, i] + [k, i][i, j] = 0$.

Define the *Dunkl element* $\theta_1, \dots, \theta_n$ in $\mathcal{E}_n$ by

$$\theta_i := \sum_{j\ne i} [i, j].$$

Then one can check the commutativity $\theta_i\theta_i - \theta_j\theta_i = 0$, $\forall i, j$, from the quadratic relations above.

THEOREM 3.1. (Fomin and Kirillov [**5**]) *The subalgebra generated by the Dunkl elements is isomorphic to the cohomology ring of the flag variety Fl_n of type A_{n-1}. The isomorphism is given by*

$$\begin{array}{rcl} \mathcal{E}_n \supset \ \ \mathbf{Q}[\theta_1, \dots, \theta_n] & \to & H^*(Fl_n), \\ \theta_1 + \cdots + \theta_i & \mapsto & \Omega_{s_i}. \end{array}$$

The key tool which connects the algebra $\mathcal{E}_n$ to the Schubert calculus is the following *Bruhat representation.*

DEFINITION 3.2. The Bruhat representation of $\mathcal{E}_n$ is defined to be the representation on the vector space $\bigoplus_{w\in W}\mathbf{Q}\cdot w$ by

$$[i,j]w = \begin{cases} ws_{ij}, & \text{if } l(ws_{ij}) = l(w)+1, \\ 0, & \text{otherwise,} \end{cases}$$

where $i<j$ and s_{ij} is the transposition of i and j.

The algebra $\mathcal{E}_n$ admits a natural quantum deformation which corresponds to the quantum cohomology ring of the flag variety. The quantum cohomology ring $QH^*(G/B)$ of the flag variety G/B also has a structure of a quotient ring of the polynomial ring $\mathrm{Sym}_{\mathbf{Q}}\mathfrak{h}^*\otimes\mathbf{Q}[q_1,\dots,q_r]$, where $q_1,\dots,q_r$ are deformation parameters corresponding to the simple roots. The generators $\tilde{f}_1,\dots,\tilde{f}_r$ of the defining ideal of $QH^*(G/B)$ are explicitly determined by Givental and Kim [**7**] for the root system of type A, and by Kim [**9**] for general root systems. Roughly speaking, they are the conserved quantities of the Toda system. Denote by R the polynomial ring $\mathbf{Q}[q_1,\dots,q_{n-1}]$.

DEFINITION 3.3. The quantum deformed quadratic algebra $\tilde{\mathcal{E}}_n$ is an R-algebra defined by the same symbols and relations as those for the algebra $\mathcal{E}_n$ except that the relation (1) for $\mathcal{E}_n$ is replaced by
(1)'

$$[i,j]^2 = \begin{cases} q_i, & \text{if } i=j-1, \\ 0, & \text{if } i<j-1. \end{cases}$$

The quantized version of the Bruhat representation of $\tilde{\mathcal{E}}_n$ is also defined on $\bigoplus_{w\in W}R\cdot w$. The action of the generator $[i,j]$, $i<j$, is given by

$$[i,j]w = \begin{cases} ws_{ij}, & \text{if } l(ws_{ij}) = l(w)+1, \\ q_iq_{i+1}\cdots q_{j-1}ws_{ij}, & \text{if } l(ws_{ij}) = l(w)-2(j-i)+1, \\ 0, & \text{otherwise,} \end{cases}$$

The Dunkl elements θ_i in the quantized algebra $\tilde{\mathcal{E}}_n$ is defined as before. The following theorem was first conjectured in [**5**] and later proved by Postnikov [**21**].

THEOREM 3.4. *The subalgebra generated by the Dunkl elements is isomorphic to the quantum cohomology ring of the flag variety Fl_n of type A_{n-1}.*

Their description of the (quantum) cohomology ring Fl_n in terms of the algebra $\mathcal{E}_n$ (or $\tilde{\mathcal{E}}_n$) is of use to consider Problem 2.1 combinatorially. See [**5**] and [**21**] for the detail on this point. A generalization to other root systems is treated in [**10**].

The algebra $\mathcal{E}_n$ is defined by generators and relations, so it is a problem to understand its meaning conceptually. The importance of the (braided) Hopf algebra structure of $\mathcal{E}_n$ has been pointed out by [**6**], [**20**] and other works. Now it is conjectured that the algebra $\mathcal{E}_n$ is a kind of the Nichols-Woronowicz algebra. Bazlov [**2**] has constructed a model of the cohomology ring of the flag variety G/B by using a Nichols-Woronowicz algebra $\mathbf{B}_W$ defined below instead of $\mathcal{E}_n$. When we work on the algebra $\mathcal{E}_n$, all the considerations are based on the defining relations and the Bruhat representation. On the other hand, the results on the Nichols-Woronowicz algebra should come from the method of the braided differential calculus. Hence, the argument for the algebra $\mathbf{B}_W$ is completely different from that for $\mathcal{E}_n$.

Let us define a Yetter-Drinfeld module $V=V_W$ over the Weyl group W. We consider a vector space V generated by the symbols $[\alpha]=-[-\alpha]$, $\alpha\in\Delta$:

$$V=\bigoplus_{\alpha\in\Delta}\mathbf{Q}\cdot[\alpha]/([\alpha]+[-\alpha]).$$

The action of $w\in W$ on V is given by $w[\alpha]=[w(\alpha)]$. If we set the W-degree of the symbol $[\alpha]$ to be the reflection s_α, then V becomes a Yetter-Drinfeld module over W. The braiding $\psi: V\otimes V\to V\otimes V$ is given by $\psi([\alpha]\otimes[\beta])=s_\alpha([\beta])\otimes[\alpha]$. The braided vector space V is identified with its dual V^* via a W-invariant inner product on V. We denote by $\mathbf{B}_W$ the Nichols-Woronowicz algebra associated to the Yetter-Drinfeld module V.

REMARK 3.5. It is conjectured that the Nichols-Woronowicz algebra $\mathbf{B}_W$ for A_{n-1} should be isomorphic to the Fomin-Kirillov quadratic algebra $\mathcal{E}_n$. This conjecture is now confirmed up to $n=6$.

Consider a W-homomorphism $\mu_0 : \mathfrak{h}^* \to V$. The homomorphism μ_0 can be written as

$$\mu_0(x) = \sum_{\alpha \in \Delta_+} c_\alpha(\alpha, x)[\alpha],$$

by using a set of W-invariant constants $(c_\alpha)_{\alpha\in\Delta}$. Then the following result corresponds to the commutativity of the Dunkl elements.

PROPOSITION 3.1. *The image of μ_0 generates a commutative subalgebra in $\mathbf{B}_W$.*

Then μ_0 can be extended to an algebra homomorphism $\mu : \mathrm{Sym}_{\mathbf{Q}}\mathfrak{h}^* \to \mathbf{B}_W$.

THEOREM 3.6. (Bazlov [**2**]) *If μ_0 is injective, the image of μ is isomorphic to the cohomology ring $H^*(G/B, \mathbf{Q})$.*

REMARK 3.7. Bazlov proved the theorem above for arbitrary finite Coxeter groups and for their coinvariant algebras (over $\mathbf{R}$).

The braided differential operator $\overleftarrow{D}_\alpha = \overleftarrow{D}_{[\alpha]}$ plays an important role for the proof of Theorem 2.3. Indeed, the following properties
(1) $\mu(f)\overleftarrow{D}_\alpha = c_\alpha\mu(\partial_\alpha f)$,
(2) $\cap_{\alpha\in\Delta_+}\mathrm{Ker}(\overleftarrow{D}_\alpha) = \mathbf{B}_W^0 (= \mathbf{Q})$
imply the result. Here, we denote by ∂_α the divided difference operator on $\mathrm{Sym}_{\mathbf{Q}}\mathfrak{h}^*$:

$$\partial_\alpha(f) := \frac{f - s_\alpha(f)}{\alpha}.$$

We introduce a quantum deformed version of Bazlov's construction. Let $R = \mathbf{Q}[q^{\alpha^\vee} | \alpha \in \Delta_+]$, where the parameters q^a satisfy the condition $q^{a+b} = q^a q^b$. We denote by $\mathbf{B}_{W,R}$ the scalar extension $R \otimes \mathbf{B}_W$. Since the twisted derivations $\overleftarrow{D}_\alpha$ satisfy the Coxeter relations, one can define the operators $\overleftarrow{D}_w$ for any elements $w \in W$ by $\overleftarrow{D}_w = \overleftarrow{D}_{\alpha_1} \cdots \overleftarrow{D}_{\alpha_l}$ for a reduced decomposition $w = s_{\alpha_1} \cdots s_{\alpha_l}$.

DEFINITION 3.8. Let $(c_\alpha)_{\alpha\in\Delta}$ be a set of nonzero constants with the condition $c_\alpha = c_{w\alpha}$, $w \in W$. For each root $\alpha \in \Delta_+$, we define an element $\widetilde{[\alpha]}$ in the algebra of braided differential operators $\mathbf{B}_{W,R}^{op} \bowtie \mathbf{B}_{W,R}$ by

$$\widetilde{[\alpha]} := \begin{cases} c_\alpha[\alpha] + d_\alpha q^{\alpha^\vee}\overleftarrow{D}_{s_\alpha}, & \text{if } l(s_\alpha) = 2\mathrm{ht}(\alpha^\vee) - 1, \\ c_\alpha[\alpha], & \text{otherwise.} \end{cases},$$

where $d_\alpha = (c_{\alpha_1} \cdots c_{\alpha_l})^{-1}$.

Let $\tilde{\mu}_0$ be a W-homomorphism $\mathfrak{h}_R \to \mathbf{B}_{W,R}^{op} \otimes (R \oplus V_R)$ given by

$$\tilde{\mu}_0(x) = \sum_{\alpha\in\Delta_+} (\alpha, x)\widetilde{[\alpha]}.$$

The image of $\tilde{\mu}_0$ again generates a commutative subalgebra in $\mathbf{B}_{W,R}^{op} \bowtie \mathbf{B}_{W,R}$, so it can be extended to an algebra homomorphism $\tilde{\mu} : \mathrm{Sym}_R\mathfrak{h}_R^* \to \mathbf{B}_{W,R}^{op} \bowtie \mathbf{B}_{W,R}$. Now we can state our main result:

THEOREM 3.9. ([**11**]) *The image of $\tilde{\mu}$ is isomorphic to the quantum cohomology ring of the flag variety G/B.*

The key fact to prove this theorem is that the action of the operator $\tilde{\mu}_0(x)$ on $\mathrm{Im}(\mu)$ coincides with the quantization operator by Fomin, Gelfand and Postnikov [**4**] for A_{n-1} and by Maré [**19**] for other root systems.

4. Model of the Grothendieck ring

The Nichols-Woronowicz model of the Grothendieck ring $K(G/B)$ of the holomorphic vector bundles on the flag variety G/B has also been constructed for the classical root systems and G_2 in [**12**]. In this section, we briefly show the construction of the model of $K(G/B)$ for the root system of type B_n.

Let $e_1, \ldots, e_n$ be an orthonormal basis of $\mathfrak{h}^*$, and $\{\pm e_i \pm e_j, \pm e_i \mid 1 \le i, j \le n, i \ne j\}$ be a standard realization of the root system $\Delta = \Delta(B_n)$. For simplicity, we use the symbols $[i, j]$, $\overline{[i, j]}$ and $[i]$ to denote $[e_i - e_j]$, $[e_i + e_j]$ and $[e_i]$ in $\mathbf{B}_W$ respectively. Then the elements $h_{ij} := 1 + [i, j]$, $g_{ij} := 1 + \overline{[i, j]}$ and

$h_i := 1 + [i]$ are solutions of the Yang-Baxter equations:
(1) $h_{ij}h_{kl} = h_{kl}h_{ij}$, $g_{ij}g_{kl} = g_{kl}g_{ij}$, for $\{i,j\} \cap \{k,l\} = \emptyset$,

$h_ih_j = h_jh_i$, $h_{ij}g_{ij} = g_{ij}h_{ij}$,

(2) $h_{ij}h_{ik}h_{jk} = h_{jk}h_{ik}h_{ij}$, $h_{ij}g_{ik}g_{jk} = g_{jk}g_{ik}h_{ij}$,

(3) $h_{ij}h_ig_{ij}h_j = h_jg_{ij}h_ih_{ij}$.

The equations (1), (2) and (3) are respectively corresponding to the subsystems of types $A_1 \times A_1$, A_2 and B_2.

Definition 4.1. We define the multiplicative Dunkl elements or the Ruijsenaars-Schneider-Macdonald elements $\Theta_1^B, \dots, \Theta_n^B$ of type B_n by the formula

$$\Theta_i^B := h_{i-1\,i}^{-1}h_{i-2\,i}^{-1}\cdots h_{1\,i}^{-1}\cdot h_i \cdot g_{1\,i}g_{2\,i}\cdots g_{n\,i}\cdot h_i \cdot h_{i\,n}h_{i\,n-1}\cdots h_{i\,i+1}.$$

The multiplicative Dunkl elements Θ_i^D (resp. Θ_i^A) of type D_n (resp. A_{n-1}) are obtained by the specialzation $h_i \mapsto 1$ (resp. $g_{ij} \mapsto 1$ and $h_i \mapsto 1$).

Remark 4.2. The multiplicative Dunkl elements have been also introduced by Lenart and Yong [**15**], [**16**] for the root system of type A.

The commutativity $\Theta_i^B\Theta_j^B = \Theta_j^B\Theta_i^B$ follows from the Yang-Baxter relations.

Theorem 4.3. ([**12**]) *The subalgebra in the Nichols-Woronowicz algera $\mathbf{B}_{B_n}$ generated by the multiplicative Dunkl elements $\Theta_1^B, \dots, \Theta_n^B$ is isomorphic to the Grothendieck ring $K(G/B)$ of the flag variety G/B of type B_n.*

Corollary 4.1. The following identity in the algebra $\mathbf{B}_{B_n}$ holds:

$$\sum_{j=1}^{n}(\Theta_j^B + (\Theta_j^B)^{-1})^k = n \cdot 2^k$$

for all $k \in \mathbf{Z}_{\geq 0}$.

Remark 4.4. The results for D_n and A_{n-1} are obtained by the specializations $h_i \mapsto 1$, $\forall i$, and $h_i \mapsto 1$, $g_{ij} \mapsto 1$, $\forall i,j$, respectively.

The (small) quantum K-ring $QK(Fl_n)$ of the flag variety Fl_n has the following expression by generators and relations:

$$QK(Fl_n) \cong \mathbf{Z}[q_1, \dots, q_{n-1}][X_1, \dots, X_n]/(\varphi_k^q(X),\ k = 1, \dots, n),$$

where

$$\varphi_k^q(X) = \sum_{I \subset \{1,\dots n\}, |I|=k} \prod_{i \in I} X_i \prod_{i \notin I, i+1 \in I}(1 - q_i) - \binom{n}{k}.$$

Let us introduce the quantized multiplicative Dunkl elements by substituting $\widetilde{[ij]}$ defined in Definition 3.8 for $[ij]$ in the definition of Θ_i^A. Here, we put $c_\alpha = 1$. More precisely, we define the quantized multiplicative Dunkl elements $\widetilde{\Theta}_i^A$, $i = 1, \dots, n$, of type A_{n-1} by the formula

$$\widetilde{\Theta}_i^A = (1 - q_{i-1})\tilde{h}_{i-1\,i}^{-1}\tilde{h}_{i-2\,i}^{-1}\cdots\tilde{h}_{1\,i}^{-1}\cdot\tilde{h}_{i\,n}\tilde{h}_{i\,n-1}\cdots\tilde{h}_{i\,i+1},$$

where $\tilde{h}_{ij} := 1 + \widetilde{[ij]} = 1 + [ij] + q_i \cdots q_{j-1}\overleftarrow{D}_{s_{ij}}$, $i < j$.

Theorem 4.5. ([**13**]) *The equalities*

$$\varphi_k^q(\widetilde{\Theta}_1^A, \dots, \widetilde{\Theta}_n^A) = 0,\ k = 1, \dots n,$$

hold in the algebra $\mathbf{B}_{A_{n-1},R}^{op} \rtimes \mathbf{B}_{A_{n-1},R}$.

References

[1] N. Andruskiewitsch and H.-J. Schneider, *Pointed Hopf algebras,* New directions in Hopf algebras, Math. Sci. Res. Inst. Publ., **43**, Cambridge Univ. Press, Cambridge, 2002, 1-68.

[2] Y. Bazlov, *Nichols-Woronowicz algebra model for Schubert calculus on Coxeter groups,* preprint, math.QA/0409206.

[3] I. N. Bernstein, I. M. Gelfand and S. I. Gelfand, *Schubert cells and cohomology of the spaces G/P*, Russian Math. Surveys, **28** (1973), 1-26.

[4] S. Fomin, S. Gelfand and A. Postnikov, *Quantum Schubert polynomials,* J. Amer. Math. Soc., **10** (1997), 565-596.

[5] S. Fomin and A. N. Kirillov, *Quadratic algebras, Dunkl elements and Schubert calculus,* Advances in Geometry, (J.-L. Brylinski, R. Brylinski, V. Nistor, B. Tsygan, and P. Xu, eds.) Progress in Math., **172**, Birkhäuser, 1995, 147-182.

[6] S. Fomin and C. Procesi, *Fibered quadratic Hopf algebras related to Schubert calculus,* J. Algebra, **230** (2000), 174-183.

[7] A. Givental and B. Kim, *Quantum cohomology of flag manifolds and Toda lattices,* Commun. Math. Phys., **168** (1995), 609-641.

[8] A. Givental and Y.-P. Lee, *Quantum K-theory on flag manifolds, finite-difference Toda lattices and quantum groups,* Invent. Math., **151** (2003), 193-219.

[9] B. Kim, *Quantum cohomology of flag manifolds G/B and quantum Toda lattices,* Ann. of Math., **149** (1999), 129-148.

[10] A. N. Kirillov and T. Maeno, *Noncommutative algebras related with Schubert calculus on Coxeter groups,* European J. Combin., **25** (2004), 1301-1325.

[11] A. N. Kirillov and T. Maeno, *A note on quantization operators on Nichols algebra model for Schubert calculus on Weyl groups,* Lett. Math. Phys., **72** (2005). 233-241.

[12] A. N. Kirillov and T. Maeno, *On some noncommutative algebras related to K-theory of flag varieties, part I,* Int. Math. Res. Notices **2005**, no. 60, 3753-3789.

[13] A. N. Kirillov and T. Maeno, *A note on quantum K-theory of flag varieties,* in preparation.

[14] A. Lascoux and M.-P. Schützenberger, *Polynômes de Schubert,* C. R. Acad. Sci. Paris, **294** (1982), 447-450.

[15] C. Lenart, *The K-theory of the flag variety and the Fomin-Kirillov quadratic algebra*, J. Algebra, **285** (2005), 120–135.

[16] C. Lenart and A. Yong, *Lecture notes on the K-theory of the flag variety and the Fomin-Kirillov quadratic algebra,* `http://www.math.umn.edu/~ayong/`

[17] S. Majid, *Free braided differential calculus, braided binomial theorem, and the braided exponential map,* J. Math. Phys., **34** (1993), 4843-4856.

[18] S. Majid, *Foundations of quantum group theory,* Cambridge University Press, Cambridge, 1995.

[19] A.-L. Maré, *The combinatorial quantum cohomology ring of G/B*, J. Alg. Combin., **21** (2005), 331-349.

[20] A. Milinski and H.-J. Schneider, *Pointed indecomposable Hopf algebras over Coxeter groups,* Contemp. Math. **267** (2000), 215-236.

[21] A. Postnikov, *On a quantum version of Pieri's formula,* Advances in Geometry, (J.-L. Brylinski, R. Brylinski, V. Nistor, B. Tsygan and P. Xu, eds.) Progress in Math., **172** Birkhäuser, 1995, 371-383.

[22] S. L. Woronowicz, *Differential calculus on compact matrix pseudogroups (quantum groups),* Commun. Math. Phys., **122** (1989), 125-170.

Research Institute for Mathematical Sciences, Kyoto University, Kyoto 606-8502, Japan
E-mail address: `kirillov@kurims.kyoto-u.ac.jp`

Department of Mathematics, Kyoto University, Kyoto 606-8502, Japan
E-mail address: `maeno@math.kyoto-u.ac.jp`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Kempf collapsing and quiver loci

Allen Knutson and Mark Shimozono

Abstract. Let Q be a Dynkin quiver, that is, a directed graph whose underlying undirected graph has connected components given by Dynkin diagrams of root systems of types A, D, or E. Assign a fixed vector space to each vertex. Consider the set Rep of representations of the quiver Q with these fixed vector spaces. A product G of general linear groups acts on Rep by change of basis at each vertex. A quiver locus Ω is the closure of a G-orbit in Rep. The equivariant cohomology class (resp. K-class) of Ω is known as a quiver polynomial (resp. K-quiver polynomial).

Reineke proved that Ω is the image of a Kempf collapsing, which is a G-equivariant map from a vector bundle over a partial flag manifold. From this we deduce a formula for the quiver polynomial of Ω.

We extend Kempf's construction. On the numerical side, we give a formula for the equivariant cohomology class of the image of a Kempf collapsing. On the geometric side, we give sufficient conditions under which we can compute the equivariant K-class of the image. We observe that these conditions hold for Reineke's Kempf collapsings in types A and D, yielding a formula for the K-quiver polynomials for these loci.

The formulae are BGG/Demazure divided difference operators applied to a product of linear forms.

Résumé. Soit Q un carquois de Dynkin, c'est-à-dire un graphe orienté dont le graphe non-orienté sous-jacent est formé de composantes connexes de diagrammes de Dynkin de type A,D et E. Fixons un espace vectoriel à chaque sommet du graphe. Considérons l'ensemble Rep des représentations du carquois Q avec ces espaces vectoriels. Un produit G de groupes générals linéaires agit sur Rep en effectuant un changement de base à chacun des sommets. Le locus du carquois Ω est la fermeture d'une G-orbite dans Rep. La classe équivariante de la cohomologie (resp. K-classe) de Ω est un polynôme carquois (resp. K-polynôme carquois). Reineke a prouvé que Ω est l'image d'une application de Kempf, qui est une application G-équivariante d'un fibré vectoriel sur une variété de drapeau partielle. De ceci, nous pouvons en déduire une formule pour le polynôme carquois de Ω. Nous étendons la construction de Kempf. Du côté numérique, nous donnons une formule pour la classe équivariante de cohomologie de l'image d'une application de Kempf. Du côté géométrique, nous donnons des conditions suffisantes avec lesquelles nous pouvons calculer la classe K-invariante de l'image. Nous observons que ces conditions sont les mêmes pour l'application de Kempf pour les types A et D, générant une formule pour le K-polynôme carquois pour ces loci. Les formules sont des opérateurs BGG/Demazure de différence divisée appliqués à un produit de formes linéaires.

1. Introduction

Given a quiver representation one may define a torus-stable affine variety called a quiver locus. The universal torus-equivariant cohomology class of a quiver locus is called a **quiver polynomial**. The polynomials associated with the type A quiver admit many beautiful combinatorial formulae involving tableaux [**8**], rc-graphs [**3**] [**17**], lacing diagrams [**23**], factor sequences [**4**], etc. These quiver polynomials have been studied extensively due to their connection with Thom's theory of degeneracy loci [**29**], intersection theory, and Schubert calculus. We list some cases of quiver polynomials in order of increasing generality.

(1) Double Schur polynomials via the Giambelli-Thom-Porteous formula [**26**].
(2) Double Schubert polynomials [**18**].

2000 *Mathematics Subject Classification.* Primary 14N15; Secondary 05E05.
Key words and phrases. quiver polynomials, divided differences.
A. K. was partially supported by the NSF.
M. S. was partially supported by NSF grant DMS-0401012.

(3) Universal Schubert polynomials [**19**]. These specialize to quantum Schubert polynomials [**16**] among others.
(4) Quiver polynomials for the equioriented type A quiver [**4**] [**5**] [**23**].
(5) Quiver polynomials for the type A quiver with arbitrary orientation [**10**].

We present a divided difference formula for the quiver polynomial of any quiver locus belonging to a Dynkin quiver. We also give a divided difference formula for the more refined information given by the K-quiver polynomial, which is a Laurent polynomial associated with a quiver locus. The literature on the K-theoretic classes of degeneracy loci include [**11**] for Grassmannians, [**22**] for matrix Schubert varieties, [**9**] for the K-analogue of universal Schubert polynomials, [**7**] [**12**] [**13**] [**23**] [**24**] for the equioriented type A quiver, and [**10**] for a conjecture for type A with arbitrary orientation.

Our divided difference formulae are obtained through Kempf collapsings. A Kempf collapsing is a suitable map from a fiber bundle over a partial flag variety, to a vector space. This extends a construction of Kempf [**21**], who used it to derive geometric properties of the image of the collapsing map. The instance of this construction as applied to quiver loci has already been given by Reineke [**28**]; we found it independently.

We expect that our method applies to a suitable nontrivial family of quiver loci for quivers that are not necessarily of type ADE.

Since quiver loci are equivariant classes of subvarieties it follows that their multidegrees (the quiver polynomials) satisfy a certain kind of positivity: it is always possible to equivariantly and flatly degenerate a quiver locus Ω to a union $\Omega(0)$ of coordinate subspaces with multiplicities. This leaves the multidegree invariant. The multidegree is additive on maximum degree components, so the quiver polynomial is the positive sum of products of linear forms. Moreover the forms correspond to vectors that lie in an open half space (assuming the torus action was positive, as it is for quiver loci of Dynkin quivers), so positivity is well-defined. A similar formulation of positivity holds for the K-quiver polynomials.

Our divided difference formulae for the quiver and K-quiver polynomials are not obviously positive in the above sense. It would be desirable to obtain manifestly positive combinatorial formulae.

We give some recent examples of positive formulae for quiver polynomials. In the paper [**23**] (which circulated as a preprint in 2003) four positive formulae (pipe, tableau/Schur, component/Schubert, and ratio) were given for the quiver polynomials for the equioriented type A quiver. The pipe formula is positive in the above sense. The Schur formula was previously conjectured to be positive in [**4**]. The component formula was proved independently in [**5**] after its authors were shown the formula in the form of a conjecture. In [**10**] the component formula was generalized to the type A quiver with arbitrary orientation; this formula can also be obtained via Gröbner degeneration as in [**23**]. Previously in [**8**] a Schur-type formula was proved for Fulton's universal Schubert polynomials. In [**13**] [**24**] positive formulae were given for the K-quiver polynomials for the type A equioriented quiver.

2. Vague statement of "numerical" results

Theorem 2.1. (1) *Let Q be a quiver whose underlying undirected graph is a Dynkin diagram of type ADE, d any dimension vector and $\Omega \subset \mathrm{Rep} = \mathrm{Rep}(Q,d)$ a quiver locus. Then the quiver polynomial $H_{\mathrm{Rep}}(\Omega)$ is obtained by applying a divided difference operator to an explicit product of linear forms.*
(2) *For quivers of types A and D, the K-quiver polynomial $K_{\mathrm{Rep}}(\Omega)$ is obtained by applying a divided difference operator to an explicit product of linear forms.*

Conjecturally the formula for $K_{\mathrm{Rep}}(\Omega)$ also holds for quivers of type E. This kind of formula for quiver polynomials, is reminiscent of those for double Schubert and Grothendieck polynomials. However it is different in that for each quiver locus, one starts with a different product of linear forms, whereas all the Schubert and Grothendieck polynomials indexed by a permutation in a given symmetric group, are obtained by applying divided difference operators to a single product of linear forms. These formulae are new even in the equioriented type A case, where the quiver and K-quiver polynomials are known to be certain double Schubert and Grothendieck polynomials respectively, with the y variables set equal to the reverse of the x variables [**23**]. The most important ingredient is the product of linear forms, which depends in a subtle way on cohomological data that is calculated from the quiver locus.

Example 2.2. Let Q consist of two vertices connected by a single arrow, $d = (3,4)$ and $\Omega \subset \mathrm{Rep} = M_{3\times 4}$ the determinantal variety of 3×4 matrices of rank at most two. It is well-known that (in suitable coordinates)

$H_{\mathrm{Rep}}(\Omega)$ is the double Schubert polynomial $\mathfrak{S}_{1,2,5,3,4}(x;y)$, which coincides with the double Schur polynomial $s_2[x-y]$ where $x=(x_1,x_2,x_3)$ and $y=(y_1,y_2,y_3,y_4)$. Let $\partial_i^x f=(f-s_if)/(x_i-x_{i+1})$ where s_i exchanges x_i and x_{i+1}, and define ∂_i^y for the similar operator in the y variables. Then our formula reads

$$\begin{aligned}H_{\mathrm{Rep}}(\Omega)=\partial_2^y\partial_1^y\partial_3^y\partial_2^y\partial_1^x\partial_2^x\partial_1^x\circ\\ (x_1-y_1)(x_1-y_2)(x_1-y_3)(x_1-y_4)(x_2-y_2)(x_2-y_3)(x_2-y_4)(x_3-y_3)(x_3-y_4)\end{aligned}$$

3. Hilbert numerators and multidegrees

Quiver polynomials and K-quiver polynomials are instances of the constructions of the multidegree and Hilbert numerator. We recall these notions, following [**25**].

Let $T=(\mathbb{C}^*)^r$ be an algebraic torus and $X(T)\cong\mathbb{Z}^r$ be the group of algebraic group homomorphisms $T\to\mathbb{C}^*$. We write the group operation on $X(T)$ additively. Let $x_1,\dots,x_r$ be the standard basis of $X(T)$.

Let M be a T-module, that is, a vector space over $\mathbb{C}$ endowed with a rational T-action. For $\lambda\in X(T)$ a vector of weight λ is a nonzero vector $v\in M$ such that $t\cdot v=\lambda(t)v$ for all $t\in T$. Let $M_\lambda\subset M$ be the subspace of vectors of weight λ. Then $M\bigoplus_{\lambda\in X(T)}M_\lambda$. If $\dim M_\lambda<\infty$ for all λ then one may define

$$\mathrm{ch}_T\ M=\sum_{\lambda\in X(T)}\dim M_\lambda\, e^\lambda,$$

which is a formal Laurent series in the variables e^{x_i}.

Let Y be a finite-dimensional T-module. Suppose Y is **positive**, that is, all the weights of Y ($\lambda\in X(T)$ such that $Y_\lambda\neq 0$) lie on one side of a hyperplane in $\mathbb{R}^r$ through the origin. Consider the coordinate ring $\mathbb{C}[Y]$ of Y; it is a polynomial ring in a set $\mathcal{B}$ of coordinate functions on Y, which can be taken to be weight vectors. A basis of weight vectors in $\mathbb{C}[Y]$ is given by the set of monomials with variables in $\mathcal{B}$. Therefore the weight spaces of $\mathbb{C}[Y]$ are finite-dimensional, and using geometric series one obtains

$$\mathrm{ch}_T\ \mathbb{C}[Y]=\prod_{v\in\mathcal{B}}(1-e^{\mathrm{wt}(v)})^{-1}\tag{3.1}$$

where $\mathrm{wt}(v)\in X(T)$ is the weight of v.

Let $Z\subset Y$ be a T-stable closed subscheme, with defining ideal $I(Z)\subset\mathbb{C}[Y]$. Its coordinate ring is $\mathbb{C}[Z]\cong\mathbb{C}[Y]/I(Z)$. Since $\mathbb{C}[Z]$ is a quotient of $\mathbb{C}[Y]$ by a T-stable ideal, it has a basis of weight vectors given by a subset of that of $\mathbb{C}[Y]$. Thus $\mathbb{C}[Z]$ has finite-dimensional weight spaces and $\mathrm{ch}_T\ \mathbb{C}[Z]$ is a well-defined formal Laurent series. The **T-equivariant Hilbert numerator** of Z in the positive T-module Y is the formal Laurent series in the variables e^{x_i} defined by

$$K_Y(Z)=\frac{\mathrm{ch}_T\ \mathbb{C}[Z]}{\mathrm{ch}_T\ \mathbb{C}[Y]}.\tag{3.2}$$

Using a T-equivariant version of the Hilbert Syzygy Theorem it follows that $K_Y(Z)$ is in fact a Laurent polynomial: the formal series $\mathrm{ch}_T\ \mathbb{C}[Z]$ can always be expressed as a Laurent polynomial (namely, $K_Y(Z)$) divided by the denominator of $\mathrm{ch}_T\ \mathbb{C}[Y]$.

There are natural isomorphisms $K_T^*(Y)\cong K_T^*(\mathrm{pt})\cong R(T)=\mathbb{Z}[X(T)]=\mathbb{Z}[e^{\pm x_1},\dots,e^{\pm x_r}]$ where $R(T)$ is the ring of rational representations of T. The Hilbert numerator $K_Y(Z)$ may be regarded as an element in the T-equivariant K-theory $K_T^*(Y)$ of Y; it is the equivariant K-class of the structure sheaf $\mathcal{O}_Z$ of Z.

There is a surjective ring homomorphism $\mathbb{Z}[e^{\pm x_1},\dots,e^{\pm x_r}]\to\mathbb{Z}[x_1,\dots,x_r]$ that sends a Laurent polynomial to its lowest degree nonvanishing homogeneous term, where e^λ is formally expressed as $e^\lambda=\sum_{i\geq 0}\lambda^i/i!$. The **multidegree** of Z is the polynomial $H_Y(Z)$ given by the image of the Hilbert numerator $K_Y(Z)$ under this map. It can be shown that $H_Y(Z)$ is a polynomial with integer coefficients: $H_Y(Z)\in\mathbb{Z}[x_1,\dots,x_r]$. More canonically, there are isomorphisms $H_T^*(Y)\cong H_T^*(\mathrm{pt})\cong\mathrm{Sym}_{\mathbb{Z}}(X(T))\cong\mathbb{Z}[x_1,\dots,x_r]$ where $\mathrm{Sym}_{\mathbb{Z}}$ is the symmetric algebra with integer coefficients. Then $H_Y(Z)$ is identified with the element of the equivariant cohomology ring $H_T^*(Y)$ given by the T-equivariant fundamental class of the T-stable subvariety Z of Y, and the above ring homomorphism is the T-equivariant Chern map $K_T^*(Y)\to H_T^*(Y)$.

Suppose Z is a coordinate subspace, that is, it is defined by the vanishing of some subset $\mathcal{B}'\subset\mathcal{B}$ of the set of coordinates $\mathcal{B}$ of Y. Then directly from the definitions one may easily compute the Hilbert numerator and multidegree:

$$K_Y(Z) = \prod_{v \in \mathcal{B}'} (1 - e^{\mathrm{wt}(v)}) \qquad H_Y(Z) = \prod_{v \in \mathcal{B}'} (-\mathrm{wt}(v)). \tag{3.3}$$

The Hilbert numerator $K_Y(Z)$ is a more subtle geometric invariant than the multidegree $H_Y(Z)$ since the latter is only the leading term of the former.

4. Quiver polynomials

To each quiver representation we define its quiver polynomial and K-quiver polynomial as the multidegree and Hilbert numerator of its associated quiver locus.

A **quiver** is a finite directed graph $Q = (Q_0, Q_1)$ where Q_0 is the set of vertices and Q_1 is the set of directed edges. Each directed edge $a \in Q_1$ has a head $ha \in Q_0$ and a tail $ta \in Q_0$. A **dimension vector** is a function $d : Q_0 \to \mathbb{Z}_{\geq 0}$: it assigns to each vertex $i \in Q_0$ a nonnegative integer $d(i)$. A **representation** V of the quiver Q of dimension d, is a collection of linear maps V_a, one for each arrow $a \in Q_1$, with $V_a : \mathbb{C}^{d(ta)} \to \mathbb{C}^{d(ha)}$. Equivalently, V is a list of matrices where $V_a \in M_{d(ta) \times d(ha)}$; here matrices act on row vectors. Let $\mathrm{Rep} = \mathrm{Rep}(Q, d) = \prod_{a \in Q_1} M_{d(ta) \times d(ha)}$ be the set of representations of Q of dimension d. Say that $V, W \in \mathrm{Rep}$ are **equivalent** if V is taken to W by a change of basis in the vector spaces at the vertices, that is, there is an element $g = (g_i)_{i \in Q_0} \in G = G(Q, d) = \prod_{i \in Q_0} GL(d(i))$ such that $W_a = g_{ta} V_a g_{ha}^{-1}$ for all $a \in Q_1$. Thus an equivalence class of quiver representations of Q of dimension d is a G-orbit in $\mathrm{Rep}(Q, d)$.

A **quiver locus** in Rep is a subvariety of the form $\Omega = \overline{G \cdot V}$ for some $V \in \mathrm{Rep}$. Let $T \subset G$ be the maximal torus consisting of tuples of diagonal matrices. Since Ω is G-stable and closed it is also T-stable and therefore defines T-equivariant classes $K_{\mathrm{Rep}}(\Omega) \in K_T^*(\mathrm{Rep})$ and $H_{\mathrm{Rep}}(\Omega) \in H_T^*(\mathrm{Rep})$. These are by definition the K**-quiver polynomial** and **quiver polynomial** of the quiver locus Ω.

More specifically, let $T^i \subset GL(d(i))$ be the subgroup of diagonal matrices in the i-th component of G for $i \in Q_0$ and let $T = \prod_{i \in Q_0} T^i \subset G$. Let $X^i = \{x_1^i, x_2^i, \ldots, x_{d(i)}^i\}$ be a basis of $X(T^i)$. Then $K_T(\mathrm{Rep}) \cong \mathbb{Z}[e^{\pm x_j^i}]$. Since Ω is G-stable it defines a G-equivariant class in $K_G^*(\mathrm{Rep})$. But there are natural isomorphisms $K_G^*(\mathrm{Rep}) \cong K_T^*(\mathrm{Rep})^W \cong \mathbb{Z}[e^{\pm x_j^i}]^W$ where $W = \prod_{i \in Q_0} S_{d(i)}$ is the Weyl group of G, the product of symmetric groups where $S_{d(i)}$ permutes the i-th set of variables X^i. So $K_{\mathrm{Rep}}(\Omega)$ is a W-symmetric Laurent polynomial. Similarly $H_G^*(\mathrm{Rep}) \cong H_T^*(\mathrm{Rep})^W \cong \mathbb{Z}[x_j^i]^W$, and the quiver polynomial $H_{\mathrm{Rep}}(\Omega)$ is W-symmetric.

Remark 4.1. The above action of T on $\mathrm{Rep}(Q, d)$ is *not* positive if and only if there is some directed cycle C in Q such that for every vertex i on C, $d(i) > 0$. In this situation the Hilbert numerator of some quiver loci in $\mathrm{Rep}(Q, d)$ are not well-defined. However we may consider the action of a bigger group $G \times T'$ on Rep where $T' = (\mathbb{C}^*)^{|Q_1|}$ is a torus with a copy of $\mathbb{C}^*$ for each arrow $a \in Q_1$, where the a-th copy of $\mathbb{C}^*$ acts on the a-th component of $\mathrm{Rep}(Q, d)$ by scaling. The torus $T^+ = T \times T'$ in G^+ acts positively on Rep. In particular if $\Omega \subset \mathrm{Rep}$ is a quiver locus that is also stable under G^+ then its quiver and K-quiver polynomial with respect to the T^+-module Rep, are well-defined. If Q has no directed cycles then the T^+ polynomials specialize to the usual quiver polynomials by setting to zero the basis elements of $X(T')$.

More generally one may consider the Hilbert numerators and multidegrees of G^+-orbit closures or other G^+-stable subvarieties of Rep with respect to the T^+-module Rep.

Example 4.2. Let Q consist of a single vertex and a single loop and fix the dimension n. Then $\mathrm{Rep} = M_n$ is the $n \times n$ matrices and $G = GL(n)$ acts by conjugation. The indecomposables of $\mathbb{C}Q$ are Jordan blocks. With the notation of the previous Remark, the $G^+ = G \times \mathbb{C}^*$-stable quiver loci are the closures of conjugacy classes of nilpotent matrices.

More generally if Q is a directed cycle then the quiver loci given by G-orbits of nilpotent elements of Rep, are also G^+-stable.

5. Representations of Q

We recall some of the representation theory of Dynkin quivers. This provides an indexing set for the quiver loci and other key ingredients for our divided difference formula for the quiver polynomials. See [**14**] [**15**] for excellent survey information.

5.1. Path Algebra. The **path algebra** $\mathbb{C}Q$ is the associative algebra over $\mathbb{C}$ with generating set $Q_0 \cup Q_1$ and relations (for all $i, j \in Q_0$ and $a \in Q_1$)

$$\begin{aligned} i \cdot j &= \delta_{i,j} i \\ i \cdot a &= \delta_{i,ta} a \\ a \cdot j &= \delta_{ha,j} a. \end{aligned} \tag{5.1}$$

Using these relations it follows that for $a, b \in Q_1$, the product ab is zero unless $ha = tb$. Hence $\mathbb{C}Q$ has a basis given by paths, where a path of length zero is an element of Q_0, and a path of length $m > 0$ is a sequence $a_1 a_2 \cdots a_m$ with $a_i \in Q_1$ where $ha_k = ta_{k+1}$ for $1 \leq k \leq m-1$. Since a path has a unique starting vertex and unique ending vertex, it follows that the elements $i \in Q_0$ are a complete set of orthogonal idempotents in $\mathbb{C}Q$. Let $\mathbb{C}Q$-Mod be the category of finite-dimensional right $\mathbb{C}Q$-modules. Let $V \in \mathbb{C}Q$-Mod. We have $V = \bigoplus_{i \in Q_0} V_i$ where $V_i = V \cdot i$. One easily checks that the linear map V_a given by the action of a on V, is zero on V_j for $j \neq ta$ and its image lies in V_{ha}. So without loss we may consider V_a as a linear map from $V_{ta} \to V_{ha}$. Thus we see that a $\mathbb{C}Q$-module is just a quiver representation and vice versa. Let $g : V \to W$ be a $\mathbb{C}Q$-module isomorphism. Firstly g is a linear isomorphism. Since g intertwines the action of $i \in Q_0$, g restricts to an isomorphism $g_i : V_i \to W_i$ for all i. In particular V and W have the same dimension vector d. So we may regard V and W as being elements of $\mathrm{Rep}(Q, d)$. Since g intertwines the action of $a \in Q_1$, it must satisfy $g_{ta} V_a = W_a g_{ha}$ or equivalently $g_{ta} V_a g_{ha}^{-1} = W_a$. Therefore V and W are isomorphic if and only if the corresponding elements of Rep are in the same G-orbit.

So the problem of classifying G-orbits on Rep is the same as that of classifying finite-dimensional $\mathbb{C}Q$-modules up to isomorphism.

5.2. An index set for quiver loci. An **indecomposable** module is one that is not the direct sum of two nonzero submodules. By definition every module is the direct sum of indecomposables. So the isomorphism class of a $\mathbb{C}Q$-module is determined by the multiplicities of its indecomposable summands. Let Ind_Q be the set of isomorphism classes of indecomposable $\mathbb{C}Q$-modules. One special kind of indecomposable module is a **simple** module, one that has no proper submodule. For each vertex $i \in Q_0$ there is a corresponding simple $\mathbb{C}Q$-module S_i: it has $\mathbb{C}^1$ at vertex i and zero vector spaces at the other vertices, and all maps are zero.

Gabriel's Theorem characterizes the quivers Q with finitely many indecomposables.

THEOREM 5.1. [**20**] *The following are equivalent for a quiver Q.*

(1) *$G(Q, d)$ has finitely many orbits on $\mathrm{Rep}(Q, d)$ for all d.*
(2) *Ind_Q is finite.*
(3) *The undirected graph X underlying Q is the Dynkin diagram of a simply-laced root system Φ (that is, its connected components are Dynkin diagrams of type ADE).*

Suppose this holds. Then there is a bijection $\mathrm{Ind}_Q \to \Phi^+$ of the indecomposables with the positive roots Φ^+ of Φ. This bijection sends the simple $\mathbb{C}Q$-module S_i to the simple root α_i and in general sends $I \in \mathrm{Ind}_Q$ to its dimension vector, where a function $d : Q_0 \to \mathbb{Z}$ is identified with the element $\sum_{i \in Q_0} d(i)\alpha_i$ of the root lattice of Φ.

For $\beta \in \Phi^+$ let I_β be the indecomposable with dimension vector β. The modules in Ind_Q may be constructed explicitly using reflection functors [**6**] but we don't require this construction.

EXAMPLE 5.2. The equioriented type A quiver (A_{n+1}) is depicted below; we use the vertex set $Q_0 = \{0, 1, 2, \dots, n\}$ and directed edges going from $a-1$ to a for $1 \leq a \leq n$.

$$\underset{0}{\bullet} \longrightarrow \underset{1}{\bullet} \longrightarrow \underset{2}{\bullet} \longrightarrow \cdots \longrightarrow \underset{n}{\bullet}$$

For $0 \leq i \leq j \leq n$, let I_{ij} be the indecomposable corresponding to the root $\alpha_{ij} = \alpha_i + \alpha_{i+1} + \cdots + \alpha_j$. It can be realized by placing $\mathbb{C}^1$ at vertices i through j with identity maps connecting them and zero maps elsewhere.

REMARK 5.3. Let $\beta \in Q^+$. Say that a map $m : \Phi^+ \to \mathbb{Z}_{\geq 0}$ is a **Kostant partition** of β if $\beta = \sum_{\alpha \in \Phi^+} m(\alpha)\alpha$. By Gabriel's Theorem, the isomorphism classes of $\mathbb{C}Q$-modules of dimension d, are parametrized by the Kostant partitions of d. Write $\Omega_m = \overline{G \cdot \phi}$ for any $\phi \in \mathrm{Rep}$ such that $\phi \cong \bigoplus_{\alpha\in\Phi^+} I_\alpha^{\oplus m(\alpha)}$.

5.3. Hom, Ext, and the Euler form. For $M, N \in \mathbb{C}Q$-Mod let $\mathrm{Hom}_Q(M,N)$ be the vector space of $\mathbb{C}Q$-module homomorphisms from M to N. Let $\mathrm{Ext}^i_Q(-,N)$ be the i-th cohomology group of the functor $\mathrm{Hom}_Q(-,N)$ applied to a projective resolution of M. The homological form on $\mathbb{C}Q$-Mod is defined by

$$\langle M\,,\,N\rangle = \sum_{i\geq 0} (-1)^i \dim \mathrm{Ext}^i_Q(M,N).$$

It is *not* symmetric. The category $\mathbb{C}Q$-Mod is hereditary. In particular $\mathrm{Ext}^i_Q(M,N) = 0$ for all $i \geq 2$, so that

$$\langle M\,,\,N\rangle = \dim \mathrm{Hom}_Q(M,N) - \dim \mathrm{Ext}^1_Q(M,N). \tag{5.2}$$

Ringel [**27**] observed that the homological form $\langle M\,,\,N\rangle$ depends only on the dimension vectors d_M and d_N of M and N. Define the **Euler form** on functions $Q_0 \to \mathbb{Z}$ by

$$\langle d\,,\,e\rangle = \sum_{i\in Q_0} d(i)e(i) - \sum_{a\in Q_1} d(ta)e(ha).$$

Then the homological form on M and N is the Euler form on their dimension vectors:

$$\langle M\,,\,N\rangle = \langle d_M\,,\,d_N\rangle.$$

Remark 5.4. Dynkin quivers are precisely those with positive definite Euler form. For Q a Dynkin quiver, a vector $d : Q_0 \to \mathbb{Z}_{\geq 0}$ is a positive root if and only if $\langle d\,,\,d\rangle = 1$.

5.4. Auslander-Reiten quiver. This material comes from [**1**]. Say that a $\mathbb{C}Q$-module homomorphism f is **irreducible** if it is nonzero, and for every factorization $f = h \circ g$ as a composition of $\mathbb{C}Q$-module homomorphisms, either g is split injective or h is split surjective. The **Auslander-Reiten quiver** Γ_Q of Q is the directed graph with vertex set given by Ind_Q and with a directed edge from M to N if there is an irreducible map $M \to N$.

Proposition 5.1. *Let Q be a Dynkin quiver. Then for $\beta, \gamma \in \Phi^+$, then there is an arrow from I_β to I_γ in Γ_Q if and only if $\beta \neq \gamma$ and $\langle \beta\,,\,\gamma\rangle > 0$.*

Remark 5.5. For Dynkin quivers Q the Auslander-Reiten quiver Q has no cycles. Therefore there is a partial order $\preccurlyeq_Q$ on Ind_Q or Φ^+ given by $\beta \preccurlyeq_Q \gamma$ if there is a directed path from I_β to I_γ in Γ_Q.

Proposition 5.2. *For Q Dynkin and $\alpha, \beta \in \Phi^+$:*

1. $\mathrm{Hom}_Q(I_\alpha, I_\beta) = 0$ *if* $\alpha > \beta$.
2. $\mathrm{Ext}^1_Q(I_\alpha, I_\beta) = 0$ *if* $\alpha \leq \beta$.
3. $\langle \alpha\,,\,\beta\rangle = \dim \mathrm{Hom}_Q(\alpha,\beta)$ *for* $\alpha \leq \beta$.
4. $\langle \alpha\,,\,\beta\rangle = -\dim \mathrm{Ext}^1_Q(\alpha,\beta)$ *for* $\alpha > \beta$.

Proposition 5.2 says that the matrix $\langle \alpha\,,\,\beta\rangle$ for $\alpha, \beta \in \Phi^+$, written with respect to any linear extension of $\preccurlyeq_Q$, agrees with $\dim \mathrm{Hom}_Q(I_\alpha, I_\beta)$ on or above the diagonal and with $-\dim \mathrm{Ext}^1_Q(I_\alpha, I_\beta)$ below the diagonal. Thus one can read off all the important homological information about $\mathbb{C}Q$-Mod just from the Euler form under an appropriate ordering of positive roots.

5.5. Reduced expressions. We recall from [**2**] [**31**] a combinatorial way to construct the Auslander-Reiten quiver Γ_Q when Q is Dynkin. Suppose X is an undirected graph that is the Dynkin diagram of a simply-laced root system Φ, with Weyl group W and distinguished set $\{s_i \in W \mid i \in Q_0\}$ of simple reflections. An **orientation** of X is a directed graph Q that yields X if the directions on edges are forgotten. The Weyl group acts on the set $\mathrm{Or}(X)$ of orientations of X: for $Q \in \mathrm{Or}(X)$ and $i \in Q_0$, $s_i Q \in \mathrm{Or}(X)$ is obtained from Q by reversing all the directed edges that touch the vertex i.

Let $w_0 \in W$ be the longest element. Let Red be the set of reduced words for w_0, that is, the set of sequences $i_\bullet = (i_N, \ldots, i_2, i_1)$ such that $w_0 = s_{i_N} \cdots s_{i_2} s_{i_1}$ with N minimal. Say that $i_\bullet \in \mathrm{Red}$ is adapted to $Q \in \mathrm{Or}(X)$ and write $i_\bullet \in \mathrm{Red}_Q$, if for every j the vertex i_j is a sink in the directed graph $s_{i_{j-1}} \cdots s_{i_2} s_{i_1} Q$. For every orientation Q of X, $\mathrm{Red}_Q \neq \emptyset$. Moreover Red_Q is a commutation class (two reduced words are in the same commutation class, if they are reachable from each other by commuting Coxeter relations $s_i s_j = s_j s_i$ where i and j are nonadjacent vertices in X). However $\bigcup_{Q\in \mathrm{Or}(X)} \mathrm{Red}_Q \subsetneq \mathrm{Red}$.

Fix $i_\bullet \in \mathrm{Red}_Q$. It defines a total ordering $\leq_{i_\bullet}$ on the set of positive roots Φ^+ by the sequence $\gamma_1 < \gamma_2 < \ldots$ where $\gamma_j = s_{i_j} \cdots s_{i_2} \alpha_{i_1}$. Then it is a theorem of [**2**] that the total orders $\leq_{i_\bullet}$ for $i_\bullet \in \mathrm{Red}_Q$, are the set of linear extensions of the partial order $\preccurlyeq_Q$.

The Auslander-Reiten quiver Γ_Q of $Q \in \mathrm{Or}(X)$ is traditionally drawn with arrows going from right to left and smaller elements pointing towards bigger ones. It turns out that there is a nice planar embedding of Γ_Q such that the poset element γ_j is placed in the i_j-th row for all j. Even better, this graph is the 1-skeleton of a topological complex [**31**].

Example 5.6. Let Q be the A_3 quiver with both arrows pointing to the middle:

$$0 \longrightarrow 1 \longleftarrow 2$$

We use $i_\bullet = (2,0,1,2,0,1) \in \mathrm{Red}_Q$. The induced total ordering $\leq_{i_\bullet}$ on Ind_Q is given below, with labeling of Φ^+ as in Example 5.2.

$$\alpha_{11} < \alpha_{01} < \alpha_{12} < \alpha_{02} < \alpha_{22} < \alpha_{00}. \tag{5.3}$$

The Auslander-Reiten quiver Γ_Q is depicted below.

$$\begin{tikzcd}
0 & & I_{01} \arrow[dr] & & I_{22} \\
1 & I_{11} \arrow[ur] \arrow[dr] & & I_{02} \arrow[ur] \arrow[dr] & \\
2 & & I_{12} \arrow[ur] & & I_{00}
\end{tikzcd}$$

The matrix for the Euler form on pairs of elements of Φ^+ with respect to the total order (5.3) is given by

$$(\langle \alpha\,,\, \beta \rangle)_{\alpha,\beta\in\Phi^+} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \end{pmatrix}.$$

Example 5.7. Let Q be the D_4 quiver with the following orientation:

$$\begin{tikzcd}
 & 1 \arrow[d] & \\
4 & 2 \arrow[l] \arrow[r] & 3
\end{tikzcd}$$

We use $w_0 = s_1 s_2 s_4 s_3 s_1 s_2 s_4 s_3 s_1 s_2 s_4 s_3$. We label the indecomposables by their dimension vectors. For example, 1211 means $(1,2,1,1)$ or $\alpha_1 + 2\alpha_2 + \alpha_3 + \alpha_4$. The total ordering for the above reduced word is given by the list

$$0010, 0001, 0111, 1111, 0101, 0110, 1211, 0100, 1110, 1101, 1100, 1000 \tag{5.4}$$

and the AR quiver is given by

$$\begin{tikzcd}
1 & & & 1111 \arrow[dr] & & 0100 \arrow[dr] & & 1000 \\
2 & & 0111 \arrow[ur] \arrow[dr] \arrow[ddr] & & 1211 \arrow[ur, leftrightarrow] \arrow[dr] \arrow[ddr] & & 1100 \arrow[ur] & \\
3 & 0010 \arrow[ur] & & 0101 \arrow[ur] & & 1110 \arrow[ur] & & \\
4 & 0001 \arrow[uur] & & 0110 \arrow[uur] & & 1101 \arrow[uur] & &
\end{tikzcd}$$

5.6. Orbit representatives. Let Q be a Dynkin quiver, $d : Q_0 \to \mathbb{Z}_{\geq 0}$ a dimension vector and m a Kostant partition of d. For our divided difference formula for quiver polynomials we define a representative element $\phi_m \in \mathrm{Rep}(Q,d)$ in the $G = G(Q,d)$-orbit indexed by m, that is, $\phi_m \cong \bigoplus_{\alpha\in\Phi^+} I_\alpha^{\oplus m(\alpha)}$.

Pick any particular matrix representation for each indecomposable I_α and by abuse of notation denote it by I_α. Consider an *ordered* direct sum $I_\bullet = I_1 \oplus I_2 \oplus \cdots \oplus I_M$ that has $m(\alpha)$ summands I_α for all $\alpha \in \Phi^+$, with the property that

$$\mathrm{Ext}^1_Q(I_j, I_i) = 0 \qquad \text{if } i < j. \tag{5.5}$$

This condition holds if we list the indecomposables in the *reverse* of the total order on Ind_Q given by $\leq_{i_\bullet}$ for any $i_\bullet \in \mathrm{Red}_Q$.

We view $I_\bullet$ as a point in $\mathrm{Rep}(Q,d)$. As such $I_\bullet$ is "block diagonal": for each $a \in Q_1$ the a-th component of $I_\bullet$ is "block diagonal" with "diagonal" blocks given by the a-th components of $I_1, I_2, \ldots, I_M$ in that order.

EXAMPLE 5.8. Take Q to be the equioriented A_2 quiver, $d(0) = e$ and $d(1) = f$. Take the quiver locus X_r given by the determinantal variety of $e \times f$ matrices of rank at most r. Then the $G = GL(e) \times GL(f)$-orbit associated to X_r has Kostant partition m with $m(\alpha_0) = e - r$, $m(\alpha_0 + \alpha_1) = r$, and $m(\alpha_1) = f - r$. So if $e = 3$, $f = 4$, and $r = 2$ then an appropriate ordering of the indecomposables in $I_\bullet$ is given by $I_\bullet = I_{\alpha_0} \oplus I_{\alpha_0+\alpha_1} \oplus I_{\alpha_0+\alpha_1} \oplus I_{\alpha_1} \oplus I_{\alpha_1}$. The element $I_\bullet \in M_{3\times 4}$ is the matrix

$$I_\bullet = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

where each I_{α_0} is a 1×0 matrix, each $I_{\alpha_0+\alpha_1}$ is a 1×1 identity matrix, and each I_{α_1} is a 0×1 matrix.

Fix $I_\bullet$ as above. Define the Levi subgroup $L(I_\bullet) \subset G(Q,d)$ by

$$L(I_\bullet) = \prod_{k=1}^{M} G(Q, d(I_k)).$$

We regard $L(I_\bullet)$ as a block diagonal subgroup of $G = G(Q,d)$: for each $i \in Q_0$, the i-th component of $L(I_\bullet)$ are the block diagonal matrices in the i-th component of $G(Q,d)$ with block sizes coming from the i-th components of $G(Q, d(I_k))$. It acts on the direct product

$$\mathrm{Rep}(I_\bullet) = \prod_{k=1}^{M} \mathrm{Rep}(Q, d(I_k)). \tag{5.6}$$

We regard $\mathrm{Rep}(I_\bullet) \subset \mathrm{Rep}$ similarly as the "block diagonal" elements of Rep.

If I is an indecomposable $\mathbb{C}Q$-module with dimension vector d_I, then it is easy to show for our situation that $\overline{G(Q,d_I)\cdot I} = \mathrm{Rep}(Q, d_I)$. It follows that

$$\mathrm{Rep}(I_\bullet) = \overline{L(I_\bullet)\cdot I_\bullet}. \tag{5.7}$$

Let $P(I_\bullet) \subset G$ be the parabolic subgroup given by the lower block triangular subgroup of G with Levi factor $L(I_\bullet)$. For $i \in Q_0$ its i-th component is lower block triangular with diagonal blocks given by those of the i-th component of $L(I_\bullet)$.

Finally, let $Z(I_\bullet) \subset \mathrm{Rep}$ be the "lower block triangular" coordinate subspace of Rep, such that for $a \in Q_1$ the a-th component of $Z(I_\bullet)$ consists of the matrices with zeroes in the entries strictly above the "block diagonal" given by the a-th component of $\mathrm{Rep}(I_\bullet)$ and arbitrary entries allowed elsewhere.

LEMMA 5.9.

$$Z(I_\bullet) = \overline{P(I_\bullet)\cdot I_\bullet}. \tag{5.8}$$

The proof of this fact, which is equivalent to the condition (5.5), follows easily from the definition of Ext. This is precisely the point where the careful ordering of the indecomposables in $I_\bullet$ is used.

EXAMPLE 5.10. For $I_\bullet$ as in Example 5.8,

$$L(I_\bullet) = T_3 \times T_4 \subset GL(3) \times GL(4)$$
$$P(I_\bullet) = B_- \times B_- \subset GL(3) \times GL(4).$$

So $L(I_\bullet)$ is the maximal torus and $P(I_\bullet)$ is the product of lower triangular Borels. (This always holds in type A: each positive root α_{ij} contains at most one copy of each simple root). $\mathrm{Rep}(I_\bullet)$ and $Z(I_\bullet)$ are the coordinate subspaces of Rep given by

$$\mathrm{Rep}(I_\bullet) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \end{pmatrix} \qquad Z(I_\bullet) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 \\ * & * & 0 & 0 \end{pmatrix}. \tag{5.9}$$

6. Kempf collapsings

The geometric construction of a Kempf collapsing κ leads to divided difference formulae for the equivariant cohomology class and K-theory class of the image of κ. We recall Reineke's construction, which realizes a Dynkin quiver locus as the image of a Kempf collapsing. This yields divided difference formula for quiver and K-quiver polynomials.

Let G be a reductive algebraic group over $\mathbb{C}$ and P a parabolic subgroup. Let Y be a finite-dimensional G-module and $Z \subset Y$ a P-stable closed subscheme. In our application G is the product of general linear groups of the form $G(Q,d)$ and Z is a linear subspace of $\mathrm{Rep}(Q,d)$. Consider the G-equivariant fiber bundle $G \times^P Z$ over the partial flag variety G/P with fiber Z over the identity:

$$G \times^P Z = (G \times Z)/P.$$

Here P acts diagonally on the right by $(g,z)p = (gp, p^{-1} \cdot z)$. Consider the map

$$\begin{aligned} \kappa : G \times^P Z &\to Y \\ (g,z)P &\mapsto gz. \end{aligned}$$

We call κ a **Kempf collapsing**. The map κ is proper so its image is closed.

Theorem 6.1. [**21**] *Suppose that*

- *Z has rational singularities.*
- *$\mathcal{O}_Y \to \kappa_* \mathcal{O}_{G\times^P Z}$ is surjective.*
- *$R^j \kappa_* \mathcal{O}_{G\times^P Z} = 0$ for $j > 0$.*

Then Im κ *is normal and Cohen-Macaulay. If in addition κ is birational to its image, then* Im κ *has rational singularities.*

Kempf suggests a condition to guarantee these criteria: that Z is a linear subspace and a completely reducible P-module. In our application the latter condition doesn't hold so we don't assume it. Here is our extension of Kempf's result.

Theorem 6.2. *Suppose Z has rational singularities and $R^j \kappa_* \mathcal{O}_{G\times^P Z} = 0$ for $j > 0$. Let $\widetilde{\mathrm{Im}\ \kappa}$ be the normalization of the image of κ.*

- *If the general fiber of κ is connected, then $\widetilde{\mathrm{Im}\ \kappa}$ has rational singularities.*
- *If the general fiber of κ is connected and* Im κ *is normal (hence has rational singularities), then $\kappa_* \mathcal{O}_{G\times^P Z} = \mathcal{O}_{\mathrm{Im}\ \kappa}$.*
- *Conversely, if $\kappa_* \mathcal{O}_{G\times^P Z} = \mathcal{O}_{\mathrm{Im}\ \kappa}$, then all fibers of κ are connected, and* Im κ *is normal (hence has rational singularities).*

Even without the last two conditions, the Kempf collapsing still determines the multidegree of Im κ.

Theorem 6.3. *Suppose Z has rational singularities. Let $m_0 = H_Z(Y)$. Construct a sequence of polynomials $m_1, m_2, \ldots$ where each polynomial is obtained from the previous one by a divided difference operator $\partial_\alpha = \frac{1}{\alpha}(1 - s_\alpha)$, where α varies over the set of simple roots of G (taken in the Borel opposite to one contained in P), and the action of s_α on $\mathrm{Sym}_{\mathbb{Z}}(X(T))$ is induced from the reflection action on $X(T)$. Don't apply a divided difference operator if the result is 0, and only stop when all ∂_α give the result 0. This process always terminates after the same number of steps, and the last polynomial in this sequence is c times $H_Y(\mathrm{Im}\ \kappa)$, where c is the number of components in a general fiber of κ.*

When we have both connected fibers and the vanishing of higher direct images of κ, then we can compute the Hilbert numerator $K_Y(\widetilde{\mathrm{Im}\ \kappa})$.

THEOREM 6.4. *Suppose Z has at worst rational singularities, the general fiber of κ is connected, and $R^j\kappa_*\mathcal{O}_{G\times^P Z} = 0$ for $j > 0$.*

Let $m_0 = K_Y(Z)$, and construct a sequence of Laurent polynomials $m_1, m_2, \ldots$ by applying Demazure operators $\pi_\alpha := (1-\exp(-\alpha))^{-1}(1-\exp(\alpha)s_\alpha)$ to m_0, where α varies over the set of simple roots of G. Stop when the application of any π_α leaves the result unchanged. This process terminates after finitely many steps. The last Laurent polynomial in this sequence is $K_Y(X)$ where X is the pushforward of $\mathcal{O}_{\widetilde{\mathrm{Im}\ \kappa}}$ under the normalization map $\widetilde{\mathrm{Im}\ \kappa} \to \mathrm{Im}\ \kappa \to Y$.

Explicitly,

$$K_Y(X) = \sum_{w\in W} w \cdot \frac{K_Y(Z)}{\prod_{\beta\in\Phi^+}(1-\exp(-\beta))}$$

where Φ^+ is the set of positive roots relative to the opposite of some Borel subgroup between T and P.

REMARK 6.5. In Theorems 6.3 and 6.4, let w_0 be the longest element in the Weyl group W. One may take a reduced word for w_0 and apply the sequence of divided differences indicated by the reduced word. In cohomology one should skip an operator if its result is zero.

The general machine of Kempf collapsings and divided differences may be applied to quiver loci via Reineke's construction [**28**].

THEOREM 6.6. *Suppose Q is a Dynkin quiver and d is any dimension vector. Then each orbit closure $\Omega \subseteq \mathrm{Rep}(Q,d)$ is the image of a linear Kempf collapsing, i.e. there exists a parabolic subgroup $P \subset G$ and a P-invariant linear subspace $Z \subset \mathrm{Rep}$ such that $\Omega = G\cdot Z$.*

By Lemma 5.9 a suitable choice for P and Z is given by $P(I_\bullet)$ and $Z(I_\bullet)$ where $I_\bullet$ is chosen as in section 5.6. Then one may use the product formulae (3.3) for the starting element of the divided difference formulae and apply divided differences to get the desired quiver or K-quiver polynomial.

EXAMPLE 6.7. Continuing Examples 5.8 and 5.10, let $x_1, x_2, x_3, y_1, y_2, y_3, y_4$ be the standard basis of $X(T)$ where $T \subset GL(3)\times GL(4)$ is the maximal torus. $[Z(I_\bullet)]_T$ is the product of linear forms $(x_i - y_j)$ where (i,j) runs over the positions in $M_{3\times 4}$ where $Z(I_\bullet)$ contains a zero entry. We recover Example 2.2:

$$\begin{aligned}
[Z(I_\bullet)]_T &= (x_1-y_1)(x_1-y_2)(x_1-y_3)(x_1-y_4)(x_2-y_2)(x_2-y_3)(x_2-y_4)\\
&\quad\times(x_3-y_3)(x_3-y_4)\\
[X_2]_G &= \partial^y_2\partial^y_1\partial^y_3\partial^y_2\partial^x_1\partial^x_2\partial^x_1[Z(I_\bullet)]_T\\
&= s_2[x-y]
\end{aligned}$$

where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3, y_4)$. Note that two divided difference operators must be omitted from a reduced decomposition of the longest element of $W(G(Q,d)) = S_3 \times S_4$.

EXAMPLE 6.8. Let Q be the type D4 quiver in Example 5.7, $d = (2,3,2,2)$, T the maximal torus in $G(Q,d)$, and let $X(T)$ have basis a_1, a_2, b_1, b_2, b_3, c_1, c_2, d_1, d_2. Consider

$$I_\bullet = I_{(1,1,0,1)} \bigoplus I_{(1,1,1,0)} \bigoplus I_{(0,1,1,1)},$$

ordering terms as in the reverse of the total order (5.4) on Ind_Q. Then $P(I_\bullet) = B$ consists of the product $\prod_{i\in Q_0} B^i$ of lower triangular subgroups $B^i \subset GL(d(i))$. Let $z^1 \in M_{2\times 3}$, $z^2 \in M_{3\times 2}$, and $z^3 \in M_{3\times 2}$ be the matrices corresponding to the arrows $(1,2)$, $(2,3)$, and $(2,4)$ in Q_1 respectively. Then the point $I_\bullet$ and the subspace $Z = Z(I_\bullet)$ are given by

$$I^1_\bullet = \begin{pmatrix}1&0&0\\0&1&0\end{pmatrix} \qquad I^2_\bullet = \begin{pmatrix}0&0\\1&0\\0&1\end{pmatrix} \qquad I^3_\bullet = \begin{pmatrix}1&0\\0&0\\0&1\end{pmatrix}$$

$$z^1 = \begin{pmatrix}*&0&0\\ *&*&0\end{pmatrix} \qquad z^2 = \begin{pmatrix}0&0\\ *&0\\ *&*\end{pmatrix} \qquad z^3 = \begin{pmatrix}*&0\\ *&0\\ *&*\end{pmatrix}$$

$Z(I_\bullet)$ has the equations $z^1_{12} = z^1_{13} = z^1_{23} = 0$, $z^2_{11} = z^2_{12} = z^2_{22} = 0$, and $z^3_{12} = z^3_{22} = 0$. To compute the multidegree of the corresponding quiver locus Ω, we start with the multidegree $H_{\mathrm{Rep}}(Z) = (a_1 -$

$b_2)(a_1 - b_3)(a_2 - b_3)(b_1 - c_1)(b_1 - c_2)(b_2 - c_2)(b_1 - d_2)(b_2 - d_2)$. Applying ∂_1^a, ∂_1^c, and ∂_1^d we obtain $(a_1 - b_3)(a_2 - b_3)(b_1 - c_1)(b_1 - c_2)(b_1 + b_2 - d_1 - d_2)$. Applying $\partial_1^b \partial_2^b \partial_1^b$ we obtain the answer

$$H_{\mathrm{Rep}}(\Omega) = s_{11}[a - b] + s_1[a - b]s_1[b - c] + s_1[a - b]s_1[b - d] + s_1[b - c]s_1[b - d]$$

where $s_\lambda[X - Y]$ is the double Schur polynomial. Note how the answer can be expressed as a positive sum of products of double Schur polynomials in differences of sets of variables, where the differences correspond to arrows in Q_1. This seems to be an instance of a Schur or component formula (a la [**23**]) in type D.

7. Future directions

We believe that the method of Kempf collapsing yields divided difference formulae for a nontrivial family of quiver loci for any quiver Q. In Remark 4.1 it was explained how one may define multidegrees and Hilbert numerators for an arbitrary Q but with a condition on the quiver locus. Under those conditions, consider Example 4.2 consisting of $M_{n\times n}$ under the adjoint action of $G = GL(n)$ and in particular the closure X of a nilpotent conjugacy class. Then X has a Kempf collapsing [**30**], but the best choice of the space Z is not the direct sum of the indecomposables as in the Dynkin case. One may choose Z to be the set of matrices that are strictly lower block triangular with diagonal block of sizes given by the transpose of the partition coming from the nilpotent Jordan blocks, and P to be the lower block triangular parabolic for the same set of diagonal blocks.

As indicated in the introduction, it would also be nice to obtain "positive" formulae for the (K-)quiver polynomials.

References

[1] M. Auslander, I. Reiten, and S. O. Smalø, Representation theory of Artin algebras, Cambridge Stud. Adv. Math. 36, Cambridge University Press, 1995.

[2] R. Bedard, On commutation classes of reduced words in Weyl groups, European J. Combin. **20** (1999), no. 6, 483–505.

[3] N. Bergeron and S. Billey, RC-graphs and Schubert polynomials, Experiment. Math. **2** (1993), no. 4, 257–269.

[4] A. S. Buch and W. Fulton, Chern class formulas for quiver varieties, Invent. Math. **135** (1999), no. 3, 665–687.

[5] A. S. Buch, L. M. Fehér, and R. Rimányi, Positivity of quiver coefficients through Thom polynomials, Adv. Math. 197 (2005), no. 1, 306–320.

[6] I. N. Bernstein, I. M. Gelfand, and V. A. Ponomarev, Coxeter functors and Gabriel's theorem, Uspehi Mat. Nauk **28** (1973), no. 2 (170), 19–33.

[7] A. S. Buch, A. Kresch, H. Tamvakis, M. Shimozono, and A. Yong, Stable Grothendieck polynomials and K-theoretic factor sequences, preprint.

[8] A. S. Buch, A. Kresch, H. Tamvakis, and A. Yong, Schubert polynomials and quiver formulas, Duke Math. J. **122** (2004), no. 1, 125–143.

[9] A. S. Buch, A. Kresch, H. Tamvakis, and A. Yong, Grothendieck polynomials and quiver formulas, Amer. J. Math. **127** (2005), no. 3, 551–567.

[10] A. S. Buch and R. Rimanyi, A formula for non-equioriented quiver orbits of type A, arXiv:math.AG/0412073.

[11] A. S. Buch, A Littlewood-Richardson rule for the K-theory of Grassmannians. Acta Math. **189** (2002), no. 1, 37–78.

[12] A. S. Buch, Grothendieck classes of quiver varieties. Duke Math. J. 115 (2002), no. 1, 75–103.

[13] A. S. Buch, Alternating signs of quiver coefficients, J. Amer. Math. Soc. **18** (2005), no. 1, 217–237.

[14] H. Derksen, Lecture notes on quiver representations, http://www.math.lsa.umich.edu/%7Ehderksen/math711.w01/math711.html.

[15] H. Derksen and J. Weyman, Quiver representations, Notices Amer. Math. Soc. **52** (2005), no. 2, 200–206.

[16] S. V. Fomin, S. Gelfand, and A. Postnikov, Quantum Schubert polynomials, JAMS **10** (1997) 565–596.

[17] S. V. Fomin and A. N. Kirillov, The Yang-Baxter equation, symmetric functions, and Schubert polynomials. Proceedings of the 5th Conference on Formal Power Series and Algebraic Combinatorics (Florence, 1993). Discrete Math. **153** (1996), no. 1-3, 123–143.

[18] W. Fulton, Flags, Schubert polynomials, degeneracy loci, and determinantal formulas, Duke Math. J. **65** (1992) 381–420.

[19] W. Fulton, Universal Schubert polynomials, Duke Math. J. **96** (1999) 575–594.

[20] P. Gabriel, Unzerlegbare Darstellungen, I, Manuscripta Math. **6** (1972), 71–103; correction, ibid. **6** (1972), 309.

[21] G. Kempf, On the collapsing of homogeneous bundles. Invent. Math. **37** (1976) 229–239.

[22] A. Knutson and E. Miller, Gröbner geometry of Schubert polynomials, to appear in Ann. of Math.

[23] A. Knutson, E. Miller, and M. Shimozono, Four positive formulae for type A quiver polynomials, to appear in Invent. Math.

[24] E. Miller, Alternating formulas for K-theoretic quiver polynomials, Duke Math. J. **128** (2005), no. 1, 1–17.

[25] E. Miller and B. Sturmfels, Combinatorial Commutative Algebra, Graduate Texts in Mathematics, 227. Springer-Verlag, New York, 2005.

[26] I. R. Porteous, Simple singularities of maps, Proceedings of Liverpool Singularities Symposium, I (1969/70), 286–307, Lecture Notes in Math., Vol. 192, Springer, Berlin, 1971.

[27] C. M. Ringel, Representations of K-species and bimodules, J. Algebra **41** (1976) 269–302.

[28] M. Reineke, Quivers, desingularizations and canonical bases, Studies in memory of Issai Schur (Chevaleret/Rehovot, 2000), 325–344, Progr. Math., 210, Birkhäuser Boston, Boston, MA, 2003.

[29] R. Thom, Les singularités des applications différentiables, Ann. Inst. Fourier, Grenoble **6** (1955–1956), 43–87.

[30] J. Weyman, Cohomology of vector bundles and syzygies, Cambridge Tracts in Mathematics, 149, Cambridge University Press, Cambridge, 2003.

[31] S. Zelikson, Auslander-Reiten quivers and the Coxeter complex, Algebr. Represent. Theory **8** (2005), no. 1, 35–55.

Department of Mathematics, University of California, San Diego
E-mail address: `allenk@math.ucsd.edu`

Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123 USA
E-mail address: `mshimo@vt.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Schur positivity and Cell Transfer

Thomas Lam, Alexander Postnikov, and Pavlo Pylyavskyy

Abstract. We give combinatorial proofs that certain families of differences of products of Schur functions are monomial-positive. We show in addition that such monomial-positivity is to be expected of a large class of generating functions with combinatorial definitions similar to Schur functions. These generating functions are defined on posets with labelled Hasse diagrams and include for example generating functions of Stanley's (P,ω)-partitions. Then we prove Okounkov's conjecture, a conjecture of Fomin-Fulton-Li-Poon, and a special case of Lascoux-Leclerc-Thibon's conjecture on Schur positivity and give several more general statements using a recent result of Rhoades and Skandera. An alternative proof of this result is provided. We also give an intriguing log-concavity property of Schur functions. This text contains the material from [**LP, LPP**].

Résumé. Nous prouvons combinatoirement que certaines familles de différences de produits de fonctions de Schur sont monomiales-positives. Nous montrons de plus que l'on peut attendre une telle propriété pour une importante classe de fonctions génératrices définies combinatoirement d'une façon similaire aux fonctions de Schur. Ces fonctions génératrices sont définies en termes d'ensembles partiellement ordonnés dont le diagramme de Hasse est étiqueté et comprennent par exemple la fonction génératrice des (P,ω)-partitions de Stanley. Nous prouvons aussi la conjecture d'Okounkov, une conjecture de Fomin-Fulton-Li-Poon, et un cas particulier de la conjecture de Lascoux-Leclerc-Thibon sur la positivité de Schur, et nous donnons plusieurs énoncés plus généraux en utilisant un résultat récent de Rhoades et Skandera. Nous donnons aussi une nouvelle preuve de ce résultat et une propriété surprenante de log-concavité des fonctions de Schur.

1. Schur positivity conjectures

The Schur functions s_λ form an orthonormal basis of the ring of symmetric functions Λ. They have a remarkable number of combinatorial and algebraic properties, and are simultaneously the irreducible characters of $GL(N)$ and representatives of Schubert classes in the cohomology $H^*(Gr_{kn})$ of the Grassmannian; see [**Mac, Sta**]. In recent years, a lot of work has gone into studying whether certain expressions of the form

$$s_\lambda s_\mu - s_\nu s_\rho \tag{1.1}$$

The first aim of this article is to provide a large class of expressions of the form (1.1) which are *monomial-positive*, that is, expressible as a non-negative linear combination of monomials. In particular, we show that (1.1) is monomial-positive when $\lambda = \nu \vee \rho$ and $\mu = \nu \wedge \rho$ are the union and intersections of the Young diagrams of ν and ρ. However, we show in addition that such monomial-positivity is to be expected of many families of generating functions with combinatorial definitions similar to Schur functions, which are generating functions for semistandard Young tableaux.

We define a new combinatorial object called a $\mathbb{T}$-*labelled poset* and given a $\mathbb{T}$-labelled poset (P,O) we define another combinatorial object which we call (P,O)-*tableaux*. These (P,O)-tableaux include as special cases standard Young tableaux, semistandard Young tableaux, cylindric tableaux, plane partitions, and Stanley's (P,ω)-partitions. Our main theorem is the *cell transfer theorem*. It says that for a fixed $\mathbb{T}$-labelled poset (P,O), one obtains many expressions of the form (1.1) which are monomial-positive, where the Schur functions in (1.1) are replaced by generating functions for (P,O)-tableaux.

Key words and phrases. Schur functions, Schur positivity, Schur log-concavity, immanants, Kazhdan-Lusztig polynomials, Temperley-Lieb algebra, minors.

A.P. was supported in part by NSF grant DMS-0201494. P.P. would like to thank Institut Mittag-Leffler for their hospitality.

A symmetric function is called *Schur nonnegative* if it is a linear combination with nonnegative coefficients of the Schur functions, or, equivalently, if it is the character of a certain representation of GL_n. In particular, *skew Schur functions* $s_{\lambda/\mu}$ are Schur nonnegative. We prove that our cell-transfer results for Schur functions hold not just for monomial-positivity but also for Schur-positivity. In particular, we prove the following theorem.

For two partitions $\lambda = (\lambda_1, \lambda_2, \dots)$ and $\mu = (\mu_1, \mu_2, \dots)$, let us define partitions

$$\lambda \vee \mu := (\max(\lambda_1, \mu_1), \max(\lambda_2, \mu_2), \dots)$$

and

$$\lambda \wedge \mu := (\min(\lambda_1, \mu_1), \min(\lambda_2, \mu_2), \dots).$$

The Young diagram of $\lambda \vee \mu$ is the set-theoretical union of the Young diagrams of λ and μ. Similarly, the Young diagram of $\lambda \wedge \mu$ is the set-theoretical intersection of the Young diagrams of λ and μ. For two skew shapes, define $(\lambda/\mu) \vee (\nu/\rho) := \lambda \vee \nu / \mu \vee \rho$ and $(\lambda/\mu) \wedge (\nu/\rho) := \lambda \wedge \nu / \mu \wedge \rho$.

Theorem 1.1. *Let λ/μ and ν/ρ be any two skew shapes. Then we have*

$$s_{(\lambda/\mu)\vee(\nu/\rho)}\, s_{(\lambda/\mu)\wedge(\nu/\rho)} \geq_s s_{\lambda/\mu}\, s_{\nu/\rho}.$$

Using this theorem, we prove the following several Schur positivity conjectures due to Okounkov, Fomin-Fulton-Li-Poon, and Lascoux-Leclerc-Thibon.

Okounkov [**Oko**] studied branching rules for classical Lie groups and proved that the multiplicities were "monomial log-concave" in some sense. An essential combinatorial ingredient in his construction was the theorem about monomial nonnegativity of some symmetric functions. He conjectured that these functions are Schur nonnegative, as well. For a partition λ with all even parts, let $\frac{\lambda}{2}$ denote the partition $(\frac{\lambda_1}{2}, \frac{\lambda_2}{2}, \ldots)$. For two symmetric functions f and g, the notation $f \geq_s g$ means that $f - g$ is Schur nonnegative.

Conjecture 1.2. Okounkov [**Oko**] *For two skew shapes λ/μ and ν/ρ such that $\lambda + \nu$ and $\mu + \rho$ both have all even parts, we have $(s_{\frac{(\lambda+\nu)}{2}/\frac{(\mu+\rho)}{2}})^2 \geq_s s_{\lambda/\mu}\, s_{\nu/\rho}$.*

Fomin, Fulton, Li, and Poon [**FFLP**] studied the eigenvalues and singular values of sums of Hermitian and of complex matrices. Their study led to two combinatorial conjectures concerning differences of products of Schur functions. Let us formulate one of these conjectures, which was also studied recently by Bergeron and McNamara [**BM**]. For two partitions λ and μ, let $\lambda \cup \mu = (\nu_1, \nu_2, \nu_3, \dots)$ be the partition obtained by rearranging all parts of λ and μ in the weakly decreasing order. Let $\mathrm{sort}_1(\lambda, \mu) := (\nu_1, \nu_3, \nu_5, \dots)$ and $\mathrm{sort}_2(\lambda, \mu) := (\nu_2, \nu_4, \nu_6, \dots)$.

Conjecture 1.3. Fomin-Fulton-Li-Poon [**FFLP**, Conjecture 2.7] *For two partitions λ and μ, we have $s_{\mathrm{sort}_1(\lambda,\mu)} s_{\mathrm{sort}_2(\lambda,\mu)} \geq_s s_\lambda s_\mu$.*

Lascoux, Leclerc, and Thibon [**LLT**] studied a family of symmetric functions $\mathcal{G}_\lambda^{(n)}(q, x)$ arising combinatorially from ribbon tableaux and algebraically from the Fock space representations of the quantum affine algebra $U_q(\widehat{\mathfrak{sl}}_n)$. They conjectured that $\mathcal{G}_{n\lambda}^{(n)}(q, x) \geq_s \mathcal{G}_{m\lambda}^{(m)}(q, x)$ for $m \leq n$. For the case $q = 1$, their conjecture can be reformulated, as follows. For a partition λ and $1 \leq i \leq n$, let $\lambda^{[i,n]} := (\lambda_i, \lambda_{i+n}, \lambda_{i+2n}, \dots)$. In particular, $\mathrm{sort}_i(\lambda, \mu) = (\lambda \cup \mu)^{[i,2]}$, for $i = 1, 2$.

Conjecture 1.4. Lascoux-Leclerc-Thibon [**LLT**, Conjecture 6.4] *For integers $1 \leq m \leq n$ and a partition λ, we have $\prod_{i=1}^n s_{\lambda^{[i,n]}} \geq_s \prod_{i=1}^m s_{\lambda^{[i,m]}}$.*

Theorem 1.5. *Conjectures 1.2, 1.3 and 1.4 are true.*

In Section 6, we present and prove more general versions of these conjectures.

2. Posets and Tableaux

Let $(P, \leq)$ be a possibly infinite poset. Let $s, t \in P$. We say that s covers t and write $s \gtrdot t$ if for any $r \in P$ such that $s \geq r \geq t$ we have $r = s$ or $r = t$. The *Hasse diagram* of a poset P is the graph with vertex set equal to the elements of P and edge set equal to the set of covering relations in P. If $Q \subset P$ is a subset of the elements of P then Q has a natural induced subposet structure. If $s, t \in Q$ then $s \leq t$ in Q if and only if $s \leq t$ in P. Call a subset $Q \subset P$ *connected* if the elements in Q induce a connected subgraph in the Hasse diagram of P.

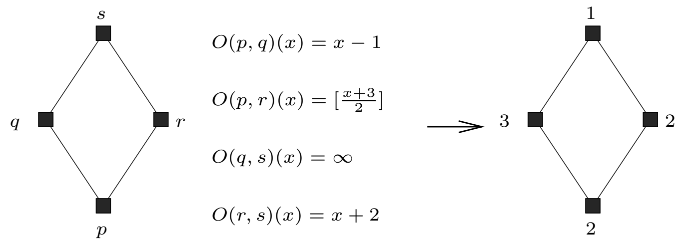


FIGURE 1. An example of a $\mathbb{T}$-labelled poset (P,O) and a (P,O)-tableaux.

An *order ideal* I of P is an induced subposet of P such that if $s \in I$ and $s \geq t \in P$ then $t \in I$. A subposet $Q \subset P$ is called *convex* if for any $s, t \in Q$ and $r \in P$ satisfying $s \leq r \leq t$ we have $r \in Q$. Alternatively, a convex subposet is one which is closed under taking intervals. A convex subset Q is determined by specifying two order ideals J and I so that $J \subset I$ and $Q = \{s \in I \mid s \notin J\}$. We write $Q = I/J$. If $s \notin Q$ then we write $s < Q$ if $s < t$ for some $t \in Q$ and similarly for $s > Q$. If $s \in Q$ or s is incomparable with all elements in Q we write $s \sim Q$. Thus for any $s \in P$, exactly one of $s < Q$, $s > Q$ and $s \sim Q$ is true.

Let $\mathbb{P}$ denote the set of positive integers and $\mathbb{Z}$ denote the set of integers. Let $\mathbb{T}$ denote the set of all weakly increasing functions $f : \mathbb{P} \to \mathbb{Z} \cup \{\infty\}$.

DEFINITION 2.1. A $\mathbb{T}$-*labelling* O of a poset P is a map $O : \{(s,t) \in P^2 \mid s \gtrdot t\} \to \mathbb{T}$ labelling each edge (s,t) of the Hasse diagram by a weakly increasing function $O(s,t) : \mathbb{P} \to \mathbb{Z} \cup \{\infty\}$. A $\mathbb{T}$-*labelled poset* is an an ordered pair (P,O) where P is a poset, and O is a $\mathbb{T}$-labelling of P.

We shall refer to a $\mathbb{T}$-labelled poset (P,O) as P when no ambiguity arises. If $Q \subset P$ is a convex subposet of P then the covering relations of Q are also covering relations in P. Thus a $\mathbb{T}$-labelling O of P naturally induces a $\mathbb{T}$-labelling $O|_Q$ of Q. We denote the resulting $\mathbb{T}$-labelled poset by $(Q,O) := (Q, O|_Q)$.

DEFINITION 2.2. A (P,O)-tableau is a map $\sigma : P \to \mathbb{P}$ such that for each covering relation $s \lessdot t$ in P we have

$$\sigma(s) \leq O(s,t)(\sigma(t)).$$

If $\sigma : P \to \mathbb{P}$ is any map, then we say that σ *respects* O if σ is a (P,O)-tableau.

Figure 1 contains an example of a $\mathbb{T}$-labelled poset (P,O) and a corresponding (P,O)-tableau.

Denote by $\mathcal{A}(P,O)$ the set of all (P,O)-tableaux. If P is finite then one can define the formal power series $K_{P,O}(x_1, x_2, \ldots) \in \mathbb{Q}[[x_1, x_2, \ldots]]$ by

$$K_{P,O}(x_1, x_2, \ldots) = \sum_{\sigma \in \mathcal{A}(P,O)} x_1^{\#\sigma^{-1}(1)} x_2^{\#\sigma^{-1}(2)} \cdots .$$

The composition $\mathrm{wt}(\sigma) = (\#\sigma^{-1}(1), \#\sigma^{-1}(2), \ldots)$ is called the weight of σ.

EXAMPLE 2.3. Any Young diagram $P = \lambda$ can be considered as a $\mathbb{T}$-labelled poset. Indeed, consider its cells to be elements of the poset, and let O be the labelling of the horizontal edges with the function $f^{\mathrm{weak}}(x) = x$ and label the vertical edges with the function $f^{\mathrm{strict}}(x) = x - 1$. A (λ, O)-tableau is just a semistandard Young tableaux and $K_{\lambda,O}(x_1, x_2, \cdots)$ is the Schur function $s_\lambda(x_1, x_2, \cdots)$.

EXAMPLE 2.4. Another interesting example are cylindric tableaux and cylindric Schur functions. Let $1 \leq k < n$ be two positive integers. Let $\mathbb{C}_{k,n}$ be the quotient of $\mathbb{Z}^2$ given by

$$\mathbb{C}_{k,n} = \mathbb{Z}^2/(k-n,k)/Z.$$

In other words, the integer points (a,b) and $(a+k-n, b+k)$ are identified in $\mathbb{C}_{k,n}$. We can give $\mathbb{C}_{k,n}$ the structure of a poset by the generating relations $(i,j) \lessdot (i+1,j)$ and $(i,j) \lessdot (i,j+1)$. We give $\mathbb{C}_{k,n}$ a $\mathbb{T}$-labelling O by labelling the edges $(i,j) \lessdot (i+1,j)$ with the function $f^{\mathrm{weak}}(x) = x$ and the edges $(i,j) \lessdot (i,j+1)$ with the function $f^{\mathrm{strict}}(x) = x-1$. A finite convex subposet P of $\mathbb{C}_{k,n}$ is known as a *cylindric skew shape*; see [**GK, Pos, McN**]. The (P,O)-tableau are known as *semistandard cylindric tableaux* of shape P and the generating function $K_{P,O}(x_1, x_2, \cdots)$ is the cylindric Schur function defined in [**BS, Pos**].

Example 2.5. Let N be the number of elements in a poset P, and let $\omega : P \longrightarrow [N]$ be a bijective labelling of elements of P with numbers from 1 to N. Recall that a (P,ω)-partition (see [**Sta**]) is a map $\sigma : P \longrightarrow \mathbb{P}$ such that $s \leq t$ in P implies $\sigma(s) \leq \sigma(t)$, while if in addition $\omega(s) > \omega(t)$ then $\sigma(s) < \sigma(t)$. Label now each edge (s,t) of the Hasse diagram of P with f^{weak} or f^{strict}, depending on whether $\omega(s) \leq \omega(t)$ or $\omega(s) > \omega(t)$ correspondingly. It is not hard to see that for this labelling O the (P,O)-tableaux are exactly the (P,ω)-partitions. Similarly, if we allow any labelling of the edges of P with f^{weak} and f^{strict}, we get the *oriented posets* of McNamara; see [**McN**].

3. The Cell Transfer Theorem

A generating function $f \in \mathbb{Q}[[x_1, x_2, \ldots]]$ is *monomial-positive* if all coefficients in its expansion into monomials are non-negative. If f is actually a symmetric function then this is equivalent to f being a non-negative linear combination of monomial symmetric functions.

Let (P,O) be a $\mathbb{T}$-labelled poset. Let Q and R be two finite convex subposets of P. The subset $Q \cap R$ is also a convex subposet. Define two convex subposets $Q \wedge R$ and $Q \vee R$ by

$$Q \wedge R = \{s \in R \mid s < Q\} \cup \{s \in Q \mid s \sim R \text{ or } s < R\} \tag{3.1}$$

and

$$Q \vee R = \{s \in Q \mid s > R\} \cup \{s \in R \mid s \sim Q \text{ or } s > Q\} \tag{3.2}$$

Recall that if A and B are sets then $A \backslash B = \{a \in A \mid a \notin B\}$ denotes the *set difference*.

Lemma 3.1. *The subposets $Q \wedge R$ and $Q \vee R$ are both convex subposets of P. We have $(Q \wedge R) \cup (Q \vee R) = Q \cup R$ and $(Q \wedge R) \cap (Q \vee R) = Q \cap R$.*

Proof. Suppose $s < t$ lie in $Q \wedge R$ and $s < r < t$ for some $r \in P$ but $r \notin Q \wedge R$. Then either $s \in R \backslash Q$ and $t \in Q \backslash R$ or $s \in Q \backslash R$ and $t \in R \backslash Q$. In the first case, since $t > s$ we must have $t > R$ which is impossible by definition. In the second case, we have $t > Q$ which is again impossible. The proof for $Q \vee R$ is analogous. The second statement of the lemma is straightforward. $\square$

Note that the operations $\wedge$ and $\vee$ are stable so that $(Q \wedge R) \wedge (Q \vee R) = Q \wedge R$ and $(Q \wedge R) \vee (Q \vee R) = Q \vee R$.

Theorem 3.2 (Cell Transfer Theorem). *The difference*

$$K_{Q \wedge R, O} K_{Q \vee R, O} - K_{Q,O} K_{R,O}$$

is monomial-positive.

Proof. We prove Theorem 3.2 by exhibiting an injection

$$\eta : \mathcal{A}(Q,O) \times \mathcal{A}(R,O) \longrightarrow \mathcal{A}(Q \wedge R, O) \times \mathcal{A}(Q \vee R, O)$$

which is weight preserving. We call this the *cell transfer procedure*. The name comes from our main examples where elements of a poset are cells of a Young diagram. For convenience, in this paper we call elements of any poset *cells*.

Let ω be a (Q,O)-tableau and σ be a (R,O)-tableau. We now describe how to construct a $(Q \wedge R, O)$-tableau $\omega \wedge \sigma$ and a $(Q \vee R, O)$-tableau $\omega \vee \sigma$. Define a subset of $Q \cap R$, depending on ω and σ, by

$$(Q \cap R)^+ = \{x \in Q \cap R \mid \omega(x) < \sigma(x)\}.$$

We give $(Q \cap R)^+$ the structure of a graph by inducing from the Hasse diagram of $Q \cap R$.

Let $\mathrm{bd}(R) = \{x \in Q \cap R \mid x \gtrdot y \text{ for some } y \in R \backslash Q\}$ be the "lower boundary" of $Q \cap R$ which touches elements in R. Let $\mathrm{bd}(R)^+ \subset (Q \cap R)^+$ be the union of the connected components of $(Q \cap R)^+$ which contain an element of $\mathrm{bd}(R)$. Similarly, let $\mathrm{bd}(Q) = \{x \in Q \cap R \mid x \lessdot y \text{ for some } y \in Q \backslash R\}$ be the "upper boundary" of $Q \cap R$ which touches elements in Q. Let $\mathrm{bd}(Q)^+ \subset (Q \cap R)^+$ be the union of the connected components of $(Q \cap R)^+$ which contain an element of $\mathrm{bd}(Q)$. The elements in $\mathrm{bd}(Q)^+ \cup \mathrm{bd}(R)^+$ are amongst the cells that we might "transfer".

Let $S \subset Q \cap R$. Define $(\omega \wedge \sigma)_S : Q \wedge R \to \mathbb{P}$ by

$$(\omega \wedge \sigma)_S(x) = \begin{cases} \sigma(x) & \text{if } x \in R \backslash Q \text{ or } x \in S, \\ \omega(x) & \text{otherwise.} \end{cases}$$

And define $(\omega \vee \sigma)_S : Q \vee R \to \mathbb{P}$ by

$$(\omega \vee \sigma)_S(x) = \begin{cases} \omega(x) & \text{if } x \in Q \backslash R \text{ or } x \in S, \\ \sigma(x) & \text{otherwise.} \end{cases}$$

One checks directly that $\mathrm{wt}(\sigma) + \mathrm{wt}(\omega) = \mathrm{wt}((\omega \wedge \sigma)_S) + \mathrm{wt}((\omega \vee \sigma)_S)$. We claim that when $S = S^* := \mathrm{bd}(Q)^+ \cup \mathrm{bd}(R)^+$, both $(\omega \wedge \sigma)^*_S$ and $(\omega \vee \sigma)_{S^*}$ respect O. We check this for $(\omega \wedge \sigma)_{S^*}$ and the claim for $(\omega \vee \sigma)_{S^*}$ follows from symmetry.

Let $s \lessdot t$ be a covering relation in $Q \wedge R$. Since σ and ω are assumed to respect O, we need only check the conditions when $(\omega \wedge \sigma)_{S^*}(s) = \omega(s)(\neq \sigma(s))$ and $(\omega \wedge \sigma)_{S^*}(t) = \sigma(t)(\neq \omega(t))$; or when $(\omega \wedge \sigma)_{S^*}(s) = \sigma(s)(\neq \omega(s))$ and $(\omega \wedge \sigma)_{S^*}(t) = \omega(t)(\neq \sigma(t))$.

In the first case, we must have $s \in Q$ and $t \in R$. If $t \in R$ but $t \notin Q$ then by the definition of $Q \wedge R$ we must have $t < Q$ and so $t < t'$ for some $t' \in Q$. This is impossible since Q is convex. Thus $t \in Q \cap R$ and so $t \in S^*$. We compute that $\omega(s) \leq O(s,t)(\omega(t)) \leq O(s,t)(\sigma(t))$ since $\omega(t) < \sigma(t)$ and $O(s,t)$ is weakly increasing.

In the second case, we must have $s \in R$ and $t \in Q$. By the definition of $Q \wedge R$ we must have $t \in R$ as well. So $t \in Q \cap R$ but $t \notin S^*$ which means that $\omega(t) > \sigma(t)$. Thus $\sigma(s) \leq O(s,t)(\sigma(t)) \leq O(s,t)(\omega(t))$ and $\omega \wedge \sigma$ respects O here.

For each (ω, σ), say a subset $S \subseteq S^*$ is *transferrable* if both $(\omega \wedge \sigma)_S$ and $(\omega \vee \sigma)_S$ respect O. If S' and S'' are both transferrable then it is easy to check that so is $S' \cap S''$. Thus there exists a unique smallest transferrable subset $S^\diamond \subseteq S^*$. Now define $\eta : \mathcal{A}(Q,O) \times \mathcal{A}(R,O) \to \mathcal{A}(Q \wedge R, O) \times \mathcal{A}(Q \vee R, O)$ by

$$(\omega, \sigma) \longmapsto ((\omega \wedge \sigma)_{S^\diamond}, (\omega \vee \sigma)_{S^\diamond}).$$

Note that $S^\diamond$ depends on ω and σ, though we have suppressed the dependence from the notation.

We now show that this η is injective. Given $(\alpha, \beta) \in \eta(\mathcal{A}(Q,O) \times \mathcal{A}(R,O))$, we show how to recover ω and σ. As before, for a subset $S \subset Q \cap R$, define $\omega_S = \omega(\alpha,\beta)_S : Q \to \mathbb{P}$ by

$$\omega_S(x) = \begin{cases} \beta(x) & \text{if } x \in (Q \backslash R) \cap (Q \vee R) \text{ or } x \in S, \\ \alpha(x) & \text{otherwise.} \end{cases}$$

And define $\sigma_S = \sigma(\alpha,\beta)_S : R \to \mathbb{P}$ by

$$\sigma_S(x) = \begin{cases} \alpha(x) & \text{if } x \in (R \backslash Q) \cap (Q \wedge R) \text{ or } x \in S, \\ \beta(x) & \text{otherwise.} \end{cases}$$

Note that if $(\alpha,\beta) = ((\omega \wedge \sigma)_{S^\diamond}, (\omega \vee \sigma)_{S^\diamond}))$ then $\omega = \omega_{S^\diamond}$ and $\sigma = \sigma_{S^\diamond}$. Let $S^\square \subset Q \cap R$ be the unique smallest subset such that $\omega_{S^\square}$ and $\sigma_{S^\square}$ both respect O. Since we have assumed that $(\alpha,\beta) \in \eta(\mathcal{A}(Q,O) \times \mathcal{A}(R,O))$, such a $S^\square$ must exist. (As before the intersection of two transferrable subsets with respect to (α, β) is transferrable.)

We now show that if $(\alpha,\beta) = ((\omega \wedge \sigma)_{S^\diamond}, (\omega \vee \sigma)_{S^\diamond}))$ then $S^\square = S^\diamond$. By definition, $S^\square \subset S^\diamond$. Let $C \subset S^\diamond \backslash S^\square$ be a connected component of $S^\diamond \backslash S^\square$, viewed as an induced subgraph of the Hasse diagram of P. We claim that $S^\diamond \backslash C$ is a transferrable set for (ω, σ); this means that changing $\alpha|_C$ to $\omega|_C$ and $\beta|_C$ to $\sigma|_C$ gives a pair in $\mathcal{A}(Q \wedge R, O) \times \mathcal{A}(Q \vee R, O)$. Suppose first that $c \in C$ and $s \in S^\square$ is so that $c \lessdot s$. By the definition of $S^\square$, we must have $\alpha(c) \leq O(c,s)(\beta(s))$ and $\beta(c) \leq O(c,s)(\alpha(s))$. Now suppose that $c \in C$ and $s \in Q \backslash R$ such that $c \lessdot s$. Then we must have $O(c,s)(\omega(s)) = O(c,s)(\beta(s)) \geq \alpha(c) = \sigma(c)$. Similar conclusions hold for $c \gtrdot s$. Thus we have checked that $S^\diamond \backslash C$ is a transferrable set for (ω, σ).

This shows that the map $(\omega, \sigma) \mapsto ((\omega \wedge \sigma)_{S^\diamond}, (\omega \vee \sigma)_{S^\diamond}))$ is injective, completing the proof. $\square$

On Figure 2 we can see how shapes $P \vee Q$ and $P \wedge Q$ are formed in the case of SSYT. On Figure 3 an example of cell transfer for those shapes is given. Note that $S^\diamond$ does not contain one cell which is is in S^*.

Note that $(\omega, \sigma) \mapsto ((\omega \wedge \sigma)_{S^*}, (\omega \vee \sigma)_{S^*})$ also defines a weight-preserving map $\eta^* : \mathcal{A}(Q,O) \times \mathcal{A}(R,O) \to \mathcal{A}(Q \wedge R, O) \times \mathcal{A}(Q \vee R, O)$. Unfortunately, η^* is not always injective.

Suppose P is a locally-finite poset with a minimal element. Let $J(P)$ be the lattice of finite order ideals of P; see [**Sta**]. If $I, J \in J(P)$ then the sets $I \wedge J$ and $I \vee J$ defined in (3.1) and (3.2) are finite order ideals and agree with the the meet $\wedge_{J(P)}$ and join $\vee_{J(P)}$ of I and J respectively within $J(P)$. In fact, by defining $(Q \wedge R)' = \{s \in R \mid s < Q\} \cup \{s \in Q \mid s \in R \text{ or } s < R\}$ and $(Q \vee R)' = \{s \in Q \mid s \sim R\} \cup \{s \in R \mid s \sim$

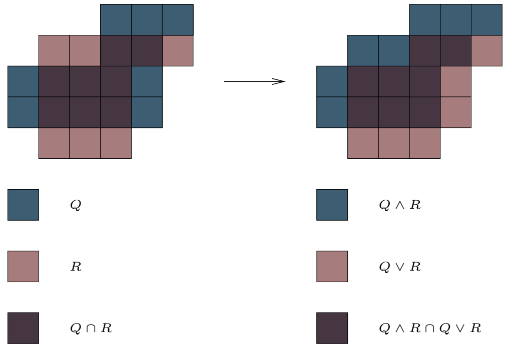


Figure 2. An example of $P \wedge Q$ and $P \vee Q$ for semistandard Young tableaux.

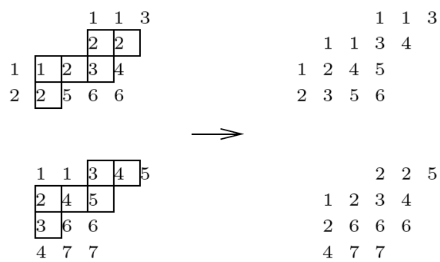


Figure 3. An example of cell transfer for semistandard Young tableaux, cells in $S^\diamond$ are marked.

Q or $s > Q\}$, the order ideals $(I \wedge J)' = I \wedge_{J(P)} J$ and $(I \vee J)' = I \vee_{J(P)} J$ agree with the meet and join in $J(P)$ even when P does not contain a minimal element.

Corollary 3.3. *Let P be a locally-finite poset and $I, J \in J(P)$. Then the generating function*

$$K_{I \wedge_{J(P)} J, O} K_{I \vee_{J(P)} J, O} - K_{I,O} K_{J,O}$$

is monomial-positive.

Proof. The elements altered going from $(Q \wedge R)$ to $(Q \wedge R)'$ do not involve the intersection $Q \cap R$, and in fact are incomparable to the elements of $Q \cap R$. The cells being transferred in the proof of Theorem 3.2 are not affected by changing $(Q \wedge R)$ to $(Q \wedge R)'$ and changing $(Q \vee R)$ to $(Q \vee R)'$. Thus the same proof works here. $\square$

4. Background for Schur positivity proof

In this section we give an overview of some results of Haiman [**Hai**] and Rhoades-Skandera [**RS2, RS1**]. We include an alternative proof Rhoades-Skandera's result.

4.1. Haiman's Schur positivity result. Let $H_n(q)$ be the *Hecke algebra* associated with the symmetric group S_n. The Hecke algebra has the standard basis $\{T_w \mid w \in S_n\}$ and the *Kazhdan-Lusztig basis* $\{C'_w(q) \mid w \in S_n\}$ related by

$$q^{l(v)/2} C'_v(q) = \sum_{w \le v} P_{w,v}(q)\, T_w \quad \text{and} \quad T_w = \sum_{v \le w} (-1)^{l(vw)} Q_{v,w}(q)\, q^{l(v)/2} C'_v(q),$$

where $P_{w,v}(q)$ are the *Kazhdan-Lusztig polynomials* and $Q_{v,w}(q) = P_{w_\circ w, w_\circ v}(q)$, for the longest permutation $w_\circ \in S_n$, see [**Hum**] for more details.

For $w \in S_n$ and a $n \times n$ matrix $X = (x_{ij})$, the *Kazhdan-Lusztig immanant* was defined in [**RS2**] as

$$\mathrm{Imm}_w(X) := \sum_{v \in S_n} (-1)^{l(vw)} Q_{w,v}(1)\, x_{1,v(1)} \cdots x_{n,v(n)},$$

Let $h_k = \sum_{i_1 \leq \cdots \leq i_k} x_{i_1} \cdots x_{i_k}$ be the k-th homogeneous symmetric function, where $h_0 = 1$ and $h_k = 0$ for $k < 0$. A *generalized Jacobi-Trudi matrix* is a $n \times n$ matrix of the form $\left(h_{\mu_i - \nu_j}\right)_{i,j=1}^n$, for partitions $\mu = (\mu_1 \geq \mu_2 \cdots \geq \mu_n \geq 0)$ and $\nu = (\nu_1 \geq \nu_2 \cdots \geq \nu_n \geq 0)$. Haiman's result can be reformulated as follows, see [**RS2**].

THEOREM 4.1. Haiman [**Hai**, Theorem 1.5] *The immanants* Imm_w *of a generalized Jacobi-Trudi matrix are Schur non-negative.*

Haiman's proof of this result is based on the Kazhdan-Lusztig conjecture proven by Beilinson-Bernstein and Brylinski-Kashiwara. This conjecture expresses the characters of Verma modules as sums of the characters of some irreducible highest weight representations of $\mathfrak{sl}_n$ with multiplicities equal to $P_{w,v}(1)$. One can derive from this conjecture that the coefficients of Schur functions in Imm_w are certain tensor product multiplicities of irreducible representations.

4.2. Temperley-Lieb algebra. The *Temperley-Lieb algebra* $TL_n(\xi)$ is the $\mathbb{C}[\xi]$-algebra generated by $t_1, \ldots, t_{n-1}$ subject to the relations $t_i^2 = \xi\, t_i$, $t_i t_j t_i = t_i$ if $|i-j| = 1$, $t_i t_j = t_j t_i$ if $|i-j| \geq 2$. The dimension of $TL_n(\xi)$ equals the n-th Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$. A *321-avoiding permutation* is a permutation $w \in S_n$ that has no reduced decomposition of the form $w = \cdots s_i s_j s_i \cdots$ with $|i-j| = 1$. (These permutations are also called *fully-commutative.*) A natural basis of the Temperley-Lieb algebra is $\{t_w \mid w \text{ is a 321-avoiding permutation in } S_n\}$, where $t_w := t_{i_1} \cdots t_{i_l}$, for a reduced decomposition $w = s_{i_1} \cdots s_{i_l}$.

The map $\theta : T_{s_i} \mapsto t_i - 1$ determines a homomorphism $\theta : H_n(1) = \mathbb{C}[S_n] \to TL_n(2)$. Indeed, the elements $t_i - 1$ in $TL_n(2)$ satisfy the Coxeter relations.

THEOREM 4.2. Fan-Green [**FG**] *The homomorphism* θ *acts on the Kazhdan-Lusztig basis* $\{C'_w(1)\}$ *of* $H_n(1)$ *as follows:*

$$\theta(C'_w(1)) = \begin{cases} t_w & \text{if } w \text{ is 321-avoiding,} \\ 0 & \text{otherwise.} \end{cases}$$

For any permutation $v \in S_n$ and a 321-avoiding permutation $w \in S_n$, let $f_w(v)$ be the coefficient of the basis element $t_w \in TL_n(2)$ in the basis expansion of $\theta(T_v) = (t_{i_1} - 1) \cdots (t_{i_l} - 1) \in TL_n(2)$, for a reduced decomposition $v = s_{i_1} \cdots s_{i_l}$. Rhoades and Skandera [**RS1**] defined the *Temperley-Lieb immanant* $\mathrm{Imm}_w^{\mathrm{TL}}(x)$ of an $n \times n$ matrix $X = (x_{ij})$ by

$$\mathrm{Imm}_w^{\mathrm{TL}}(X) := \sum_{v \in S_n} f_w(v)\, x_{1,v(1)} \cdots x_{n,v(n)}.$$

THEOREM 4.3. Rhoades-Skandera [**RS1**] *For a 321-avoiding permutation* $w \in S_n$*, we have* $\mathrm{Imm}_w^{\mathrm{TL}}(X) = \mathrm{Imm}_w(X)$.

PROOF. Applying the map θ to $T_v = \sum_{w \leq v} (-1)^{l(vw)} Q_{w,v}(1)\, C'_w(1)$ and using Theorem 4.2 we obtain $\theta(T_v) = \sum (-1)^{l(vw)} Q_{w,v}(1)\, t_w$, where the sum is over 321-avoiding permutations w. Thus $f_w(v) = (-1)^{l(vw)} Q_{w,v}(1)$ and $\mathrm{Imm}_w^{\mathrm{TL}} = \mathrm{Imm}_w$. ☐

A product of generators (decomposition) $t_{i_1} \cdots t_{i_l}$ in the Temperley-Lieb algebra TL_n can be graphically presented by a *Temperley-Lieb diagram* with n non-crossing strands connecting the vertices $1, \ldots, 2n$ and, possibly, with some internal loops. This diagram is obtained from the wiring diagram of the decomposition $w = s_{i_1} \cdots s_{i_l} \in S_n$ by replacing each crossing "$\times$" with a *vertical uncrossing* "$)($". For example, the following figure shows the wiring diagram for $s_1 s_2 s_2 s_3 s_2 \in S_4$ and the Temperley-Lieb diagram for $t_1 t_2 t_2 t_3 t_2 \in TL_4$.

Pairs of vertices connected by strands of a wiring diagram are $(2n+1-i, w(i))$, for $i = 1, \ldots, n$. Pairs of vertices connected by strands in a Temperley-Lieb diagram form a *non-crossing matching*, i.e., a graph on the vertices $1, \ldots, 2n$ with n disjoint edges that contains no pair of edges (a, c) and (b, d) with $a < b < c < d$. If two Temperley-Lieb diagrams give the same matching and have the same number of internal loops, then the corresponding products of generators of TL_n are equal to each other. If the diagram of a is obtained from the diagram of b by removing k internal loops, then $b = \xi^k a$ in TL_n.

The map that sends t_w to the non-crossing matching given by its Temperley-Lieb diagram is a bijection between basis elements t_w of TL_n, where w is 321-avoiding, and non-crossing matchings on the vertex set $[2n]$. For example, the basis element $t_1 t_3 t_2$ of TL_4 corresponds to the non-crosssing matching with the edges $(1, 2), (3, 4), (5, 8), (6, 7)$.

4.3. An identity for products of minors. For a subset $S \subset [2n]$, let us say that a Temperley-Lieb diagram (or the associated element in TL_n) is *S-compatible* if each strand of the diagram has one end-point in S and the other end-point in its complement $[2n] \setminus S$. Coloring vertices in S black and the remaining vertices white, a basis element t_w is S-compatible if and only if each edge in the associated matching has two vertices of different colors. Let $\Theta(S)$ denote the set of all 321-avoiding permutation $w \in S_n$ such that t_w is S-compatible.

For two subsets $I, J \subset [n]$ of the same cardinality let $\Delta_{I,J}(X)$ denote the *minor* of an $n \times n$ matrix X in the row set I and the column set J. Let $\bar{I} := [n] \setminus I$ and let $I^\wedge := \{2n+1-i \mid i \in I\}$.

Theorem 4.4. Rhoades-Skandera [**RS1**, Proposition 4.3], cf. Skandera [**Ska**] *For two subsets $I, J \subset [n]$ of the same cardinality and $S = J \cup (\bar{I})^\wedge$, we have*

$$\Delta_{I,J}(X) \cdot \Delta_{\bar{I},\bar{J}}(X) = \sum_{w \in \Theta(S)} \mathrm{Imm}_w^{\mathrm{TL}}(X).$$

The proof given in [**RS1**] employs planar networks. We give a more direct proof that uses the involution principle.

Proof. Let us fix a permutation $v \in S_n$ with a reduced decomposition $v = s_{i_1} \cdots s_{i_l}$. The coefficient of the monomial $x_{1,v(1)} \cdots x_{n,v(n)}$ in the expansion of the product of two minors $\Delta_{I,J}(X) \cdot \Delta_{\bar{I},\bar{J}}(X)$ equals

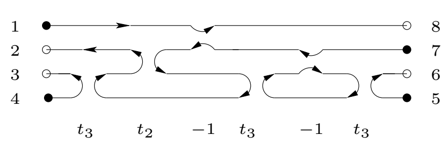
$$\begin{cases} (-1)^{\mathrm{inv}(I)+\mathrm{inv}(\bar{I})} & \text{if } v(I) = J, \\ 0 & \text{if } v(I) \neq J, \end{cases}$$

where $\mathrm{inv}(I)$ is the number of inversions $i < j$, $v(i) > v(j)$ such that $i, j \in I$.

On the other hand, by the definition of $\mathrm{Imm}_w^{\mathrm{TL}}$, the coefficient of $x_{1,v(1)} \cdots x_{n,v(n)}$ in the right-hand side of the identity equals the sum $\sum (-1)^r\, 2^s$ over all diagrams obtained from the wiring diagram of the reduced decomposition $s_{i_1} \cdots s_{i_l}$ by replacing each crossing " $\times$ " with either a *vertical uncrossing* ") (" or a *horizontal uncrossing* " $\asymp$ " so that the resulting diagram is S-compatible, where r is the number of horizontal uncrossings " $\asymp$ " and s is the number of internal loops in the resulting diagram. Indeed, the choice of ") (" corresponds to the choice of "t_{i_k}" and the choice of " $\asymp$ " corresponds to the choice of "-1" in the k-th term of the product $(t_{i_1} - 1) \cdots (t_{i_l} - 1) \in TL_n(2)$, for $k = 1, \ldots, l$.

Let us pick directions of all strands and loops in such diagrams so that the initial vertex in each strand belongs to S (and, thus, the end-point is not in S). There are 2^s ways to pick directions of s internal loops. Thus the above sum can be written as the sum $\sum (-1)^r$ over such *directed Temperley-Lieb* diagrams.

Here is an example of a directed diagram for $v = s_3 s_2 s_1 s_3 s_2 s_3$ and $S = \{1, 4, 5, 7\}$ corresponding to the term $t_3 t_2 (-1) t_3 (-1) t_3$ in the expansion of the product $(t_3 - 1)(t_2 - 1)(t_1 - 1)(t_3 - 1)(t_2 - 1)(t_3 - 1)$. This diagram comes with the sign $(-1)^2$.



Let us construct a sign reversing partial involution ι on the set of such directed Temperley-Lieb diagrams. If a diagram has a *misaligned uncrossing*, i.e., an uncrossing of the form ") (", ") (", " $\asymp$ ", or " $\asymp$ ", then ι switches the leftmost such uncrossing according to the rules ι :) ($\leftrightarrow$ $\asymp$ and ι :) ($\leftrightarrow$ $\asymp$. Otherwise, when the diagram involves only *aligned uncrossings* " ", " ", " ", " ", the involution ι is not defined.

For example, in the above diagram, the involution ι switches the second uncrossing, which has the form " ", to " ". The resulting diagram corresponds to the term $t_3(-1)(-1)t_3(-1)t_3$.

Since the involution ι reverses signs, this shows that the total contribution of all diagrams with at least one misaligned uncrossing is zero. Let us show that there is at most one S-compatible directed Temperley-Lieb diagram with all aligned uncrossings. If we have a such diagram, then we can direct the strands of the wiring diagram for $v = s_{i_1} \dots s_{i_l}$ so that each segment of the wiring diagram has the same direction as in the Temperley-Lieb diagram. In particular, the end-points of strands in the wiring diagram should have different colors. Thus each strand starting at an element of J should finish at an element of $I^\wedge$, or, equivalently, $v(I) = J$. The directed Temperley-Lieb diagram can be uniquely recovered from this directed wiring diagram by replacing the crossings with uncrossings, as follows: $\to$, $\to$, $\to$, $\to$. Thus the coefficient of $x_{1,v(1)} \cdots x_{n,v(n)}$ in the right-hand side of the needed identity is zero, if $v(I) \neq J$, and is $(-1)^r$, if $v(I) = J$, where r is the number of crossings of the form " " or " " in the wiring diagram. In other words, r equals the number of crossings such that the right end-points of the pair of crossing strands have the same color. This is exactly the same as the expression for the coefficient in the left-hand side of the needed identity. $\square$

5. Proof of Theorem 1.1

For two subsets $I, J \subseteq [n]$ of the same cardinality, let $\Delta_{I,J}(H)$ denote the minor of the Jacobi-Trudi matrix $H = (h_{j-i})_{1 \le i,j \le n}$ with row set I and column set J, where h_i is the i-th homogeneous symmetric function, as before. According to the Jacobi-Trudi formula, see [**Mac**], the minors $\Delta_{I,J}(H)$ are precisely the skew Schur functions

$$\Delta_{I,J}(H) = s_{\lambda/\mu},$$

where $\lambda = (\lambda_1 \ge \cdots \ge \lambda_k \ge 0)$, $\mu = (\mu_1 \ge \cdots \ge \mu_k \ge 0)$ and the associated subsets are $I = \{\mu_k + 1 < \mu_{k-1} + 2 < \cdots < \mu_1 + k\}$, $J = \{\lambda_k + 1 < \lambda_{k-1} + 2 < \cdots < \lambda_1 + k\}$.

For two sets $I = \{i_1 < \cdots < i_k\}$ and $J = \{j_1 < \cdots < j_k\}$, let us define $I \vee J := \{\max(i_1, j_1) < \cdots < \max(i_k, j_k)\}$ and $I \wedge J := \{\min(i_1, j_1) < \cdots < \min(i_k, j_k)\}$.

Theorem 1.1 can be reformulated in terms of minors, as follows. Without loss of generality we can assume that all partitions λ, μ, ν, ρ in Theorem 1.1 have the same number k of parts, some of which might be zero. Note that generalized Jacobi-Trudi matrices are obtained from H by skipping or duplicating rows and columns.

Theorem 5.1. *Let I, J, I', J' be k element subsets in $[n]$. Then we have*

$$\Delta_{I\vee I',\, J \vee J'}(X) \cdot \Delta_{I \wedge I',\, J \wedge J'}(X) \ge_s \Delta_{I,J}(X) \cdot \Delta_{I',J'}(X),$$

for a generalized Jacobi-Trudi matrix X.

Proof. Let us denote $\bar{I} := [n] \setminus I$ and $\check{S} := [2n] \setminus S$. By skipping or duplicating rows and columns of the matrix X, we may assume that $I' = \bar{I}$ and $J' = \bar{J}$. Then $I \vee I' = \overline{I \wedge I'}$ and $J \vee J' = \overline{J \wedge J'}$. Let $S := J \cup (\bar{I})^\wedge$ and $T := (J \vee J') \cup (\overline{I \vee I'})^\wedge$. Then we have $T = S \vee \check{S}$ and $\check{T} = S \wedge \check{S}$.

Let us show that $\Theta(S) \subseteq \Theta(T)$, i.e., every S-compatible non-crossing matching on $[2n]$ is also T-compatible. Let $S = \{s_1 < \cdots < s_n\}$ and $\check{S} = \{\check{s}_1 < \cdots < \check{s}_n\}$. Then $T = \{\max(s_1, \check{s}_1), \dots, \max(s_n, \check{s}_n)\}$ and $\check{T} = \{\min(s_1, \check{s}_1), \dots, \min(s_n, \check{s}_n)\}$. Let M be an S-compatible non-crossing matching on $[2n]$ and let $(a < b)$ be an edge of M. Without loss of generality we may assume that $a = s_i \in S$ and $b = \check{s}_j \in \check{S}$. We must show that either ($a \in T$ and $b \in \check{T}$) or ($a \in \check{T}$ and $b \in T$). Since no edge of M can cross (a, b), the elements of S in the interval $[a+1, b-1]$ are matched with the elements of $\check{S}$ in this interval. Let $k = \#(S \cap [a+1, b-1]) = \#(\check{S} \cap [a+1, b-1])$. Suppose that $a, b \in T$, or, equivalently, $\check{s}_i < s_i$ and $s_j < \check{s}_j$. Since there are at least k elements of $\check{S}$ in the interval $[\check{s}_i + 1, \check{s}_j - 1]$, we have $i + k + 1 \le j$. On the other hand, since there are at most $k - 1$ elements of S in the interval $[s_i + 1, s_j - 1]$, we have $i + k \ge j$. We obtain a contradiction. The case $a, b \in \check{T}$ is analogous.

Now Theorem 4.4 implies that the difference $\Delta_{I \vee I',\, J \vee J'} \cdot \Delta_{I \wedge I',\, J \wedge J'} - \Delta_{I,J} \cdot \Delta_{I',J'}$ is a nonnegative combination of Temperley-Lieb immanants. Theorems 4.1 and 4.3 imply its Schur nonnegativity. $\square$

6. Proof of conjectures and generalizations

In this section we prove generalized versions of Conjectures 1.2-1.4, which were conjectured by Kirillov [**Kir**, Section 6.8]. Corollary 6.2 was also conjectured by Bergeron-McNamara [**BM**, Conjecture 5.2] who showed that it implies Theorem 6.3.

Let $\lfloor x \rfloor$ be the maximal integer $\leq x$ and $\lceil x \rceil$ be the minimal integer $\geq x$. For vectors v and w and a positive integer n, we assume that the operations $v+w$, $\frac{v}{n}$, $\lfloor v \rfloor$, $\lceil v \rceil$ are performed coordinate-wise. In particular, we have well-defined operations $\lfloor \frac{\lambda+\nu}{2} \rfloor$ and $\lceil \frac{\lambda+\nu}{2} \rceil$ on pairs of partitions.

The next claim extends Okounkov's conjecture (Conjecture 1.2).

THEOREM 6.1. *Let λ/μ and ν/ρ be any two skew shapes. Then we have*

$$s_{\lfloor \frac{\lambda+\nu}{2} \rfloor / \lfloor \frac{\mu+\rho}{2} \rfloor}\, s_{\lceil \frac{\lambda+\nu}{2} \rceil / \lceil \frac{\mu+\rho}{2} \rceil} \geq_s s_{\lambda/\mu}\, s_{\nu/\rho}.$$

PROOF. We will assume that all partitions have the same fixed number k of parts, some of which might be zero. For a skew shape $\lambda/\mu = (\lambda_1, \ldots, \lambda_k)/(\mu_1, \ldots, \mu_k)$, define

$$\overrightarrow{\lambda/\mu} := (\lambda_1+1, \ldots, \lambda_k+1)/(\mu_1+1, \ldots, \mu_k+1),$$

that is, $\overrightarrow{\lambda/\mu}$ is the skew shape obtained by shifting the shape λ/μ one step to the right. Similarly, define the left shift of λ/μ by

$$\overleftarrow{\lambda/\mu} := (\lambda_1-1, \ldots, \lambda_k-1)/(\mu_1-1, \ldots, \mu_k-1),$$

assuming that the result is a legitimate skew shape. Note that $s_{\lambda/\mu} = s_{\overleftarrow{\lambda/\mu}} = s_{\overrightarrow{\lambda/\mu}}$.

Let θ be the operation on pairs of skew shapes given by

$$\theta : (\lambda/\mu, \nu/\rho) \longmapsto ((\lambda/\mu) \vee (\nu/\rho), (\lambda/\mu) \wedge (\nu/\rho)).$$

According to Theorem 1.1, the product of the two skew Schur functions corresponding to the shapes in $\theta(\lambda/\mu, \nu/\rho)$ is $\geq_s s_{\lambda/\mu}\, s_{\nu/\rho}$. Let us show that we can repeatedly apply the operation θ together with the left and right shifts of shapes and the flips $(\lambda/\mu, \nu/\rho) \mapsto (\nu/\rho, \lambda/\mu)$ in order to obtain the pair of skew shapes $(\lfloor \frac{\lambda+\nu}{2} \rfloor / \lfloor \frac{\mu+\rho}{2} \rfloor, \lceil \frac{\lambda+\nu}{2} \rceil / \lceil \frac{\mu+\rho}{2} \rceil)$ from $(\lambda/\mu, \nu/\rho)$.

Let us define two operations ϕ and ψ on ordered pairs of skew shapes by conjugating θ with the right and left shifts and the flips, as follows:

$$\phi : (\lambda/\mu, \nu/\rho) \longmapsto ((\lambda/\mu) \wedge (\overrightarrow{\nu/\rho}), \overleftarrow{(\lambda/\mu) \vee (\overrightarrow{\nu/\rho})}),$$

$$\psi : (\lambda/\mu, \nu/\rho) \longmapsto (\overleftarrow{(\overrightarrow{\lambda/\mu}) \vee (\nu/\rho)}, (\overrightarrow{\lambda/\mu}) \wedge (\nu/\rho)).$$

In this definition the application of the left shift "$\leftarrow$" always makes sense. Indeed, in both cases, before the application of "$\leftarrow$", we apply "$\rightarrow$" and then "$\vee$". As we noted above, both products of skew Schur functions for shapes in $\phi(\lambda/\mu, \nu/\rho)$ and in $\psi(\lambda/\mu, \nu/\rho)$ are $\geq_s s_{\lambda/\mu}\, s_{\nu/\rho}$.

It is convenient to write the operations ϕ and ψ in the coordinates $\lambda_i, \mu_i, \nu_i, \rho_i$, for $i = 1, \ldots, k$. These operations independently act on the pairs (λ_i, ν_i) by

$$\begin{aligned} \phi &: (\lambda_i, \nu_i) \mapsto (\min(\lambda_i, \nu_i+1), \max(\lambda_i, \nu_i+1) - 1), \\ \psi &: (\lambda_i, \nu_i) \mapsto (\max(\lambda_i+1, \nu_i) - 1, \min(\lambda_i+1, \nu_i)), \end{aligned}$$

and independently act on the pairs (μ_i, ρ_i) by exactly the same formulas. Note that both operations ϕ and ψ preserve the sums $\lambda_i + \nu_i$ and $\mu_i + \rho_i$.

The operations ϕ and ψ transform the differences $\lambda_i - \nu_i$ and $\mu_i - \rho_i$ according to the following piecewise-linear maps:

$$\bar{\phi}(x) = \begin{cases} x & \text{if } x \leq 1, \\ 2 - x & \text{if } x \geq 1, \end{cases} \qquad \text{and} \qquad \bar{\psi}(x) = \begin{cases} x & \text{if } x \geq -1, \\ -2 - x & \text{if } x \leq -1. \end{cases}$$

Whenever we apply the composition $\phi \circ \psi$ of these operations, all absolute values $|\lambda_i - \nu_i|$ and $|\mu_i - \rho_i|$ strictly decrease, if these absolute values are ≥ 2. It follows that, for a sufficiently large integer N, we have $(\phi \circ \psi)^N(\lambda/\mu, \nu/\rho) = (\tilde{\lambda}/\tilde{\mu}, \tilde{\nu}/\tilde{\rho})$ with $\tilde{\lambda}_i + \tilde{\nu}_i = \lambda_i + \nu_i$, $\tilde{\mu}_i + \tilde{\rho}_i = \mu_i + \rho_i$, and $|\tilde{\lambda}_i - \tilde{\nu}_i| \leq 1$, $|\tilde{\mu}_i - \tilde{\rho}_i| \leq 1$, for all i. Finally, applying the operation θ, we obtain $\theta(\tilde{\lambda}/\tilde{\mu}, \tilde{\nu}/\tilde{\rho}) = (\lceil \frac{\lambda+\nu}{2} \rceil / \lceil \frac{\mu+\rho}{2} \rceil, \lfloor \frac{\lambda+\nu}{2} \rfloor / \lfloor \frac{\mu+\rho}{2} \rfloor)$, as needed. $\square$

The following conjugate version of Theorem 6.1 extends Fomin-Fulton-Li-Poon's conjecture (Conjecture 1.3) to skew shapes.

Corollary 6.2. *Let λ/μ and ν/ρ be two skew shapes. Then we have*

$$s_{\mathrm{sort}_1(\lambda,\nu)/\mathrm{sort}_1(\mu,\rho)}\, s_{\mathrm{sort}_2(\lambda,\nu)/\mathrm{sort}_2(\mu,\rho)} \geq_s s_{\lambda/\mu}\, s_{\nu/\rho}.$$

Proof. This statement is obtained from Theorem 6.1 by conjugating the shapes. Indeed, $\lceil\frac{\lambda+\mu}{2}\rceil' = \mathrm{sort}_1(\lambda',\mu')$ and $\lfloor\frac{\lambda+\mu}{2}\rfloor' = \mathrm{sort}_2(\lambda',\mu')$. Here λ' denote the partition conjugate to λ. $\square$

Theorem 6.3. *Let $\lambda^{(1)}/\mu^{(1)},\ldots,\lambda^{(n)}/\mu^{(n)}$ be n skew shapes, let $\lambda = \bigcup\lambda^{(i)}$ be the partition obtained by the decreasing rearrangement of the parts in all $\lambda^{(i)}$, and, similarly, let $\mu = \bigcup\mu^{(i)}$. Then we have $\prod_{i=1}^{n} s_{\lambda^{[i,n]}/\mu^{[i,n]}} \geq_s \prod_{i=1}^{n} s_{\lambda^{(i)}/\mu^{(i)}}$.*

This theorem extends Corollary 6.2 and Conjecture 1.3. Also note that Lascoux-Leclerc-Thibon's conjecture (Conjecture 1.4) is a special case of Theorem 6.3 for the n-tuple of partitions $(\lambda^{[1,m]},\ldots,\lambda^{[m,m]},\emptyset,\ldots,\emptyset)$.

Proof. Let us derive the statement by applying Corollary 6.2 repeatedly. For a sequence $v = (v_1, v_2, \ldots, v_l)$ of integers, the *anti-inversion number* is $\mathrm{ainv}(v) := \#\{(i,j) \mid i<j,\ v_i<v_j\}$. Let $L = (\lambda^{(1)}/\mu^{(1)},\ldots,\lambda^{(n)}/\mu^{(n)})$ be a sequence of skew shapes. Define its anti-inversion number as

$$\begin{aligned}\mathrm{ainv}(L) = {} & \mathrm{ainv}(\lambda_1^{(1)},\lambda_1^{(2)},\ldots,\lambda_1^{(n)},\lambda_2^{(1)},\ldots,\lambda_2^{(n)},\lambda_3^{(1)},\ldots,\lambda_3^{(n)},\ldots)\\ & + \mathrm{ainv}(\mu_1^{(1)},\mu_1^{(2)},\ldots,\mu_1^{(n)},\mu_2^{(1)},\ldots,\mu_2^{(n)},\mu_3^{(1)},\ldots,\mu_3^{(n)},\ldots).\end{aligned}$$

If $\mathrm{ainv}(L)\neq 0$ then there is a pair $k<l$ such that $\mathrm{ainv}(\lambda^{(k)}/\mu^{(k)},\lambda^{(l)}/\mu^{(l)})\neq 0$. Let $\tilde{L}$ be the sequence of skew shapes obtained from L by replacing the two terms $\lambda^{(k)}/\mu^{(k)}$ and $\lambda^{(l)}/\mu^{(l)}$ with the terms

$$\mathrm{sort}_1(\lambda^{(k)},\lambda^{(l)})/\mathrm{sort}_1(\mu^{(k)},\mu^{(l)}) \quad \text{and} \quad \mathrm{sort}_2(\lambda^{(k)},\lambda^{(l)})/\mathrm{sort}_2(\mu^{(k)},\mu^{(l)}),$$

correspondingly. Then $\mathrm{ainv}(\tilde{L}) < \mathrm{ainv}(L)$. Indeed, if we rearrange a subsequence in a sequence in the decreasing order, the total number of anti-inversions decreases. According to Corollary 6.2, we have $s_{\tilde{L}} \geq_s s_L$, where $s_L := \prod_{i=1}^{n} s_{\lambda^{(i)}/\mu^{(i)}}$. Note that the operation $L\mapsto\tilde{L}$ does not change the unions of partitions $\bigcup\lambda^{(i)}$ and $\bigcup\mu^{(i)}$. Let us apply the operations $L\mapsto\tilde{L}$ for various pairs (k,l) until we obtain a sequence of skew shapes $\hat{L} = (\hat{\lambda}^{(1)}/\hat{\mu}^{(1)},\ldots,\hat{\lambda}^{(n)}/\hat{\mu}^{(n)})$ with $\mathrm{ainv}(\hat{L}) = 0$, i.e., the parts of all partitions must be sorted as $\hat{\lambda}_1^{(1)}\geq\cdots\geq\hat{\lambda}_1^{(n)}\geq\hat{\lambda}_2^{(1)}\geq\cdots\geq\hat{\lambda}_2^{(n)}\geq\hat{\lambda}_3^{(1)}\geq\cdots\geq\hat{\lambda}_3^{(n)}\geq\cdots$, and the same inequalities hold for the $\hat{\mu}_j^{(i)}$. This means that $\hat{\lambda}^{(i)}/\hat{\mu}^{(i)} = \lambda^{[i,n]}/\mu^{[i,n]}$, for $i=1,\ldots,n$. Thus $s_{\hat{L}} = \prod s_{\lambda^{[i,n]}/\mu^{[i,n]}} \geq_s s_L$, as needed. $\square$

Let us define $\lambda^{\{i,n\}} := ((\lambda')^{[i,n]})'$, for $i=1,\ldots,n$. Here λ' again denotes the partition conjugate to λ. The partitions $\lambda^{\{i,n\}}$ are uniquely defined by the conditions $\lceil\frac{\lambda}{n}\rceil \supseteq \lambda^{\{1,n\}} \supseteq \cdots \supseteq \lambda^{\{n,n\}} \supseteq \lfloor\frac{\lambda}{n}\rfloor$ and $\sum_{i=1}^{n}\lambda^{\{i,n\}} = \lambda$. In particular, $\lambda^{\{1,2\}} = \lceil\frac{\lambda}{2}\rceil$ and $\lambda^{\{2,2\}} = \lfloor\frac{\lambda}{2}\rfloor$. If $\frac{\lambda}{n}$ is a partition, i.e., all parts of λ are divisible by n, then $\lambda^{\{i,n\}} = \frac{\lambda}{n}$ for each $1\leq i\leq n$.

Corollary 6.4. *Let $\lambda^{(1)}/\mu^{(1)},\ldots,\lambda^{(n)}/\mu^{(n)}$ be n skew shapes, let $\lambda = \lambda^{(1)}+\cdots+\lambda^{(n)}$ and $\mu = \mu^{(1)}+\cdots+\mu^{(n)}$. Then we have $\prod_{i=1}^{n} s_{\lambda^{\{i,n\}}/\mu^{\{i,n\}}} \geq_s \prod_{i=1}^{n} s_{\lambda^{(i)}/\mu^{(i)}}$.*

Proof. This claim is obtained from Theorem 6.3 by conjugating the shapes. Indeed, $\left(\bigcup\lambda^{(i)}\right)' = \sum(\lambda^{(i)})'$. $\square$

For a skew shape λ/μ and a positive integer n, define $s^{\langle n\rangle}_{\frac{\lambda}{n}/\frac{\mu}{n}} := \prod_{i=1}^{n} s_{\lambda^{\{i,n\}}/\mu^{\{i,n\}}}$. In particular, if $\frac{\lambda}{n}$ and $\frac{\mu}{n}$ are partitions, then $s^{\langle n\rangle}_{\frac{\lambda}{n}/\frac{\mu}{n}} = \left(s_{\frac{\lambda}{n}/\frac{\mu}{n}}\right)^n$.

Corollary 6.5. *Let c and d be positive integers and $n=c+d$. Let λ/μ and ν/ρ be two skew shapes. Then $s^{\langle n\rangle}_{\frac{c\lambda+d\nu}{n}/\frac{c\mu+d\rho}{n}} \geq_s s^c_{\lambda/\mu}\, s^d_{\nu/\rho}$.*

Theorem 6.1 is a special case of Corollary 6.5 for $c=d=1$.

Proof. This claim follows from Corollary 6.4 for the sequence of skew shapes that consists of λ/μ repeated c times and ν/ρ repeated d times. $\square$

Corollary 6.5 implies that the map $S:\lambda\mapsto s_\lambda$ from the set of partitions to symmetric functions satisfies the following "Schur log-concavity" property.

COROLLARY 6.6. *For positive integers c, d and partitions λ, μ such that $\frac{c\lambda+d\mu}{c+d}$ is a partition, we have* $\left(S\left(\frac{c\lambda+d\mu}{c+d}\right)\right)^{c+d} \geq_s S(\lambda)^c S(\mu)^d$.

This notion of Schur log-concavity is inspired by Okounkov's notion of log-concavity; see [**Oko**].

ACKNOWLEDGEMENTS: We thank Richard Stanley for useful conversations. We are grateful to Sergey Fomin for helpful comments and suggestions and to Mark Skandera for help with the references.

References

[BM] F. BERGERON AND P. MCNAMARA: Some positive differences of products of Schur functions, arXiv: math.CO/0412289.

[BBR] F. BERGERON, R. BIAGIOLI, AND M. ROSAS: Inequalities between Littlewood-Richardson Coefficients, preprint; math.CO/0403541.

[BS] N. BERGERON AND F. SOTTILE: Skew Schubert functions and the Pieri formula for flag manifolds, with Nantel Bergeron. *Trans. Amer. Math. Soc.*, **354** No.2, (2002), 651-673.

[FFLP] S. FOMIN, W. FULTON, C.-K. LI AND Y.-T. POON: Eigenvalues, singular values, and Littlewood-Richardson coefficients, *American Journal of Mathematics*, **127** (2005), 101-127.

[FG] K. FAN AND R.M. GREEN: Monomials and Temperley-Lieb algebras, *Journal of Algebra*, **190** (1997), 498-517.

[GK] I. GESSEL AND C. KRATTENTHALER: Cylindric Partitions, *Trans. Amer. Math. Soc.* **349** (1997), 429-479.

[Hai] M. HAIMAN: Hecke algebra characters and immanant conjectures, *J. Amer. Math. Soc.*, **6** (1993), 569-595.

[Hum] J. HUMPHREYS: *Reflection Groups and Coxeter Groups,* Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1992.

[Kir] A. KIRILLOV: An invitation to the generalized saturation conjecture, *Publications of RIMS Kyoto University* **40** (2004), 1147-1239.

[Lam] T. LAM: Affine Stanley Symmetric Functions, preprint; math.CO/0501335.

[LP] T. LAM AND P. PYLYAVSKYY: Cell transfer and monomial positivity, arXiv: math.CO/ 0505273.

[LPP] T. LAM, A. POSTNIKOV AND P. PYLYAVSKYY: Schur positivity and Schur log-concavity, arXiv: math.CO/ 0502446.

[LLT] A. LASCOUX, B. LECLERC AND J.-Y. THIBON: Ribbon tableaux, Hall-Littlewood symmetric functions, quantum affine algebras, and unipotent varieties, *Journal of Mathematical Physics*, **38**(3) (1997), 1041-1068.

[Mac] I. G. MACDONALD: *Symmetric Functions and Hall Polynomials,* Oxford, 1970.

[McN] P. MCNAMARA: Cylindric Skew Schur Functions, preprint; math.CO/0410301.

[Oko] A. OKOUNKOV: Log-concavity of multiplicities with applications to characters of $U(\infty)$, *Advances in Mathematics*, **127** no. 2 (1997), 258-282.

[Pos] A. POSTNIKOV: Affine approach to quantum Schubert calculus, *Duke Math. J.*, to appear; math.CO/0205165.

[RS1] B. RHOADES AND M. SKANDERA: Temperley-Lieb immanants, to appear in *Annals of Combinatorics*; http://www.math.dartmouth.edu/~skan/papers.htm.

[RS2] B. RHOADES AND M. SKANDERA: Kazhdan-Lusztig immanants and products of matrix minors, preprint, November 19, 2004; http://www.math.dartmouth.edu/~skan/papers.htm.

[Ska] M. SKANDERA: Inequalities in products of minors of totally nonnegative matrices, *Journal of Algebraic Combinatorics* **20** (2004), no. 2, 195–211.

[Sta] R. STANLEY: *Enumerative Combinatorics, Vol 2*, Cambridge, 1999.

DEPARTMENT OF MATHEMATICS, HARVARD, USA
E-mail address: tfylam@fas.harvard.edu

DEPARTMENT OF MATHEMATICS, MIT, USA, 02139
E-mail address: apost@math.mit.edu

DEPARTMENT OF MATHEMATICS, MIT, USA, 02139
E-mail address: pasha@mit.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

On the Combinatorics of Crystal Graphs, I

Cristian Lenart

Abstract. In this paper, we continue the development of a new combinatorial model for the irreducible characters of a complex semisimple Lie group. The main results of this paper are: (1) a combinatorial description of the crystal graphs corresponding to the irreducible representations (this result includes a transparent proof, based on the Yang-Baxter equation, of the fact that the mentioned description does not depend on the choice involved in our model); (2) a combinatorial realization (which is the first direct generalization of *Schützenberger's involution* on tableaux) of a certain fundamental involution on the canonical basis exhibiting the crystals as self-dual posets; (3) an analog for arbitrary root systems, based on the Yang-Baxter equation, of Schützenberger's sliding algorithm, which is also known as jeu de taquin (this algorithm has many applications to the representation theory of the Lie algebra of type A). Our approach is type-independent.

Résumé.
Dans cet article, nous continuons le développement d'un nouveau modèle combinatoire pour les caractères irréductibles d'un groupe de Lie complexe semisimple. Les résultats principaux de cet article sont : (1) une description combinatoire des graphes cristallins correspondant aux représentations irréductibles (ce résultat inclut une preuve transparente, basée sur l'équation de Yang-Baxter, du fait que la description mentionnée ne dépend pas du choix impliqué dans notre modèle) ; (2) une réalisation combinatoire (qui est la première généralisation directe de *l'involution de Schützenberger* sur les tableaux) d'une involution fondamentale sur la base canonique pour laquelle les cristaux sont des ensembles partiellement ordonnés auto-dual ; (3) un analogue de l'algorithme coulissant de Schützenberger, qui est également connu sous le nom "jeu de taquin", pour les systèmes de racine. Cet analogue est basé sur l'équation de Yang-Baxter. Notre approche est indépendante du choix du type du système de racine.

1. Introduction

We have recently given a simple combinatorial model for the irreducible characters of a complex semisimple Lie group G and, more generally, for the *Demazure characters* [**12**]. For reasons explained below, we call our model the *alcove path model*. This was extended to complex symmetrizable Kac-Moody algebras in [**13**] (that is, to infinite root systems).

The alcove path model leads to an extensive generalization of the combinatorics of irreducible characters from Lie type A (where the combinatorics is based on Young tableaux, for instance) to arbitrary type; our approach is type-independent. The present paper continues the study of the combinatorics of the new model, which was started in [**12, 13**].

The main results of this paper are:

(1) a combinatorial description of the crystal graphs corresponding to the irreducible representations (Corollary 4.4); this result includes a transparent proof, based on the Yang-Baxter equation, of the fact that the mentioned description does not depend on the choice involved in our model (Corollary 4.3);

2000 *Mathematics Subject Classification.* Primary 05E15; Secondary 17B10, 20G42, 22E46.
Key words and phrases. Bruhat order, crystals, root operators, Schützenberger's involution, λ-chains, Yang-Baxter moves.
Cristian Lenart was supported by National Science Foundation grant DMS-0403029.

(2) a combinatorial realization of a certain fundamental involution on the canonical basis (Theorem 5.4, see also Example 5.6); this involution [**18**] exhibits the crystals as self-dual posets, and corresponds to the action of the longest Weyl group element on an irreducible representation; our combinatorial realization is the first direct generalization of *Schützenberger's involution* on tableaux (see e.g. [**6**]);
(3) an analog for arbitrary root systems, based on the Yang-Baxter equation, of Schützenberger's *sliding* algorithm, which is also known as *jeu de taquin* (Section 4); this algorithm has many applications to the representation theory of the Lie algebra of type A (see e.g. [**6**]).

Our model is based on the choice of an *alcove path*, which is a sequence of adjacent alcoves for the *affine Weyl group* W_{aff} of the Langland's dual group $G^\vee$. An alcove path is best represented as a λ-*chain*, that is, as a sequence of positive roots corresponding to the common walls of successive alcoves in the mentioned sequence of alcoves. These chains extend the notion of a *reflection ordering* [**5**]. Given a fixed λ-chain, the objects that generalize semistandard Young tableaux are all the subsequences of roots that give rise to saturated increasing chains in Bruhat order (on the Weyl group W) upon multiplying on the right by the corresponding reflections. We call these subsequences *admissible subsets*. In [**13**] we defined *root operators* on admissible subsets, which are certain partial operators associated with the simple roots; in type A, they correspond to the *coplactic operations* on tableaux [**17**]. The root operators produce a directed colored graph structure and a poset structure on admissible subsets. We showed in [**13**] that this graph is isomorphic to the crystal graph of the corresponding irreducible representation if the chosen λ-chain is a special one. All this background information on the alcove path model is explained in more detail in Section 3, following some general background material discussed in Section 2.

In Section 4, we study certain discrete moves which allow us to deform any λ-chain into any other λ-chain (for a fixed dominant weight λ), and to biject the corresponding admissible subsets. We call these moves *Yang-Baxter moves* since they express the fact that certain operators satisfy the *Yang-Baxter equation*. We will explain below the reason for which the Yang-Baxter moves can be considered an analog of jeu de taquin for arbitrary root systems. We show that the Yang-Baxter moves commute with the root operators; this means that the directed colored graph defined by the root operators is invariant under Yang-Baxter moves, and it is thus independent from the choice of a λ-chain. Based on the special case in [**13**] discussed above, this immediately implies that the mentioned graph is isomorphic to the corresponding crystal graph for any choice of a λ-chain.

In Section 5, we present a combinatorial description of a certain fundamental involution η_λ on the *canonical basis*. Such a description was given by Schützenberger in type A in terms of tableaux, and the corresponding procedure is known as *evacuation*. The importance of this involution stems from the fact that it exhibits the crystals as self-dual posets, and it corresponds to the action of the longest Weyl group element on an irreducible representation; it also appears in other contexts, such as the recent realization of the category of crystals as a *coboundary category* [**8**]. Our description of the mentioned involution is very similar to that of the evacuation map. The main ingredient in defining the latter map, namely Schützenberger's sliding algorithm (also known as jeu de taquin), is replaced by Yang-Baxter moves. There is another ingredient, which has to do with "reversing" a λ-chain and an associated admissible subset, by analogy with reversing the word of a tableau in the definition of the evacuation map. Our construction also leads to a purely combinatorial proof of the fact that the crystals (as defined by our root operators) are self-dual posets.

The relationship between the alcove path model and other models for the irreducible characters of semisimple Lie algebras, such as the Littelmann path model, LS paths [**14, 15, 16**], and LS-galleries [**7**], was discussed in [**12, 13**].

As far as analogs of jeu de taquin are concerned, let us mention that the only such analog known in the Littelmann path model is the one due to van Leeuwen [**11**]. The goal of the mentioned paper was to use this analog in order to express in a bijective manner the symmetry of the Littlewood-Richardson rule in the Littelmann path model.

Let us also mention that an explicit description of the involution η_λ discussed above is given in [**19**] in a different model for characters, which is based on *Lusztig's parametrization* and the *string parametrization* of the dual canonical basis [**2**]. Unlike the combinatorial approach in Schützenberger's evacuation procedure, the involution is now expressed as an affine map whose coefficients are entries of the corresponding Cartan matrix. No intrinsic explanation for the fact that this map is an involution is available.

We believe that the properties of our model that were investigated in [**12, 13**] as well as in this paper represent just a small fraction of a rich combinatorial structure yet to be explored, which would generalize most of the combinatorics of Young tableaux. A future publication will be concerned with the combinatorics of the product of crystals.

2. Preliminaries

We recall some background information on finite root systems, affine Weyl groups, and crystal graphs.

2.1. Root systems. Let G be a connected, simply connected, simple complex Lie group. Fix a Borel subgroup B and a maximal torus T such that $G \supset B \supset T$. As usual, we denote by B^- be the opposite Borel subgroup, while N and N^- are the unipotent radicals of B and B^-, respectively. Let $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{n}$, and $\mathfrak{n}^-$ be the complex Lie algebras of G, T, N, and N^-, respectively. Let r be the rank of the Cartan subalgebra $\mathfrak{h}$. Let $\Phi \subset \mathfrak{h}^*$ be the corresponding irreducible *root system*, and let $\mathfrak{h}^*_{\mathbb{R}} \subset \mathfrak{h}^*$ be the real span of the roots. Let $\Phi^+ \subset \Phi$ be the set of positive roots corresponding to our choice of B. Then Φ is the disjoint union of Φ^+ and $\Phi^- := -\Phi^+$. We write $\alpha > 0$ (respectively, $\alpha < 0$) for $\alpha \in \Phi^+$ (respectively, $\alpha \in \Phi^-$), and we define $\mathrm{sgn}(\alpha)$ to be 1 (respectively -1). We also use the notation $|\alpha| := \mathrm{sgn}(\alpha)\alpha$. Let $\alpha_1, \ldots, \alpha_r \in \Phi^+$ be the corresponding *simple roots*, which form a basis of $\mathfrak{h}^*_{\mathbb{R}}$. Let $\langle \cdot\,, \cdot \rangle$ denote the nondegenerate scalar product on $\mathfrak{h}^*_{\mathbb{R}}$ induced by the Killing form. Given a root α, the corresponding *coroot* is $\alpha^\vee := 2\alpha/\langle \alpha, \alpha\rangle$. The collection of coroots $\Phi^\vee := \{\alpha^\vee \mid \alpha \in \Phi\}$ forms the *dual root system.*

The *Weyl group* $W \subset \mathrm{Aut}(\mathfrak{h}^*_{\mathbb{R}})$ of the Lie group G is generated by the reflections $s_\alpha : \mathfrak{h}^*_{\mathbb{R}} \to \mathfrak{h}^*_{\mathbb{R}}$, for $\alpha \in \Phi$, given by $s_\alpha : \lambda \mapsto \lambda - \langle \lambda, \alpha^\vee\rangle\, \alpha$. In fact, the Weyl group W is generated by the *simple reflections* $s_1, \ldots, s_r$ corresponding to the simple roots $s_i := s_{\alpha_i}$, subject to the *Coxeter relations.* An expression of a Weyl group element w as a product of generators $w = s_{i_1} \cdots s_{i_l}$ which has minimal length is called a *reduced decomposition* for w; its length $\ell(w) = l$ is called the *length* of w. The Weyl group contains a unique *longest element* $w_\circ$ with maximal length $\ell(w_\circ) = \#\Phi^+$. For $u, w \in W$, we say that u *covers* w, and write $u \gtrdot w$, if $w = us_\beta$, for some $\beta \in \Phi^+$, and $\ell(u) = \ell(w) + 1$. The transitive closure "$>$" of the relation "$\gtrdot$" is called the *Bruhat order* on W.

The *weight lattice* Λ is given by

$$\Lambda := \{\lambda \in \mathfrak{h}^*_{\mathbb{R}} \mid \langle \lambda, \alpha^\vee\rangle \in \mathbb{Z} \text{ for any } \alpha \in \Phi\}. \tag{2.1}$$

The set Λ^+ of *dominant weights* is given by

$$\Lambda^+ := \{\lambda \in \Lambda \mid \langle \lambda, \alpha^\vee\rangle \geq 0 \text{ for any } \alpha \in \Phi^+\}.$$

Let $\rho := \frac{1}{2}\sum_{\beta\in\Phi^+} \beta$. The *height* of a coroot $\alpha^\vee \in \Phi^\vee$ is $\langle \rho, \alpha^\vee\rangle = c_1 + \cdots + c_r$ if $\alpha^\vee = c_1\alpha_1^\vee + \cdots + c_r\alpha_r^\vee$. Since we assumed that Φ is irreducible, there is a unique *highest coroot* $\theta^\vee \in \Phi^\vee$ that has maximal height. (In other words, $\theta^\vee$ is the highest root of the dual root system $\Phi^\vee$. It should not be confused with the coroot of the highest root of Φ.) We will also use the *Coxeter number*, that can be defined as $h := \langle \rho, \theta^\vee\rangle + 1$.

2.2. Affine Weyl groups. In this subsection, we remind a few basic facts about affine Weyl groups and alcoves, cf. Humphreys [**9**, Chaper 4] for more details.

Let W_{aff} be the *affine Weyl group* for the Langland's dual group $G^\vee$. The affine Weyl group W_{aff} is generated by the affine reflections $s_{\alpha,k} : \mathfrak{h}^*_{\mathbb{R}} \to \mathfrak{h}^*_{\mathbb{R}}$, for $\alpha \in \Phi$ and $k \in \mathbb{Z}$, that reflect the space $\mathfrak{h}^*_{\mathbb{R}}$ with respect to the affine hyperplanes $H_{\alpha,k} := \{\lambda \in \mathfrak{h}^*_{\mathbb{R}} \mid \langle \lambda, \alpha^\vee\rangle = k\}$. The hyperplanes $H_{\alpha,k}$ divide the real vector space $\mathfrak{h}^*_{\mathbb{R}}$ into open regions, called *alcoves.*

The *fundamental alcove* $A_\circ$ is given by

$$A_\circ := \{\lambda \in \mathfrak{h}^*_{\mathbb{R}} \mid 0 < \langle \lambda, \alpha^\vee\rangle < 1 \text{ for all } \alpha \in \Phi^+\}.$$

An important property of the affine Weyl group is that it acts simply transitively on the collection of all alcoves. This fact implies that, for any alcove A, there exists a unique element v_A of the affine Weyl group W_{aff} such that $v_A(A_\circ) = A$. Hence the map $A \mapsto v_A$ is a one-to-one correspondence between alcoves and elements of the affine Weyl group.

Recall that $\theta^\vee \in \Phi^\vee$ is the highest coroot. Let $\theta \in \Phi^+$ be the corresponding root, and let $\alpha_0 := -\theta$. The affine Weyl group is a Coxeter group generated by the set of reflections $s_0, s_1, \ldots, s_r$, where $s_0 := s_{\alpha_0,-1}$ and $s_1, \ldots, s_r \in W$ are the simple reflections $s_i = s_{\alpha_i,0}$.

We say that two alcoves A and B are *adjacent* if B is obtained by an affine reflection of A with respect to one of its walls. In other words, two alcoves are adjacent if they are distinct and have a common wall.

For a pair of adjacent alcoves, let us write $A \xrightarrow{\beta} B$ if the common wall of A and B is of the form $H_{\beta,k}$ and the root $\beta \in \Phi$ points in the direction from A to B.

2.3. Crystal graphs and Schützenberger's involution. Let $U(\mathfrak{g})$ be the universal enveloping algebra of the Lie algebra $\mathfrak{g}$, which is generated by E_i, F_i, H_i, for $i = 1, \ldots, r$, subject to the Serre relations and some additional relations. Let $\mathcal{B}$ be the *canonical basis* of $U(\mathfrak{n}^-)$, and let $\mathcal{B}_\lambda := \mathcal{B} \cap V_\lambda$ be the canonical basis of the irreducible representation V_λ with highest weight λ. Let v_λ and v_λ^{low} be the highest and lowest weight vectors in $\mathcal{B}_\lambda$, respectively. Let $\widetilde{E}_i$, $\widetilde{F}_i$, for $i = 1, \ldots, r$, be Kashiwara's operators [**10, 18**]; these are also known as raising and lowering operators, respectively. The *crystal graph* of V_λ is the directed colored graph on $\mathcal{B}_\lambda$ defined by arrows $x \to y$ colored i for each $\widetilde{F}_i(x) = cy + \text{lower terms}$, or, equivalently, for each $\widetilde{E}_i(y) = cx + \text{lower terms}$, with c a constant. (In fact, Kashiwara defined the crystal graph of the q-deformation of $U(\mathfrak{g})$, also known as a *quantum group*; using the quantum deformation, one can associate a crystal graph to a $\mathfrak{g}$-representation.) One can also define partial orders $\preceq_i$ on $\mathcal{B}_\lambda$ by $x \preceq_i y$ if $x = \widetilde{F}_i^k(y)$ for some $k \geq 0$. We let $\preceq$ denote the partial order generated by all partial orders $\preceq_i$, for $i = 1, \ldots, r$. The poset $(\mathcal{B}_\lambda, \preceq)$ has maximum v_λ and minimum v_λ^{low}.

In order to proceed, we need the following general setup. Let V be a module over an associative algebra U and σ an automorphism of U. The twisted U-module V^σ is the same vector space V but with the new action $u * v := \sigma(u)v$ for $u \in U$ and $v \in V$. Clearly, $V^{\sigma\tau} = (V^\sigma)^\tau$ for every two automorphisms σ and τ of U. Furthermore, if V is a simple U-module, then so is V^σ. In particular, if $U = U(\mathfrak{g})$ and $V = V_\lambda$, then $(V_\lambda)^\sigma$ is isomorphic to $V_{\sigma(\lambda)}$ for some dominant weight $\sigma(\lambda)$. Thus there is an isomorphism of vector spaces $\sigma_\lambda \,:\, V_\lambda \to V_{\sigma(\lambda)}$ such that

$$\sigma_\lambda(uv) = \sigma(u)\sigma_\lambda(v)\,, \quad u \in U(\mathfrak{g})\,,\ v \in V_\lambda\,.$$

By Schur's lemma, σ_λ is unique up to a scalar multiple.

The longest Weyl group element $w_\circ$ defines an involution on the simple roots by $\alpha_i \mapsto \alpha_{i^*} := -w_\circ(\alpha_i)$. Consider the automorphisms of $U(\mathfrak{g})$ defined by

$$\phi(E_i) = F_i\,, \quad \phi(F_i) = E_i\,, \quad \phi(H_i) = -H_i\,, \tag{2.2}$$

$$\psi(E_i) = E_{i^*}\,, \quad \psi(F_i) = F_{i^*}\,, \quad \psi(H_i) = H_{i^*}\,, \tag{2.3}$$

and $\eta := \phi\psi$. Clearly, these three automorphisms together with the identity automorphism form a group isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. It also easily follows from (2.2)-(2.3) that

$$\phi(\lambda) = \psi(\lambda) = -w_\circ(\lambda)\,, \quad \eta(\lambda) = \lambda\,.$$

We can normalize each of the maps ϕ_λ, ψ_λ, and η_λ by the requirement that

$$\phi_\lambda(v_\lambda) = v_{-w_\circ(\lambda)}^{low}\,, \quad \psi_\lambda(v_\lambda) = v_{-w_\circ(\lambda)}\,, \quad \eta_\lambda(v_\lambda) = v_\lambda^{low}\,. \tag{2.4}$$

(Of course, we also set Id_λ to be the identity map on V_λ.) By [**18**, Proposition 21.1.2], cf. also [**1**, Proposition 7.1], we have the following result.

PROPOSITION 2.1. [**1, 18**] (1) *Each of the maps ϕ_λ and ψ_λ sends $\mathcal{B}_\lambda$ to $\mathcal{B}_{-w_\circ(\lambda)}$, while η_λ sends $\mathcal{B}_\lambda$ to itself.*

(2) *For every two (not necessarily distinct) elements σ, τ of the group $\{\mathrm{Id}, \phi, \psi, \eta\}$, we have $(\sigma\tau)_\lambda = \sigma_{\tau(\lambda)}\tau_\lambda$. In particular, the map η_λ is an involution.*

(3) *For every $i = 1, \ldots, r$, we have*

$$\phi_\lambda\widetilde{F}_i = \widetilde{E}_i\phi_\lambda\,, \quad \psi_\lambda\widetilde{F}_i = \widetilde{F}_{i^*}\psi_\lambda\,, \quad \eta_\lambda\widetilde{F}_i = \widetilde{E}_{i^*}\eta_\lambda\,. \tag{2.5}$$

In particular, the poset $(\mathcal{B}_\lambda, \preceq)$ is self-dual, and η_λ is the corresponding antiautomorphism.

Berenstein and Zelevinsky [**1**] showed that, in type A_{n-1} (that is, in the case of the Lie algebra $\mathfrak{sl}_n$), the operator η_λ is given by Schützenberger's *evacuation* procedure for semistandard Young tableaux (see e.g. [**6**]).

3. The Alcove Path Model

In this section, we recall the model for the irreducible characters of semisimple Lie algebras that we introduced in [**12, 13**]. We refer to these papers for more details, including the proofs of the results mentioned below. Although some of these results hold for infinite root systems (cf. [**13**]), the setup in this paper is that of a finite irreducible root system, as discussed in Section 2.

Our model is conveniently phrased in terms of several sequences, so let us mention some related notation. Given a totally ordered index set $I = \{i_1 < i_2 < \ldots < i_n\}$, a sequence $(a_{i_1}, a_{i_2}, \ldots, a_{i_n})$ is sometimes abbreviated to $\{a_j\}_{j\in I}$. We also let $[n] := \{1, 2, \ldots, n\}$.

3.1. λ-chains. The affine translations by weights preserve the set of affine hyperplanes $H_{\alpha,k}$. It follows that these affine translations map alcoves to alcoves. Let $A_\lambda = A_\circ + \lambda$ be the alcove obtained by the affine translation of the fundamental alcove $A_\circ$ by a weight $\lambda \in \Lambda$. Let v_λ be the corresponding element of W_{aff}, i.e,. v_λ is defined by $v_\lambda(A_\circ) = A_\lambda$. Note that the element v_λ may not be an affine translation itself.

Let us now fix a dominant weight λ. Let $v \mapsto \bar{v}$ be the homomorphism $W_{\mathrm{aff}} \to W$ defined by ignoring the affine translation. In other words, $\bar{s}_{\alpha,k} = s_\alpha \in W$.

Definition 3.1. A *λ-chain of roots* is a sequence of positive roots $(\beta_1, \ldots, \beta_n)$ which is determined as indicated below by a reduced decomposition $v_{-\lambda} = s_{i_1} \cdots s_{i_n}$ of $v_{-\lambda}$ as a product of generators of W_{aff}:

$$\beta_1 = \alpha_{i_1},\ \beta_2 = \bar{s}_{i_1}(\alpha_{i_2}),\ \beta_3 = \bar{s}_{i_1}\bar{s}_{i_2}(\alpha_{i_3}), \ldots,\ \beta_n = \bar{s}_{i_1} \cdots \bar{s}_{i_{n-1}}(\alpha_{i_n})\,.$$

When the context allows, we will abbreviate "λ-chain of roots" to "λ-chain". The *λ-chain of reflections* associated with the above λ-chain of roots is the sequence $(\widehat{r}_1, \ldots, \widehat{r}_n)$ of affine reflections in W_{aff} given by

$$\widehat{r}_1 = s_{i_1},\ \widehat{r}_2 = s_{i_1}s_{i_2}s_{i_1},\ \widehat{r}_3 = s_{i_1}s_{i_2}s_{i_3}s_{i_2}s_{i_1},\ \ldots,\ \widehat{r}_n = s_{i_1}\cdots s_{i_n}\cdots s_{i_1}.$$

We will present two equivalent definitions of a λ-chain of roots.

Definition 3.2. An *alcove path* is a sequence of alcoves $(A_0, A_1, \ldots, A_n)$ such that A_{i-1} and A_i are adjacent, for $i = 1, \ldots, n$. We say that an alcove path is *reduced* if it has minimal length among all alcove paths from A_0 to A_n.

Given a finite sequence of roots $\Gamma = (\beta_1, \ldots, \beta_n)$, we define the sequence of integers $(l_1^\emptyset, \ldots, l_n^\emptyset)$ by $l_i^\emptyset := \#\{j < i \mid \beta_j = \beta_i\}$, for $i = 1, \ldots, n$. We also need the following two conditions on Γ.

- (R1) The number of occurrences of any positive root α in Γ is $\langle \lambda, \alpha^\vee \rangle$.
- (R2) For each triple of positive roots (α, β, γ) with $\gamma^\vee = \alpha^\vee + \beta^\vee$, the subsequence of Γ consisting of α, β, γ is a concatenation of pairs (α, γ) and (β, γ) (in any order).

Theorem 3.3. [**12**] *The following statements are equivalent.*

- (a) *The sequence of roots $\Gamma = (\beta_1, \ldots, \beta_n)$ is a λ-chain, and $(\widehat{r}_1, \ldots, \widehat{r}_n)$ is the associated λ-chain of reflections.*
- (b) *We have a reduced alcove path $A_0 \xrightarrow{-\beta_1} \cdots \xrightarrow{-\beta_n} A_n$ from $A_0 = A_\circ$ to $A_n = A_{-\lambda}$, and $\widehat{r}_i$ is the affine reflection in the common wall of A_{i-1} and A_i, for $i = 1, \ldots, n$.*
- (c) *The sequence Γ satisfies conditions* (R1) *and* (R2) *above, and $\widehat{r}_i = s_{\beta_i, -l_i^\emptyset}$, for $i = 1, \ldots, n$.*

A particular choice of a λ-chain was described in [**13**].

3.2. Admissible subsets. For the remainder of this section, we fix a λ-chain $\Gamma = (\beta_1, \ldots, \beta_n)$. Let $r_i := s_{\beta_i}$. We now define the centerpiece of our combinatorial model for characters, which is our generalization of semistandard Young tableaux in type A.

Definition 3.4. An *admissible subset* is a subset of $[n]$ (possibly empty), that is, $J = \{j_1 < j_2 < \ldots < j_s\}$, such that we have the following saturated chain in the Bruhat order on W:

$$1 \lessdot r_{j_1} \lessdot r_{j_1}r_{j_2} \lessdot \ldots \lessdot r_{j_1}r_{j_2}\ldots r_{j_s}\,.$$

We denote by $\mathcal{A}(\Gamma)$ the collection of all admissible subsets corresponding to our fixed λ-chain Γ. Given an admissible subset J, we use the notation

$$\mu(J) := -\widehat{r}_{j_1}\ldots\widehat{r}_{j_s}(-\lambda)\,, \qquad w(J) := r_{j_1}\ldots r_{j_s}\,.$$

We call $\mu(J)$ the *weight* of the admissible subset J.

THEOREM 3.5. [**12, 13**] *We have the following character formula:*

$$ch(V_\lambda) = \sum_{J \in \mathcal{A}(\Gamma)} e^{\mu(J)}\,.$$

A more general Demazure character formula is also given in [**12**]. In addition to these character formulas, a Littlewood-Richardson rule for decomposing tensor products of irreducible representations is presented in terms of our model in [**13**].

EXAMPLE 3.6. Consider the Lie algebra $\mathfrak{sl}_3$ of type A_2. The corresponding root system Φ can be realized inside the vector space $V := \mathbb{R}^3/\mathbb{R}(1,1,1)$ as $\Phi = \{\alpha_{ij} := \varepsilon_i - \varepsilon_j \mid i \neq j,\ 1 \leq i,j \leq 3\}$, where $\varepsilon_1, \varepsilon_2, \varepsilon_3 \in V$ are the images of the coordinate vectors in $\mathbb{R}^3$. The reflection $s_{\alpha_{ij}}$ is denoted by s_{ij}. The simple roots are α_{12} and α_{23}, while $\alpha_{13} = \alpha_{12} + \alpha_{23}$ is the other positive root. Let $\lambda = \omega_1 = \varepsilon_1$ be the first fundamental weight. In this case, there is only one λ-chain $(\beta_1, \beta_2) = (\alpha_{12}, \alpha_{13})$. There are 3 admissible subsets: $\emptyset, \{1\}, \{1,2\}$. The subset $\{2\}$ is not admissible because the reflection s_{13} does not cover the identity element. We have $(l_1^{\emptyset}, l_2^{\emptyset}) = (0,0)$. Theorem 3.5 gives the following expression for the character of V_{ω_1}:

$$ch(V_{\omega_1}) = e^{\omega_1} + e^{s_{12}(\omega_1)} + e^{s_{12}s_{13}(\omega_1)}.$$

3.3. Root operators. We now define partial operators known as *root operators* on the collection $\mathcal{A}(\Gamma)$ of admissible subsets corresponding to our fixed λ-chain $\Gamma = (\beta_1, \ldots, \beta_n)$. They are associated with a fixed simple root α_p, and are traditionally denoted by F_p (also called a lowering operator) and E_p (also called a raising operator). The notation is the one introduced above. Recall the sequence of integers $(l_1^{\emptyset}, \ldots, l_n^{\emptyset})$ associated to Γ, and the corresponding affine reflections $\widehat{r}_i = s_{\beta_i, -l_i^{\emptyset}}$ for $i = 1, \ldots, n$. Let $J = \{j_1 < j_2 < \ldots < j_s\}$ be a fixed admissible subset. We associate with J the sequence of roots $(\gamma_1, \ldots, \gamma_n)$ and the sequence of integers $L(J) = (l_1, \ldots, l_n)$, as follows: given $i \in [n]$, we let $k := \max\{a \mid j_a < i\}$ and $\widehat{r}_{j_1} \ldots \widehat{r}_{j_k}(H_{\beta_i, -l_i^{\emptyset}}) = H_{\gamma_i, -l_i}$ for some positive root γ_i and some integer l_i. We also define $l_\infty^p := \langle \mu(J), \alpha_p^\vee \rangle$. Finally, we let

$$I(J,p) := \{i \in [n] \mid \gamma_i = \alpha_p\}\,, \quad L(J,p) := (\{l_i\}_{i \in I(J,p)},\, l_\infty^p)\,, \quad M(J,p) := \max L(J,p)\,. \tag{3.1}$$

It turns out that $M(J,p) \geq 0$.

We can define the root operator F_p on the admissible subset J whenever $M(J,p) > 0$. Let $m = m_F(J,p)$ be defined by

$$m_F(J,p) := \begin{cases} \min\{i \in I(J,p) \mid l_i = M(J,p)\} & \text{if this set is nonempty} \\ \infty & \text{otherwise}\,. \end{cases}$$

Let $k = k_F(J,p)$ be the predecessor of m in $I(J,p) \cup \{\infty\}$, which always exists. It turns out that $m \in J$ if $m \neq \infty$, but $k \notin J$ Finally, we set

$$F_p(J) := (J \setminus \{m\}) \cup \{k\}\,. \tag{3.2}$$

We showed in [**13**] that we have $\mu(F_p(J)) = \mu(J) - \alpha_p$.

Let us now define a partial inverse E_p to F_p. The operator E_p is defined on the admissible subset J whenever $M(J,p) > \langle \mu(J), \alpha_p^\vee \rangle$. Let $k = k_E(J,p)$ be defined by

$$k_E(J,p) := \max\{i \in I(J,p) \mid l_i = M(J,p)\}\,;$$

the above set turns out to be always nonempty. Let $m = m_E(J,p)$ be the successor of k in $I(J,p) \cup \{\infty\}$. It turns out that $k \in J$ but $m \notin J$. Finally, we set

$$E_p(J) := (J \setminus \{k\}) \cup (\{m\} \setminus \{\infty\})\,. \tag{3.3}$$

Similarly to Kashiwara's operators (see Subsection 2.3), the root operators above define a directed colored graph structure and a poset structure on the set $\mathcal{A}(\Gamma)$ of admissible subsets corresponding to a fixed λ-chain Γ. According to [**13**, Proposition 6.9]), the admissible subset $J_{\max} = \emptyset$ is the maximum of the poset $\mathcal{A}(\Gamma)$.

4. Yang-Baxter Moves

In this section, we define the analog of Schützenberger's sliding algorithm in our model, which we call a *Yang-Baxter move*, for reasons explained below. We start with some results on dihedral subgroups of Weyl groups.

4.1. Dihedral reflection subgroups. Let $\overline{W}$ be a dihedral Weyl group of order $2q$, that is, a Weyl group of type $A_1 \times A_1$, A_2, B_2, or G_2 (with $q = 2, 3, 4, 6$, respectively). Let $\overline{\Phi}$ be the corresponding root system with simple roots α, β. The sequence

$$\beta_1 := \alpha, \;\; \beta_2 := s_\alpha(\beta), \;\; \beta_3 := s_\alpha s_\beta(\alpha), \;\; \ldots, \;\; \beta_{q-1} := s_\beta(\alpha), \;\; \beta_q := \beta \tag{4.1}$$

is a *reflection ordering* on the positive roots of $\overline{\Phi}$ (cf. [**5**]). We present the Bruhat order on the Weyl group of type G_2 in Figure 1. Here, as well as throughout this paper, we label a cover $v \lessdot vs_\gamma$ in Bruhat order by the corresponding root γ.

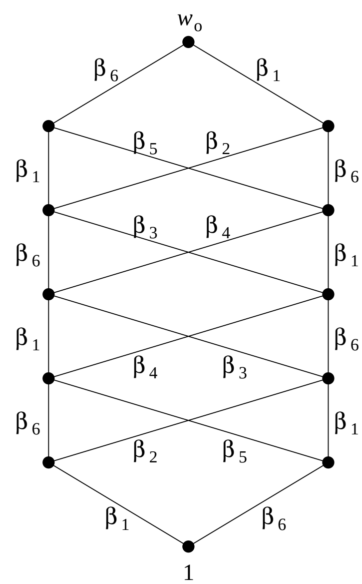


FIGURE 1. The Bruhat order on the Weyl group of type G_2.

With every pair of Weyl group elements $\overline{u} < \overline{w}$ in Bruhat order, we will associate a subset $J(\overline{u}, \overline{w})$ of $[q]$ as follows. Let $a := \ell(\overline{u})$ and $b := \ell(\overline{w})$. Given $\delta \in \{\alpha, \beta\}$, we will use the notation

$$\overline{W}_\delta := \{\overline{v} \in \overline{W} \mid \ell(\overline{v}s_\delta) > \ell(\overline{v})\}\,, \quad \overline{W}^\delta := \overline{W} \setminus \overline{W}_\delta = \{\overline{v} \in \overline{W} \mid \ell(\overline{v}s_\delta) < \ell(\overline{v})\}\,.$$

Case 0: $\overline{u} = \overline{w}$. We let $J(\overline{u}, \overline{u}) := \emptyset$.

Case 1: $b - a = 1$. We have the following disjoint subcases.

- **Case 1.1:** $\overline{u} \in \overline{W}_\alpha$, $\overline{w} \in \overline{W}^\alpha$, so $0 \le a \le q-1$. We let $J(\overline{u}, \overline{w}) := \{1\}$.
- **Case 1.2:** $\overline{u} \in \overline{W}^\beta$, $\overline{w} \in \overline{W}_\alpha$, so $0 < a < q-1$. We let $J(\overline{u}, \overline{w}) := \{q - a\}$.
- **Case 1.3:** $\overline{u} \in \overline{W}_\beta$, $\overline{w} \in \overline{W}^\beta$, so $0 \le a \le q-1$. We let $J(\overline{u}, \overline{w}) := \{q\}$.
- **Case 1.4:** $\overline{u} \in \overline{W}^\alpha$, $\overline{w} \in \overline{W}_\beta$, so $0 < a < q-1$. We let $J(\overline{u}, \overline{w}) := \{a+1\}$.

Case 2: $1 < b - a < q$. We have the following disjoint subcases.

- **Case 2.1:** $\overline{u} \in \overline{W}_\alpha$, $\overline{w} \in \overline{W}_\beta$, so $0 \le a < a+2 \le b < q$.
  We let $J(\overline{u}, \overline{w}) := \{1, a+2, a+3, \ldots, b\}$.
- **Case 2.2:** $\overline{u} \in \overline{W}^\beta$, $\overline{w} \in \overline{W}^\beta$, so $0 < a < a+2 \le b \le q$.
  We let $J(\overline{u}, \overline{w}) := \{1, a+2, a+3, \ldots, b-1, q\}$.
- **Case 2.3:** $\overline{u} \in \overline{W}_\beta$, $\overline{w} \in \overline{W}_\alpha$, so $0 \le a < a+2 \le b < q$.
  We let $J(\overline{u}, \overline{w}) := \{a+1, a+2, \ldots, b-1, q\}$.

Case 2.4: $\overline{u} \in \overline{W}^{\alpha}$, $\overline{w} \in \overline{W}^{\alpha}$, so $0 < a < a+2 \le b \le q$.
We let $J(\overline{u}, \overline{w}) := \{a+1, a+2, \dots, b\}$.

Case 3: $a = 0$ and $b = q$, that is, $\overline{u}$ is the identity and $\overline{w}$ is the longest Weyl group element $\overline{w}_\circ$. In this case, we let $J := [q]$.

In Case 2.2, if $b = a+2$ then the sequence $a+2, a+3, \dots, b-1$ is considered empty.

Let $J(\overline{u}, \overline{w}) := \{j_1 < j_2 < \dots < j_{b-a}\}$. We use the notation $r_i := s_{\beta_i}$, as above. It is easy to check that, in all cases above, we have a unique saturated increasing chain in Bruhat order from $\overline{u}$ to $\overline{w}$ whose labels form a subsequence of (4.1); this chain is

$$\overline{u} \lessdot \overline{u} r_{j_1} \lessdot \overline{u} r_{j_1} r_{j_2} \lessdot \dots \lessdot \overline{u} r_{j_1} \dots r_{j_{b-a}} = \overline{w}\,.$$

More generally, we have the result below for an arbitrary Weyl group W with a dihedral reflection subgroup $\overline{W}$ and corresponding root systems $\Phi \supseteq \overline{\Phi}$. The notation is the same as above. It is known that any element w of W can be written uniquely as $w = \lfloor w \rfloor \overline{w}$, where $\lfloor w \rfloor$ is the minimal representative of the left coset $w\overline{W}$, and $\overline{w} \in \overline{W}$. The following result can be easily deduced from the corresponding one for $W = \overline{W}$ via a standard fact about cosets modulo dihedral reflection subgroups, namely [**3**, Lemma 5.1].

PROPOSITION 4.1. *For each pair of elements $u < w$ in the same (left) coset of W modulo $\overline{W}$, we have a unique saturated increasing chain in Bruhat order from u to w whose labels form a subsequence of* (4.1)*; this chain is*

$$u \lessdot u r_{j_1} \lessdot u r_{j_1} r_{j_2} \lessdot \dots \lessdot u r_{j_1} \dots r_{j_{b-a}} = w\,,$$

where $J(\overline{u}, \overline{w}) := \{j_1 < j_2 < \dots < j_{b-a}\}$.

We obtain another reflection ordering by reversing the sequence (4.1). Let us denote the corresponding subset of $[q]$ by $J'(\overline{u}, \overline{w})$. We are interested in passing from the chain between u and w compatible with the ordering (4.1) to the chain compatible with the reverse ordering. If we fix $a := \ell(\overline{u})$ and $b := \ell(\overline{w})$, we can realize the passage from $J(\overline{u}, \overline{w})$ to $J'(\overline{u}, \overline{w})$ via the involution $Y_{q,a,b}$ described below in each of the cases mentioned above.

Case 0: $\emptyset \leftrightarrow \emptyset$ if $a = b$.

Case 1.1: $\{1\} \leftrightarrow \{q\}$ if $0 \le a = b-1 \le q-1$.

Case 1.2: $\{q-a\} \leftrightarrow \{a+1\}$ if $0 < a = b-1 < q-1$.

Case 2.1: $\{1, a+2, a+3, \dots, b\} \leftrightarrow \{a+1, a+2, \dots, b-1, q\}$ if $0 \le a < a+2 \le b < q$.

Case 2.2: $\{1, a+2, a+3, \dots, b-1, q\} \leftrightarrow \{a+1, a+2, \dots, b\}$ if $0 < a < a+2 \le b \le q$.

Case 3: $[q] \leftrightarrow [q]$ if $a = 0$ and $b = q$.

4.2. Yang-Baxter moves and their properties. Let us now consider an index set

$$I := \{\overline{1} < \dots < \overline{t} < 1 < \dots < q < \overline{t+1} < \dots < \overline{n}\}\,, \tag{4.2}$$

and let $\overline{I} := \{\overline{1}, \dots, \overline{n}\}$. Let $\Gamma = \{\beta_i\}_{i \in I}$ be a λ-chain, denote $r_i := s_{\beta_i}$ as before, and let $\Gamma' = \{\beta'_i\}_{i \in I}$ be the sequence of roots defined by

$$\beta'_i = \begin{cases} \beta_{q+1-i} & \text{if } i \in I \setminus \overline{I} \\ \beta_i & \text{otherwise}. \end{cases} \tag{4.3}$$

In other words, the sequence Γ' is obtained from the λ-chain Γ by reversing a certain segment. Now assume that $\{\beta_1, \dots, \beta_q\}$ are the positive roots of a rank two root system $\overline{\Phi}$ (without repetition). Let $\overline{W}$ be the corresponding dihedral reflection subgroup of the Weyl group W. The following result is easily proved using the correspondence between λ-chains and reduced words for the affine Weyl group element $v_{-\lambda}$ mentioned in Definition 3.1; most importantly, we need to recall from the proof of [**12**, Lemma 9.3] that the moves $\Gamma \to \Gamma'$ correspond to Coxeter moves (on the mentioned reduced words) in this context.

PROPOSITION 4.2. (1) *The sequence Γ' is also a λ-chain, and the sequence $\beta_1, \dots, \beta_q$ is a reflection ordering.*

(2) *We can obtain any λ-chain for a fixed dominant weight λ from any other λ-chain by moves of the form $\Gamma \to \Gamma'$.*

Let us now map the admissible subsets in $\mathcal{A}(\Gamma)$ to those in $\mathcal{A}(\Gamma')$. Given $J \in \mathcal{A}(\Gamma)$, let

$$\overline{J} := J \cap \overline{I}\,, \quad u := w(J \cap \{\overline{1}, \dots, \overline{t}\})\,, \quad \text{and} \quad w := w(J \cap (\{\overline{1}, \dots, \overline{t}\} \cup [q]))\,. \tag{4.4}$$

Also let

$$u = \lfloor u \rfloor \overline{u}\,, \quad w = \lfloor w \rfloor \overline{w}\,, \quad a := \ell(\overline{u})\,, \quad \text{and} \quad b := \ell(\overline{w})\,, \tag{4.5}$$

as above. It is clear that we have a bijection $Y : \mathcal{A}(\Gamma) \to \mathcal{A}(\Gamma')$ given by

$$Y(J) := \overline{J} \cup Y_{q,a,b}(J \setminus \overline{J})\,. \tag{4.6}$$

We call the moves $J \mapsto Y(J)$ *Yang-Baxter moves* (cf. the discussion following Theorem 4.1). Clearly, a Yang-Baxter move preserves the Weyl group element $w(\,\cdot\,)$ associated to an admissible subset, that is, $w(Y(J)) = w(J)$. In addition, Theorem 4.1 below holds.

THEOREM 4.1. *The map Y preserves the weight of an admissible subset. In other words, $\mu(Y(J)) = \mu(J)$ for all admissible subsets J.*

We now explain the way in which the Yang-Baxter moves are related to the *Yang-Baxter equation*, which justifies the terminology. In order to do this, we need to recall some information from [**12**]. Consider the ring $K := \mathbb{Z}[\Lambda/h] \otimes \mathbb{Z}[W]$, where $\mathbb{Z}[W]$ is the group algebra of the Weyl group W, and $\mathbb{Z}[\Lambda/h]$ is the group algebra of $\Lambda/h := \{\lambda/h \mid \lambda \in \Lambda\}$ (i.e., of the weight lattice shrunk h times, h being the Coxeter number defined in Subsection 2.1). We define $\mathbb{Z}[\Lambda/h]$-linear operators B_α and X^λ on K, where α is a positive root and λ is a weight:

$$B_\alpha : w \longmapsto \begin{cases} ws_\alpha & \text{if } \ell(ws_\alpha) = \ell(w) + 1 \\ 0 & \text{otherwise,} \end{cases} \qquad X^\lambda : w \mapsto e^{w(\lambda/h)} w.$$

Let us now consider the operators $R_\alpha := X^\rho (X^\alpha + B_\alpha) X^{-\rho}$ for $\alpha \in \Phi^+$; if $\alpha \in \Phi^-$, we define R_α by setting $B_\alpha := -B_{-\alpha}$. It was proved in [**12**, Theorem 10.1] that the operators $\{R_\alpha \mid \alpha \in \Phi\}$ satisfy the Yang-Baxter equation in the sense of Cherednik [**4**]. The main application of the operators R_α was to show that, given a λ-chain $\Gamma = (\beta_1, \dots, \beta_n)$, we have

$$R_{\beta_n} \dots R_{\beta_1}(1) = \sum_{J \in \mathcal{A}(\Gamma)} e^{\mu(J)} w(J)\,. \tag{4.7}$$

Due to the Yang-Baxter property, the right-hand side of the above formula does not change when we replace the λ-chain Γ by Γ', as defined above. The Yang-Baxter moves described above implement the passage from Γ to Γ' at the level of the individual terms in (4.7).

We now present the main result related to Yang-Baxter moves.

THEOREM 4.2. *The root operators commute with the Yang-Baxter moves, that is, a root operator F_p is defined on an admissible subset J if and only if it is defined on $Y(J)$ and we have*

$$Y(F_p(J)) = F_p(Y(J))\,.$$

Theorem 4.2 asserts that the map Y above is an isomorphism between $\mathcal{A}(\Gamma)$ and $\mathcal{A}(\Gamma')$ as directed colored graphs. Given two arbitrary λ-chains Γ and Γ', we know from Proposition 4.2 (2) that they can be related by a sequence of λ-chains $\Gamma = \Gamma_0, \Gamma_1, \dots, \Gamma_m = \Gamma'$ to which correspond Yang-Baxter moves $Y_1, \dots, Y_m$. Hence the composition $Y_m \dots Y_1$ is an isomorphism between $\mathcal{A}(\Gamma)$ and $\mathcal{A}(\Gamma')$ as directed colored graphs. Since every directed graph $\mathcal{A}(\Gamma)$ has a unique source (cf. [**13**, Proposition 6.9]), its automorphism group as a directed colored graph consists only of the identity. Thus, we have the following corollary of Theorem 4.2.

COROLLARY 4.3. *Given two arbitrary λ-chains Γ and Γ', the directed colored graph structures on $\mathcal{A}(\Gamma)$ and $\mathcal{A}(\Gamma')$ are isomorphic. This isomorphism is unique and, therefore, is given by the composition of Yang-Baxter moves corresponding to* any *sequence of λ-chains relating Γ and Γ'.*

We have thus given a combinatorial explanation for the independence of the directed colored graph defined by our root operators from the chosen λ-chain.

COROLLARY 4.4. *Given* any *λ-chain Γ, the directed colored graph on the set $\mathcal{A}(\Gamma)$ defined by the root operators is isomorphic to the crystal graph of the irreducible representation V_λ with highest weight λ. Under this isomorphism, the weight of an admissible subset gives the weight space in which the corresponding element of the canonical basis lies.*

The above result follows, based on Corollary 4.3, from its special case corresponding to the particular choice of a λ-chain Γ that was described in [**13**] and was mentioned in Subsection 3.1. Based on Corollary 4.4, we will now identify the elements of the canonical basis with the corresponding admissible subsets.

5. Schützenberger's Involution

In this section, we present an explicit description of the involution η_λ in Subsection 2.3 in the spirit of Schützenberger's evacuation. We will show that the role of jeu de taquin in the definition of the evacuation map is played by the Yang-Baxter moves.

Throughout the remainder of this paper, we fix an index set $I := \{\overline{1} < \ldots < \overline{q} < 1 < \ldots < n\}$ and a λ-chain $\Gamma = \{\beta_i\}_{i\in I}$ such that $l_i^\emptyset = 0$ if and only if $i \in \overline{I} := \{\overline{1} < \ldots < \overline{q}\}$. In other words, the second occurence of a root can never be before the first occurence of another root. We will also write $\Gamma := (\beta_{\overline{1}}, \ldots, \beta_{\overline{q}}, \beta_1, \ldots, \beta_n)$. Let us recall the notation $r_i := s_{\beta_i}$ for $i \in I$. It is easy to see that the set $J_{\min} := \overline{I}$ is the minimum of the poset $\mathcal{A}(\Gamma)$.

Given a Weyl group element w, we denote by $\lfloor w \rfloor$ and $\lceil w \rceil$ the minimal and the maximal representatives of the coset wW_λ, respectively (where W_λ is the stabilizer of the weight λ). Let $w_\circ^\lambda$ be the longest element of W_λ.

DEFINITION 5.1. Let J be an admissible subset. Let $J \cap \overline{I} = \{\overline{j}_1 < \ldots < \overline{j}_a\}$ and $J \setminus \overline{I} = \{j_1 < \ldots < j_s\}$. The *initial key* $\kappa_0(J)$ and the *final key* $\kappa_1(J)$ of J are the Weyl group elements defined by

$$\kappa_0(J) := r_{\overline{j}_1} \ldots r_{\overline{j}_a}\,, \qquad \kappa_1(J) := w(J) = \kappa_0(J) r_{j_1} \ldots r_{j_s}\,.$$

REMARK 5.2. The keys $\kappa_0(J)$ and $\kappa_1(J)$ are the analogs of the left and right keys of a semistandard Young tableau, as well as of the final and the initial directions of an LS path, respectively.

We associate with our fixed λ-chain Γ another sequence $\Gamma^{\mathrm{rev}} := \{\beta_i'\}_{i\in I}$ by

$$\beta_i' := \begin{cases} \beta_i & \text{if } i \in \overline{I} \\ w_\circ^\lambda(\beta_{n+1-i}) & \text{otherwise}\,. \end{cases}$$

In other words, we have

$$\Gamma^{\mathrm{rev}} = (\beta_{\overline{1}}, \ldots, \beta_{\overline{q}}, w_\circ^\lambda(\beta_n), w_\circ^\lambda(\beta_{n-1}), \ldots, w_\circ^\lambda(\beta_1))\,. \tag{5.1}$$

Let $r_i' := s_{\beta_i'}$ for $i \in I$. Fix an admissible subset

$$J = \{\overline{j}_1 < \ldots < \overline{j}_a < j_1 < \ldots < j_s\} \tag{5.2}$$

in $\mathcal{A}(\Gamma)$, where $\{\overline{j}_1 < \ldots < \overline{j}_a\} \subseteq \overline{I}$ and $\{j_1 < \ldots < j_s\} \subseteq I \setminus \overline{I}$. Let $u := \kappa_0(J)$ and $w := \kappa_1(J)$. We have the increasing saturated chain

$$1 \lessdot r_{\overline{j}_1} \lessdot r_{\overline{j}_1} r_{\overline{j}_2} \lessdot \ldots \lessdot r_{\overline{j}_1} \ldots r_{\overline{j}_a} = u \lessdot u r_{j_1} \lessdot u r_{j_1} r_{j_2} \lessdot \ldots \lessdot u r_{j_1} \ldots r_{j_s} = w\,. \tag{5.3}$$

According to [**5**], there is a unique saturated increasing chain in Bruhat order of the form

$$1 \lessdot r'_{\overline{k}_1} \lessdot r'_{\overline{k}_1} r'_{\overline{k}_2} \lessdot \ldots \lessdot r'_{\overline{k}_1} \ldots r'_{\overline{k}_b} = \lfloor w_\circ w \rfloor = w_\circ w w_\circ^\lambda\,,$$

where $\{\overline{k}_1 < \overline{k}_2 < \ldots < \overline{k}_b\} \subseteq \overline{I}$. Define

$$J^{\mathrm{rev}} := \{\overline{k}_1 < \ldots < \overline{k}_b < k_1 < \ldots < k_s\}\,, \tag{5.4}$$

where $k_i := n + 1 - j_{s+1-i}$ for $i = 1, \ldots, s$. Note that $\beta'_{k_i} = w_\circ^\lambda(\beta_{j_{s+1-i}})$ for $i = 1, \ldots, s$.

PROPOSITION 5.1. *Γ^{rev} is a λ-chain, and J^{rev} is an admissible subset in $\mathcal{A}(\Gamma^{\mathrm{rev}})$. We have*

$$\kappa_0(J^{\mathrm{rev}}) = \lfloor w_\circ \kappa_1(J) \rfloor\,, \quad \kappa_1(J^{\mathrm{rev}}) = \lfloor w_\circ \kappa_0(J) \rfloor\,, \quad \mu(J^{\mathrm{rev}}) = w_\circ(\mu(J))\,, \tag{5.5}$$

as well as $(J^{\mathrm{rev}})^{\mathrm{rev}} = J$.

We will now present the main result related to the map $J \mapsto J^{\mathrm{rev}}$, which involves its commutation with the root operators.

THEOREM 5.3. *A root operator F_p is defined on the admissible subset J if and only if E_{p^*} is defined on J^{rev}, and we have*

$$(F_p(J))^{\mathrm{rev}} = E_{p^*}(J^{\mathrm{rev}})\,.$$

We can summarize the construction so far as follows: given the λ-chain Γ (for a fixed dominant weight λ), we defined the λ-chain Γ^{rev}, and given $J \in \mathcal{A}(\Gamma)$, we defined $J^{\mathrm{rev}} \in \mathcal{A}(\Gamma^{\mathrm{rev}})$. Hence we can map J^{rev} to an admissible subset $J^* \in \mathcal{A}(\Gamma)$ using Yang-Baxter moves, as it is described in Section 4 and it is recalled below. To be more precise, let $R : \mathcal{A}(\Gamma) \to \mathcal{A}(\Gamma^{\mathrm{rev}})$ denote the bijection $J \mapsto J^{\mathrm{rev}}$. On the other hand, we know from Proposition 4.2 (2) that the λ-chains Γ^{rev} and Γ can be related by a sequence of λ-chains $\Gamma^{\mathrm{rev}} = \Gamma_0, \Gamma_1, \ldots, \Gamma_m = \Gamma$ to which correspond Yang-Baxter moves $Y_1, \ldots, Y_m$. By Corollary 4.3, the composition $Y := Y_m \ldots Y_1$ does not depend on the sequence of intermediate λ-chains, and it defines a bijection from $\mathcal{A}(\Gamma^{\mathrm{rev}})$ to $\mathcal{A}(\Gamma)$. We let $J^* := YR(J)$ and conclude that it is a bijection on $\mathcal{A}(\Gamma)$. The main result of this section, namely Theorem 5.4 below, now follows directly from Theorems 4.2 and 5.3.

THEOREM 5.4. *The bijection $J \mapsto J^*$ constructed above coincides with the fundamental involution η_λ on the canonical basis. In other words, a root operator F_p is defined on the admissible subset J if and only if E_{p^*} is defined on J^*, and we have*

$$(J_{\min})^* = J_{\max}\,, \quad (J_{\max})^* = J_{\min}\,, \quad \text{and} \quad (F_p(J))^* = E_{p^*}(J^*)\,, \quad \text{for } p = 1, \ldots, r\,. \tag{5.6}$$

In particular, the map $J \mapsto J^*$ expresses combinatorially the self-duality of the poset $\mathcal{A}(\Gamma)$.

REMARK 5.5. The above construction is analogous to the definition of Schützenberger's evacuation map (see, for instance, [**6**]). Below, we recall the three-step procedure defining this map and we discuss the analogy with our construction in the case of each step.

(1) REVERSE: We rotate a given semistandard Young tableau by 180°. This corresponds to reversing its word, which is similar to the procedure used to construct Γ^{rev} from Γ.
(2) COMPLEMENT: We complement each entry via the map $i \mapsto w_\circ(i)$, where $w_\circ$ is the longest element in the corresponding symmetric group. This corresponds to using $w_\circ$ for the arbitrary Weyl group in the definition (5.4) of J^{rev}.
(3) SLIDE: We apply jeu de taquin on the obtained skew tableau. This corresponds to the Yang-Baxter moves $Y_1, \ldots, Y_m$ discussed above.

EXAMPLE 5.6. Consider the Lie algebra $\mathfrak{sl}_3$ of type A_2, cf. Example 3.6. Consider the dominant weight $\lambda = 4\varepsilon_1 + 2\varepsilon_2$ and the following λ-chain:

$$\begin{array}{ccccccccc} & \overline{1} & \overline{2} & \overline{3} & 1 & 2 & 3 & 4 & 5 \\ \Gamma = (& \alpha_{12}, & \alpha_{13}, & \alpha_{23}, & \alpha_{13}, & \underline{\alpha_{12}}, & \alpha_{13}, & \underline{\alpha_{23}}, & \alpha_{13}\,)\,. \end{array}$$

Here we indicated the index corresponding to each root, using the notation above; more precisely, we have $I = \{\overline{1} < \overline{2} < \overline{3} < 1 < 2 < 3 < 4 < 5\}$ and $\overline{I} = \{\overline{1} < \overline{2} < \overline{3}\}$. By the defining relation (5.1), we have

$$\begin{array}{ccccccccc} & \overline{1} & \overline{2} & \overline{3} & 1 & 2 & 3 & 4 & 5 \\ \Gamma^{\mathrm{rev}} = (& \underline{\alpha_{12}}, & \alpha_{13}, & \alpha_{23}, & \alpha_{13}, & \underline{\alpha_{23}}, & \alpha_{13}, & \underline{\alpha_{12}}, & \alpha_{13}\,)\,. \end{array}$$

Consider the admissible subset $J = \{2, 4\}$. This is indicated above by the underlined roots in Γ. In order to define J^{rev}, cf. (5.4), we need to compute

$$\kappa_0(J^{\mathrm{rev}}) = w_\circ w(J) = (s_{12}s_{23}s_{12})(s_{12}s_{23}) = s_{12}\,.$$

Hence we have $J^{\mathrm{rev}} = \{\overline{1}, 2, 4\}$. This is indicated above by the underlined positions in Γ^{rev}.

In order to transform the λ-chain Γ^{rev} into Γ, we need to perform a single Yang-Baxter move; this consists of reversing the order of the bracketed roots below:

$$\begin{array}{cccccccccc} & \overline{1} & \overline{2} & \overline{3} & 1 & 2 & 3 & 4 & 5 & \\ \Gamma^{\mathrm{rev}} = (& \underline{\alpha_{12}}, & \alpha_{13}, & \alpha_{23}, & \alpha_{13}, & (\underline{\alpha_{23}}, & \alpha_{13}, & \underline{\alpha_{12}}), & \alpha_{13}) & \longrightarrow \\ & \overline{1} & \overline{2} & \overline{3} & 1 & 2 & 3 & 4 & 5 & \\ \Gamma \quad = (& \underline{\alpha_{12}}, & \alpha_{13}, & \alpha_{23}, & \alpha_{13}, & (\alpha_{12}, & \underline{\alpha_{13}}, & \underline{\alpha_{23}}), & \alpha_{13})\,. & \end{array}$$

The underlined roots indicate the way in which the Yang-Baxter move $J^{\mathrm{rev}} \mapsto Y(J^{\mathrm{rev}}) = J^*$ works. All we need to know is that there are two saturated chains in Bruhat order between the permutations u and w, cf. the notation in (4.4):

$$u = s_{12} \lessdot s_{12}s_{23} \lessdot s_{12}s_{23}s_{12} = w\,, \quad u = s_{12} \lessdot s_{12}s_{13} \lessdot s_{12}s_{13}s_{23} = w\,.$$

The first chain is retrieved as a subchain of Γ^{rev} and corresponds to J^{rev}, while the second one is retrieved as a subchain of Γ and corresponds to J^*. Hence we have $J^* = \{\overline{1}, 3, 4\}$.

We can give an intrinsic explanation for the fact that the map $J \mapsto J^*$ is an involution on $\mathcal{A}(\Gamma)$; this explanation is only based on the results in Sections 4 and 5, so it does not rely on Proposition 2.1 (2). Let us first recall the bijections $R : \mathcal{A}(\Gamma) \to \mathcal{A}(\Gamma^{\mathrm{rev}})$ and $Y : \mathcal{A}(\Gamma^{\mathrm{rev}}) \to \mathcal{A}(\Gamma)$ defined above. We claim that $YR = R^{-1}Y^{-1}$, which would prove that the composition YR is an involution. In the same way as we proved Theorem 5.4 (that is, as a direct consequence of Theorems 4.2 and 5.3), we can verify that the composition $R^{-1}Y^{-1}$ satisfies the conditions in (5.6). Since these conditions uniquely determine the corresponding map from $\mathcal{A}(\Gamma)$ to itself, our claim follows.

REMARK 5.7. According to the above discussion, we have a second way of realizing the fundamental involution η_λ on the canonical basis, namely as $R^{-1}Y^{-1}$. In some sense, this is the analog of the construction of the evacuation map based on the *promotion* operation (see, for instance, [**6**, p. 184]).

We have the following corollary of Proposition 5.1.

COROLLARY 5.8. *For any $J \in \mathcal{A}(\Gamma)$, we have*

$$\mu(J^*) = w_\circ(\mu(J))\,, \quad \kappa_0(J^*) = \lfloor w_\circ \kappa_1(J) \rfloor\,, \quad \kappa_1(J^*) = \lfloor w_\circ \kappa_0(J) \rfloor\,. \tag{5.7}$$

References

[1] A. Berenstein and A. Zelevinsky. Canonical bases for the quantum group of type A_r and piecewise-linear combinatorics. *Duke Math. J.*, 82:473–502, 1996.

[2] A. Berenstein and A. Zelevinsky. Tensor product multiplicities, canonical bases and totally positive varieties. *Invent. Math.*, 143:77–128, 2001.

[3] F. Brenti, S. Fomin, and A. Postnikov. Mixed Bruhat operators and Yang-Baxter equations for Weyl groups. *Internat. Math. Res. Notices*, 8:419–441, 1999.

[4] I. Cherednik. Quantum Knizhnik-Zamolodchikov equations and affine root systems. *Comm. Math. Phys.*, 150:109–136, 1992.

[5] M. J. Dyer. Hecke algebras and shellings of Bruhat intervals. *Compositio Math.*, 89(1):91–115, 1993.

[6] W. Fulton. *Young Tableaux*, volume 35 of *London Math. Soc. Student Texts*. Cambridge Univ. Press, Cambridge and New York, 1997.

[7] S. Gaussent and P. Littelmann. LS-galleries, the path model and MV-cycles. *Duke Math. J.*, 127:35–88, 2005.

[8] A. Henriques and J. Kamnitzer. Crystals and coboundary categories. `arXiv:math.QA/0406478`.

[9] J. E. Humphreys. *Reflection Groups and Coxeter Groups*, volume 29 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[10] M. Kashiwara. Crystalizing the q-analogue of universal enveloping algebras. *Commun. Math. Phys.*, 133:249–260, 1990.

[11] M. A. A. van Leeuwen. An analogue of jeu de taquin for Littelmann's crystal paths. *Sém. Loth. Comb.*, 41, Art. B41b, 23 pp., 1998.

[12] C. Lenart and A. Postnikov. Affine Weyl groups in K-theory and representation theory. `arXiv:math.RT/0309207`.

[13] C. Lenart and A. Postnikov. A combinatorial model for crystals of Kac-Moody algebras. `arXiv:math.RT/0502147`.

[14] P. Littelmann. A Littlewood-Richardson rule for symmetrizable Kac-Moody algebras. *Invent. Math.*, 116:329–346, 1994.

[15] P. Littelmann. Paths and root operators in representation theory. *Ann. of Math. (2)*, 142:499–525, 1995.

[16] P. Littelmann. Characters of representations and paths in $\mathfrak{h}^*_{\mathbb{R}}$. In *Representation Theory and Automorphic Forms (Edinburgh, 1996)*, volume 61 of *Proc. Sympos. Pure Math.*, pages 29–49. Amer. Math. Soc., Providence, RI, 1997.

[17] M. Lothaire. The plactic monoid (by A. Lascoux, B. Leclerc, and J-Y. Thibon). In *Algebraic Combinatorics on Words*, pages 144–172. Cambridge University Press, Cambridge, 2002.

[18] G. Lusztig. *Introduction to Quantum Groups*, volume 110 of *Progress in Mathematics*. Birkhäuser, Boston, 1993.

[19] S. Morier-Genoud. Relèvement géométrique de la base canonique et involution de Schützenberger. (French) [Geometrical lifting of the canonical base and Schützenberger involution]. *C. R. Math. Acad. Sci. Paris*, 337:371–374, 2003.

DEPARTMENT OF MATHEMATICS AND STATISTICS, STATE UNIVERSITY OF NEW YORK, ALBANY, NY 12222

E-mail address: `lenart@albany.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Ehrhart polynomials of lattice-face polytopes

Fu Liu

Abstract. There is a simple formula for the Ehrhart polynomial of a cyclic polytope. The purpose of this paper is to show that the same formula holds for a more general class of polytopes, lattice-face polytopes. We develop a way of decomposing any d dimensional simplex in general position into $d!$ signed sets, each of which corresponds to a permutation in the symmetric group S_d, and reduce the problem of counting lattice points in a polytope in general position to counting lattice points in these special signed sets. Applying this decomposition to a lattice-face simplex, we obtain signed sets with special properties that allow us to count the number of lattice points inside them. We are thus able to conclude the desired formula for the Ehrhart polynomials of lattice-face polytopes.

Résumé. Il y a une formule simple pour le polynôme d'Ehrhart d'un polytope cyclique. Le but de cet article est de prouver que la même formule est vraie pour une classe plus générale de polytope, les polytopes "treillis-faces". Nous donnons une manière de décomposer n'importe quel simplexe de dimension d en position générale en $d!$ ensembles signés. Chacun de ces ensembles correspond à une permutation dans le groupe symétrique S_d, et ramène le problème de compter des points de treillis dans un polytope en position générale à compter des points de treillis dans ces ensembles signés particuliers. Appliquant cette décomposition à un simplexe de treillis-faces, nous obtenons des ensembles signés dont les propriétés nous permettent de compter le nombre de points de treillis qu'ils contiennent. Nous obtenons ainsi la formule désirée pour les polynômes d'Ehrhart des polytopes de treillis-faces.

1. Introduction

A d-dimensional *lattice* $\mathbb{Z}^d = \{x = (x_1, \dots, x_d) \mid \forall x_i \in \mathbb{Z}\}$ is the collection of all points with integer coordinates in $\mathbb{R}^d$. Any point in a lattice is called a *lattice point*.

A *convex polytope* is a convex hull of a finite set of points. We often omit convex and just say polytope. For any polytope P and some positive integer $m \in \mathbb{N}$, we use $i(m, P)$ to denote the number of lattice points in mP, where $mP = \{mx | x \in P\}$ is the *mth dilated polytope* of P.

An *integral* or *lattice* polytope is a convex polytope whose vertices are all lattice points. Eugène Ehrhart [**4**] showed that for any d-dimensional integral polytope, $i(P, m)$ is a polynomial in m of degree d. Thus, we call $i(P, m)$ the Ehrhart polynomial of P when P is an integral polytope. Please see [**2, 3**] for more reference to the literature of lattice point counting. Although Ehrhart's theory was developed in 1960's, we still do not know much about the coefficients of Ehrhart polynomials for general polytopes except that the leading, second and last coefficients of $i(P, m)$ are the normalized volume of P, one half of the normalized volume of the boundary of P and 1, respectively.

In [**6**], the author showed that for any d-dimensional cyclic polytope P, we have that

$$i(P, m) = \mathrm{Vol}(mP) + i(\pi(P), m) = \sum_{k=0}^{d} \mathrm{Vol}_k(\pi^{(d-k)}(P)) m^k, \tag{1.1}$$

where $\pi^{(k)}$ is the map which ignores the last k coordinates of a point, and asked whether there are other integral polytopes that have the the same form of Ehrhart polynomials.

2000 *Mathematics Subject Classification.* Primary 05A19; Secondary 52B20.
Key words and phrases. Ehrhart polynomial, lattice-face, polytope, signed decomposition.

In this paper, we define a new family of integral polytopes, *lattice-face* polytopes, and show (Theorem 3.4) that their Ehrhart polynomials are in the form of (1.1).

The main method of [**6**] is a decomposition of an arbitrary d dimensional simplex cyclic polytope into $d!$ signed sets, each of which corresponds to a permutation in the symmetric group S_d and has the same sign as the corresponding permutation. However, for general polytopes, such a decomposition does not work.

In this paper, we develop a way of decomposing any d dimensional simplex in *general position* into $d!$ signed sets, where the sign of each set is not necessarily the same as the corresponding permutation. Applying the new decomposition to a lattice-face simplex, we are able to show (Theorem 3.5) that the number of lattice points in terms of a formula (6.1) involving Bernoulli polynomials, signs of permutations, and determinants, and then to analyze this formula further to derive the theorem. Theorem 3.5, together with some simple observation in section 2 and 3, imply Theorem 3.4.

2. Preliminaries

We first give some definitions and notations, most of which follows [**6**].

All polytopes we will consider are full-dimensional, so for any convex polytope P, we use d to denote both the dimension of the ambient space $\mathbb{R}^d$ and the dimension of P. We call a d-dimensional polytope a d-polytope. Also, We use ∂P and $I(P)$ to denote the boundary and the interior of P, respectively.

For any set S, we use $\operatorname{conv}(S)$ to denote the convex hull of all of points in S.

Recall that the projection $\pi : \mathbb{R}^d \to \mathbb{R}^{d-1}$ is the map that forgets the last coordinate. For any set $S \subset \mathbb{R}^d$ and any point $y \in \mathbb{R}^{d-1}$, let $\rho(y, S) = \pi^{-1}(y) \cap S$ be the intersection of S with the inverse image of y under π. Let $p(y, S)$ and $n(y, S)$ be the point in $\rho(y, S)$ with the largest and smallest last coordinate, respectively. If $\rho(y, S)$ is the empty set, i.e., $y \notin \pi(S)$, then let $p(y, S)$ and $n(y, S)$ be empty sets as well. Clearly, if S is a d-polytope, $p(y, S)$ and $n(y, S)$ are on the boundary of S. Also, we let $\rho^+(y, S) = \rho(y, S) \setminus n(y, S)$, and for any $T \subset \mathbb{R}^{d-1}$, $\rho^+(T, S) = \cup_{y \in T} \rho^+(y, S)$.

Definition 2.1. Define $PB(P) = \bigcup_{y \in \pi(P)} p(y, P)$ to be the *positive boundary* of P; $NB(P) = \cup_{y\in\pi(P)} n(y, P)$ to be the *negative boundary* of P and $\Omega(P) = P \setminus NB(P) = \rho^+(\pi(P), P) = \cup_{y \in \pi(P)} \rho^+(y, P)$ to be the *non-negative part* of P.

Definition 2.2. For any facet F of P, if F has an interior point in the positive boundary of P, then we call F a *positive facet* of P and define the sign of F as $+1$: $\operatorname{sign}(F) = +1$. Similarly, we can define the *negative facets* of P with associated sign -1. For the facets that are neither positive nor negative, we call them *neutral facets* and define the sign as 0.

It's easy to see that $F \subset PB(P)$ if F is a positive facet and $F \subset NB(P)$ if F is a negative facet.

Because the usual set union and set minus operation do not count the number of occurrences of an elements, which is important in our paper, from now on we will consider any polytopes or sets as *multisets* which allow *negative multiplicities*. In other words, we consider any element of a multiset as a pair (x, m), where m is the multiplicity of element x. Then for any multisets M_1, M_2 and any integers m, n and i, we define the following operators:

a) Scalar product: $iM_1 = i \cdot M_1 = \{(x, im) \mid (x, m) \in M_1\}$.
b) Addition: $M_1 \oplus M_2 = \{(x, m + n) \mid (x, m) \in M_1, (x, n) \in M_2\}$.
c) Subtraction: $M_1 \ominus M_2 = M_1 \oplus ((-1) \cdot M_2)$.

It's clear the following holds:

Lemma 2.3. *For any polytope* $P \subset \mathbb{R}^d$, $\forall R_1, \dots, R_k \subset \mathbb{R}^{d-1}$, $\forall i_1, \dots, i_k \in \mathbb{Z}$:

$$\rho^+ \left(\bigoplus_{j=1}^{k} i_j R_j,\ P \right) = \bigoplus_{j=1}^{k} i_j \rho^+(R_j,\ P).$$

Definition 2.4. We say a set S has *weight* w, if each of its elements has multiplicity either 0 or w. And S is a *signed* set if it has weight 1 or -1.

Let P be a convex polytope. For any y an interior point of $\pi(P)$, since π is a continuous open map, the inverse image of y contains an interior point of P. Thus $\pi^{-1}(y)$ intersects the boundary of P exactly twice. For any y a boundary point of $\pi(P)$, again because π is an open map, we have that $\rho(y, P) \subset \partial P$, so

$\rho(y,P) = \pi^{-1}(y) \cap \partial P$ is either one point or a line segment. The polytopes P we will be interested in are those satisfying $\rho(y,P)$ has has only one point.

Lemma 2.5. *If a polytope P satisfies:*

$$|\rho(y,P)| = 1, \forall y \in \partial\pi(P), \tag{2.1}$$

then P has the following properties:

(i) *For any $y \in I(\pi(P))$, $\pi^{-1}(y) \cap \partial P = \{p(y,P), n(y,P)\}$.*
(ii) *For any $y \in \partial\pi(P)$, $\pi^{-1}(y) \cap \partial P = \rho(y,P) = p(y,P) = n(y,P)$, so $\rho^+(y,P) = \emptyset$.*
(iii) *If $P = \bigsqcup_{i=1}^k P_i$, where the P_i's all satisfy (2.1), then $\Omega(P) = \bigoplus_{i=1}^k \Omega(P_i)$. ($P = \bigsqcup_{i=1}^k P_i$ means that P_i's give a decomposition of P, i.e., $P = \bigcup_{i=1}^k P_i$, and for any $i \neq j$, $P_i \cap P_j$ is contained in their boundaries.)*
(iv) *The set of facets of P are partitioned into the set of positive facets and the set of negative facets, i.e., there is no neutral facets.*

The proof of this lemma is straightforward, so we won't include it here.

The main purpose of this paper is to discuss the number of lattice points in a polytope. Therefore, for simplicity, for any set $S \in \mathbb{R}^d$, we denote by $\mathcal{L}(S) = S \cap \mathbb{Z}^d$ the set of lattice points in S. It's not hard to see that $\mathcal{L}$ commutes with some of the operations we defined earlier, e.g. ρ, ρ^+, Ω.

3. Lattice-face polytopes

A *d-simplex* is a polytope given as the convex hull of $d+1$ affinely independent points in $\mathbb{R}^d$.

Definition 3.1. We define *lattice-face* polytopes recursively. We call a one dimensional polytope a *lattice-face* polytope if it is integral.

For $d \geq 2$, we call a d-dimensional polytope P with vertex set V a *lattice-face* polytope if for any d-subset $U \subset V$,

a) $\pi(\mathrm{conv}(U))$ is a lattice-face polytope, and
b) $\pi(\mathcal{L}(H_U)) = \mathbb{Z}^{d-1}$, where H_U is the affine space spanned by U. In other words, after dropping the last coordinate of the lattice of H_U, we get the $(d-1)$-dimensional lattice.

To understand the definition, let's look at examples of 2-polytopes.

Example 3.2. Let P_1 be the polytope with vertices $v_1 = (0,0), v_2 = (2,0)$ and $v_3 = (2,1)$. Clearly, for any 2-subset U, condition *a*) is always satisfied. When $U = \{v_1, v_2\}$, H_U is $\{(x,0) \mid x \in \mathbb{R}\}$. So $\pi(\mathcal{L}(H_U)) = \mathbb{Z}$, i.e., *b*) holds. When $U = \{v_1, v_3\}$, H_U is $\{(x,y) \mid x = 2y\}$. Then $\mathcal{L}(H_U) = \{(2z,z) \mid z \in \mathbb{Z}\} \Rightarrow \pi(\mathcal{L}(H_U)) = 2\mathbb{Z} \neq \mathbb{Z}$. When $U = \{v_2, v_3\}$, H_U is $\{(2,y) \mid y \in \mathbb{R}\}$. Then $\pi(\mathcal{L}(H_U)) = \{2\} \neq \mathbb{Z}$. Therefore, P_1 is not a lattice-face polytope.

Let P_2 be the polytope with vertices $(0,0), (1,1)$ and $(2,0)$. One can check that P_2 is a lattice-face polytope.

Lemma 3.3. *Let P be a lattice-face d-polytope with vertex set V, then we have:*

(i) *$\pi(P)$ is a lattice-face $(d-1)$-polytope.*
(ii) *mP is a lattice-face d-polytope, for any positive integer m.*
(iii) *π induces a bijection between $\mathcal{L}(NB(P))$ and $\mathcal{L}(\pi(P))$.*
(iv) *Any d-subset U of V forms a $(d-1)$-simplex. Thus $\pi(\mathrm{conv}(U))$ is a $(d-1)$-simplex.*
(v) *Let H be the hyperplane determined by some d-subset of V. Then for any lattice point $y \in \mathbb{Z}^{d-1}$, we have that $\rho(y,H)$ is a lattice point.*
(vi) *P is an integral polytope.*

Proof. (i), (ii), (iii), (iv) and (v) are easy to prove. We prove (vi) by induction on d.

Any 1-dimensional lattice-face polytope is integral by definition.

For $d \geq 2$, suppose any $(d-1)$ dimensional lattice-face polytope is an integral polytope. Let P be a d dimensional lattice-face polytope with vertex set V. For any vertex $v_0 \in V$, let U be a subset of V that contains v_0. Let $U = \{v_0, v_1, \ldots, v_{d-1}\}$. We know that $P' = \pi(\mathrm{conv}(U))$ is a lattice-face $(d-1)$-simplex with vertices $\{\pi(v_0), \ldots, \pi(v_{d-1})\}$. Thus, by the induction hypothesis, P' is an integral polytope. In particular, $\pi(v_0)$ is a lattice point. Therefore, $v_0 = \rho(\pi(v_0), H)$ is a lattice point. ∎

The main theorem of this paper is to describe all of the coefficients of the Ehrhart polynomial of a lattice-face polytope.

Theorem 3.4. *Let P be a lattice-face d-polytope, then*

$$i(P,m) = \mathrm{Vol}(mP) + i(\pi(P),m) = \sum_{k=0}^{d} \mathrm{Vol}_k(\pi^{(d-k)}(P))m^k. \tag{3.1}$$

However, by Lemma 3.3/(ii),(iii), we have that

$$i(P,m) = |\mathcal{L}(\Omega(mP))| + i(\pi(P),m).$$

Therefore, to prove Theorem 3.4 it is sufficient to prove the following theorem:

Theorem 3.5. *For any P a lattice-face polytope,*

$$|\mathcal{L}(\Omega(P))| = \mathrm{Vol}(P).$$

Remark 3.6. We have an alternative definition of lattice-face polytopes, which is equivalent to Definition 3.1. Indeed, a d-polytope on a vertex set V is a lattice-face polytope if and only if for all k with $0 \le k \le d-1$,

$$\text{for any } (k+1)\text{-subset } U \subset V,\ \pi^{d-k}(\mathcal{L}(H_U)) \cong \mathbb{Z}^k, \tag{3.2}$$

where H_U is the affine space determined by U. In other words, after dropping the last $d-k$ coordinates of the lattice of H_U, we get the k-dimensional lattice.

4. A signed decomposition of the nonnegative part of a simplex in general position

The volume of a polytope is not very hard to characterize. So our main problem is to find the a way to describe the number of lattice points in the nonnegative part of a lattice-face polytope. We are going to do this via a signed decomposition.

4.1. Polytopes in general position. For the decomposition, we will work with a more general type of polytope (which contains the family of lattice-face polytopes).

Definition 4.1. We say that a d-polytope P with vertex set V is in *general position* if for any $k: 0 \le k \le d-1$, and any $(k+1)$-subset $U \subset V$, $\pi^{d-k}(\mathrm{conv}(U))$ is a k-simplex, where $\mathrm{conv}(U)$ is the convex hull of all of points in U.

It's easy to see that a lattice-face polytope is a polytope in general position. Therefore, the following discussion can be applied to lattice-face polytopes.

The following lemma states some properties of a polytope in general position. The proof is omitted.

Lemma 4.2. *Given a d-polytope P in general position with vertex set V, then*

- (i) *P satisfies (2.1).*
- (ii) *For any nonempty subset U of V, let $Q = \mathrm{conv}(U)$. If U has dimension $k (0 \le k \le d)$, then $\pi^{d-k}(Q)$ is a k-polytope in general position. In particular, for any facet F of P, $\pi(F)$ is a $(d-1)$-polytope in general position.*
- (iii) *For any triangulation of $P = \bigsqcup_{i=1}^k P_i$ without introducing new vertices, $\Omega(P) = \bigoplus_{i=1}^k \Omega(P_i)$. Thus, $\mathcal{L}(\Omega(P)) = \bigoplus_{i=1}^k \mathcal{L}(\Omega(P_i))$.*
- (iv) *For any hyperplane H determined by one facet of P and any $y \in \mathbb{R}^{d-1}$, $\rho(y,H)$ is one point.*

Remark 4.3. By (iii), and because $\mathrm{Vol}(\bigsqcup_{i=1}^k P_i) = \sum_{i=1}^k \mathrm{Vol}(P_i)$, to prove Theorem 3.5 it is sufficient to prove the case when P is a lattice-face simplex.

Therefore, we will only construct our decomposition in the case of simplices in general position. However, before the construction, we need one more proposition about the nonnegative part of a polytope in general position.

Proposition 4.4. *Let P be a d-polytope in general position with facets $F_1, F_2 \dots F_k$. Let H be the hyperplane determined by F_k. For $i: 1 \le i \le k$, let $F_i' = \pi^{-1}(\pi(F_i)) \cap H$ and $Q_i = \mathrm{conv}(F_i \cup F_i')$.*

Then

$$\Omega(P) = -\operatorname{sign}(F_k) \bigoplus_{i=1}^{k-1} \operatorname{sign}(F_i)\rho^+(\Omega(\pi(F_i)), Q_i). \tag{4.1}$$

The proof of Proposition 4.4 is similar to the proof of Proposition 2.6 in [**6**], so we do not include it here.

Now, we can use this proposition to inductively construct a decomposition of the nonnegative part $\Omega(P)$ of a d-simplex P in general position into $d!$ signed sets.

Decomposition of $\Omega(P)$:

- If $d = 1$, we do nothing: $\Omega(P) = \Omega(P)$.
- If $d \geq 2$, then by applying Proposition 4.4 to P and letting $k = d+1$, we have

$$\Omega(P) = -\operatorname{sign}(F_{d+1}) \bigoplus_{i=1}^{d} \operatorname{sign}(F_i)\rho^+(\Omega(\pi(F_i)), Q_i).$$

However, by Lemma 4.2/(ii), each $\pi(F_i)$ is a $(d-1)$-simplex in general position. By the induction hypothesis, $\Omega(\pi(F_i)) = \bigoplus_{j=1}^{(d-1)!} S_{i,j}$, where $S_{i,j}$'s are signed sets.

$$\rho^+(\Omega(\pi(F_i)), Q_i) = \rho^+(\bigoplus_{j=1}^{(d-1)!} S_{i,j}, Q_i) = \bigoplus_{j=1}^{(d-1)!} \rho^+(S_{i,j}, Q_i).$$

Since each $\rho^+(S_{i,j}, Q_i)$ is a signed set, we have decomposed $\Omega(P)$ into $d!$ signed sets.

Now we know that we can decompose $\Omega(P)$ into $d!$ signed sets. But we still need to figure out what these sets are and which signs they have. In the next subsection, we are going to discuss the sign of a facet of a d-simplex, which is going to help us determine the signs in our decomposition.

4.2. The sign of a facet of a d-simplex. From now on, we will always use the following setup for a d-simplex unless otherwise stated:

Suppose P is a d-simplex in general position with vertex set $V = \{v_1, v_2, \ldots, v_{d+1}\}$, where the coordinates of v_i are $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,d})$.

For any i, we denote by F_i the facet determined by vertices in $V \setminus \{v_i\}$ and H_i the hyperplane determined by F_i.

For any $\sigma \in S_d$ and $k : 1 \leq k \leq d$, we define matrices $X(\sigma, k)$ and $Y(\sigma, k)$ to be the matrices

$$X(\sigma,k) = \begin{pmatrix} 1 & x_{\sigma(1),1} & x_{\sigma(1),2} & \cdots & x_{\sigma(1),k} \\ 1 & x_{\sigma(2),1} & x_{\sigma(2),2} & \cdots & x_{\sigma(2),k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{\sigma(k),1} & x_{\sigma(k),2} & \cdots & x_{\sigma(k),k} \\ 1 & x_{d+1,1} & x_{d+1,2} & \cdots & x_{d+1,k} \end{pmatrix}_{(k+1)\times(k+1)},$$

$$Y(\sigma,k) = \begin{pmatrix} 1 & x_{\sigma(1),1} & x_{\sigma(1),2} & \cdots & x_{\sigma(1),k-1} \\ 1 & x_{\sigma(2),1} & x_{\sigma(2),2} & \cdots & x_{\sigma(2),k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{\sigma(k),1} & x_{\sigma(k),2} & \cdots & x_{\sigma(k),k-1} \end{pmatrix}_{k\times k}.$$

We also define $z(\sigma, k)$ to be

$$z(\sigma,k) = \det(X(\sigma,k))/\det(Y(\sigma,k)),$$

where $\det(M)$ is the determinant of a matrix M.

Now we can determine the sign of a facet F_i of P by looking at the determinants of these matrices, denoting by $\operatorname{sign}(x)$ the usual definition of sign of a real number x.

LEMMA 4.5. (i) *$\forall i : 1 \leq i \leq d$ and $\forall \sigma \in S_d$ with $\sigma(d) = i$,*

$$\operatorname{sign}(F_i) = \operatorname{sign}(\det(X(\sigma,d))/\det(X(\sigma,d-1))). \tag{4.2}$$

(ii) *When $i = d+1$ and for $\forall \sigma \in S_d$,*

$$\operatorname{sign}(F_{d+1}) = -\operatorname{sign}(\det(X(\sigma,d))/\det(Y(\sigma,d))) = -\operatorname{sign}(z(\sigma,d)). \tag{4.3}$$

PROOF. For any $i : 1 \leq i \leq d+1$, let $v_i' = \rho(\pi(v_i), H_i)$, i.e. v_i' is the unique point of the hyperplane spanned by F_i which has the same coordinates as v_i except for the last one. Suppose the coordinates of v_i' are $(x_{i,1}, \ldots, x_{i,d-1}, x'_{i,d})$. Then

$$\operatorname{sign}(F_i) = -\operatorname{sign}(x_{i,d} - x'_{i,d}).$$

$\forall i : 1 \leq i \leq d$ and $\forall \sigma \in S_d$ with $\sigma(d) = i$, because v_i' is in the hyperplane determined by F_i, we have that

$$\det\left(\begin{pmatrix} 1 & x_{\sigma(1),1} & \cdots & x_{\sigma(1),d-1} & x_{\sigma(1),d} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_{\sigma(d-1),1} & \cdots & x_{\sigma(d-1),d-1} & x_{\sigma(d-1),d} \\ 1 & x_{\sigma(d),1} & \cdots & x_{\sigma(d),d-1} & x'_{\sigma(d),d} \\ 1 & x_{d+1,1} & \cdots & x_{d+1,d-1} & x_{d+1,d} \end{pmatrix}\right) = 0.$$

Therefore,

$$\det(X(\sigma, d)) = (-1)^{2d+1}(x_{i,d} - x'_{i,d}) \det(X(\sigma, d-1)).$$

Thus,

$$\operatorname{sign}(\det(X(\sigma, d))/\det(X(\sigma, d-1))) = -\operatorname{sign}(x_{i,d} - x'_{i,d}) = \operatorname{sign}(F_i).$$

We can similarly prove the formula for $i = d+1$. □

4.3. Decomposition formulas. The following theorem describes the signed sets in our decomposition.

THEOREM 4.6. *Let P be a d-simplex with vertex set $V = \{v_1, v_2, \ldots, v_{d+1}\}$, where the coordinates of v_i are $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,d})$. For any $\sigma \in S_d$, and $k : 0 \leq k \leq d-1$, let $v_{\sigma,k}$ be the point with first k coordinates the same as v_{d+1} and affinely dependent with $v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}, v_{\sigma(k+1)}$. (Because P is in general position, one sees that there exists one and only one such point.) We also let $v_{\sigma,d} = v_{d+1}$. Then*

$$\Omega(P) = \bigoplus_{\sigma \in S_d} \operatorname{sign}(\sigma, P) S_\sigma,$$

where

$$\operatorname{sign}(\sigma, P) = \operatorname{sign}(\det(X(\sigma, d))) \prod_{i=1}^{d} \operatorname{sign}(z(\sigma, i)), \tag{4.4}$$

and

$$S_\sigma = \{s \in \mathbb{R}^d \mid \pi^{d-k}(s) \in \Omega(\pi^{d-k}(\operatorname{conv}(\{v_{\sigma,0}, \ldots, v_{\sigma,k}\})))\forall 1 \leq k \leq d\} \tag{4.5}$$

is a set of weight 1, i.e. a regular set.

Hence,

$$\mathcal{L}(\Omega(P)) = \bigoplus_{\sigma \in S_d} \operatorname{sign}(\sigma, P) \mathcal{L}(S_\sigma).$$

PROOF. Proof by induction. □

COROLLARY 4.7. *If P is a d-simplex in general position, then*

$$|\mathcal{L}(\Omega(P))| = \sum_{\sigma \in S_d} \operatorname{sign}(\sigma, P) |\mathcal{L}(S_\sigma)|. \tag{4.6}$$

Therefore, if we can calculate the number of lattice points in S_σ's, then we can calculate the number of lattice points in the nonnegative part of a d-simplex in general position. However, it's not so easy to find $|\mathcal{L}(S_\sigma)|$'s for an arbitrary polytope. But we can do it for any lattice-face d-polytope.

5. Lattice enumeration in S_σ and Bernoulli polynomials

In this section, we will count the number of lattice points in S_σ's when P is a lattice-face d-simplex. This calculation involves Bernoulli polynomials.

5.1. Counting lattice points in S_σ. We say a map from $\mathbb{R}^d \to \mathbb{R}^d$ is *lattice preserving* if it is invertible and it maps lattice points to lattice points. Clearly, given a lattice preserving map f, for any set $S \in \mathbb{R}^d$ we have that $|\mathcal{L}(S)| = |\mathcal{L}(f(S))|$.

Let P be a lattice face d-simplex with vertex set $V = \{v_1, \dots, v_{d+1}\}$, where we use the same setup as before for d-simplices.

Given any $\sigma \in S_d$, recall that S_σ is defined as in (4.5). To count the number of lattice points in S_σ, we want to find a lattice preserving affine transformation which simplifies the form of S_σ.

Before trying to find such a transformation, we will define more notations.

For any $\sigma \in S_d$, $k : 1 \le k \le d$ and $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$, we define matrix $\tilde{X}(\sigma, k; x)$ as

$$\tilde{X}(\sigma, k; x) = \begin{pmatrix} 1 & x_{\sigma(1),1} & x_{\sigma(1),2} & \cdots & x_{\sigma(1),k} \\ 1 & x_{\sigma(2),1} & x_{\sigma(2),2} & \cdots & x_{\sigma(2),k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{\sigma(k),1} & x_{\sigma(k),2} & \cdots & x_{\sigma(k),k} \\ 1 & x_1 & x_2 & \cdots & x_k \end{pmatrix}_{(k+1)\times(k+1)},$$

and for $j : 0 \le j \le k$, let $M(\sigma, k; j)$ be the minor of the matrix $\tilde{X}(\sigma, k; x)$ obtained by omitting the last row and the $(j+1)$th column. Then

$$\det(\tilde{X}(\sigma, k; x)) = (-1)^k (M(\sigma, k; 0) + \sum_{j=1}^{k} (-1)^j M(\sigma, k; j) x_j). \tag{5.1}$$

Note that $M(\sigma, k; k) = \det(Y(\sigma, k))$. Therefore,

$$\frac{\det(\tilde{X}(\sigma, k; x))}{\det(Y(\sigma, k))} = (-1)^k \frac{M(\sigma, k; 0)}{\det(Y(\sigma, k))} + \sum_{j=1}^{k-1} (-1)^{k+j} \frac{M(\sigma, k; j)}{\det(Y(\sigma, k))} x_j + x_k. \tag{5.2}$$

LEMMA 5.1. *Suppose P is a lattice-face d-simplex. $\forall \sigma \in S_d, \forall k : 1 \le k \le d$, and $\forall j : 0 \le j \le k-1$, we have that*

$$\frac{M(\sigma, k; j)}{\det(Y(\sigma, k))} \in \mathbb{Z}.$$

This lemma, as well as Lemma 5.6, can be directly derived from the definition of the lattice-face polytopes. We omit the proofs here.

Given this lemma, we have the following proposition.

PROPOSITION 5.2. *There exist a lattice-preserving affine transformation T_σ which maps $x = (x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ to*

$$(\frac{\det(\tilde{X}(\sigma, 1; x))}{\det(Y(\sigma, 1))}, \frac{\det(\tilde{X}(\sigma, 2; x))}{\det(Y(\sigma, 2))}, \dots, \frac{\det(\tilde{X}(\sigma, d; x))}{\det(Y(\sigma, d))}).$$

PROOF. Let $\alpha_\sigma = (-\frac{M(\sigma,1;0)}{\det(Y(\sigma,1))}, \frac{M(\sigma,2;0)}{\det(Y(\sigma,2))}, \dots, (-1)^d \frac{M(\sigma,d;0)}{\det(Y(\sigma,d))})$ and $M_\sigma = (m_{\sigma,j,k})_{d\times d}$, where

$$m_{\sigma,j,k} = \begin{cases} 1, & \text{if } j = k \\ 0, & \text{if } j > k\ . \\ (-1)^{k+j} \frac{M(\sigma,k;j)}{\det(Y(\sigma,k))} & \text{if } j < k \end{cases}$$

We define $T_\sigma : \mathbb{R}^d \to \mathbb{R}^d$ by mapping x to $\alpha_\sigma + xM_\sigma$. By (5.2),

$$\alpha_\sigma + xM_\sigma = (\frac{\det(\tilde{X}(\sigma, 1; x))}{\det(Y(\sigma, 1))}, \frac{\det(\tilde{X}(\sigma, 2; x))}{\det(Y(\sigma, 2))}, \dots, \frac{\det(\tilde{X}(\sigma, d; x))}{\det(Y(\sigma, d))}).$$

Also, because all of the entries in M_σ and α_σ are integers and the determinant of M_σ is 1, T_σ is lattice preserving. □

COROLLARY 5.3. *Given P a lattice-face polytope with vertex set $V = \{v_1, v_2, \dots, v_{d+1}\}$, we have that*

(i) *$\forall i : 1 \le i \le d$, the last $d + 1 - i$ coordinates of $T_\sigma(v_{\sigma(i)})$ are all zero.*

(ii) *$T_\sigma(v_{d+1}) = (z(\sigma, 1), z(\sigma, 2), \dots, z(\sigma, d))$.*

(iii) *Recall that for* $k : 0 \le k \le d-1$, $v_{\sigma,k}$ *is the point with first* k *coordinates the same as* v_{d+1} *and affinely dependent with* $v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}, v_{\sigma(k+1)}$. *Then the first* k *coordinates of* $T_\sigma(v_{\sigma,k})$ *are the same as* $T_\sigma(v_{d+1})$ *and the rest of the coordinates are zero.*

PROOF. (i) This follows from that fact that $\det(\tilde{X}(\sigma, k; x_{\sigma(i)})) = 0$ if $i \le k \le d$.

(ii) This follows from the fact that $\tilde{X}(\sigma, k; x_{d+1}) = X(\sigma, k)$ and $z(\sigma, k) = \det(X(\sigma, k))/\det(Y(\sigma, k))$.

(iii) Because for any $x \in \mathbb{R}^d$, the kth coordinate of T_σ only depends on the first k coordinates of x, $T_\sigma(v_{\sigma,k})$ has the same first k coordinates as $T_\sigma(v_{d+1})$. T_σ is an affine transformation. So $T_\sigma(v_{\sigma,k})$ is affinely dependent with $T_\sigma(v_{\sigma(1)}), T_\sigma(v_{\sigma(2)}), \dots, T_\sigma(v_{\sigma(k)}), T_\sigma(v_{\sigma(k+1)})$, the last $d-k$ coordinates of which are all zero. Therefore the last $d-k$ coordinates of $T_\sigma(v_{\sigma,k})$ are all zero as well. □

Recalling that $v_{\sigma,d} = v_{d+1}$, we are able to describe $T_\sigma(S_\sigma)$ now.

PROPOSITION 5.4. *Let* $\widehat{S}_\sigma = T_\sigma(S_\sigma)$. *Then*

$$s = (s_1, s_2, \dots, s_d) \in \widehat{S}_\sigma \Leftrightarrow \forall 1 \le k \le d, s_k \in \Omega(\mathrm{conv}(0, \frac{z(\sigma, k)}{z(\sigma, k-1)} s_{k-1})), \tag{5.3}$$

where by convention we let $z(\sigma, 0) = 1$ *and* $s_0 = 1$.

PROOF. This can be deduced from the fact that

$$\widehat{S}_\sigma = \{s \in \mathbb{R}^d \mid \pi^{d-k}(s) \in \Omega(\pi^{d-k}(\mathrm{conv}(\{\widehat{v}_{\sigma,0}, \dots, \widehat{v}_{\sigma,k}\})))\forall 1 \le k \le d\},$$

where $\widehat{v}_{\sigma,i} = (z(\sigma,1), \dots, z(\sigma, i), 0, \dots, 0)$, for $0 \le i \le d$. □

Because T_σ is a lattice preserving map, $|\mathcal{L}(S_\sigma)| = |\mathcal{L}(\widehat{S}_\sigma)|$. Hence, our problem becomes to find the number of lattice points in $\widehat{S}_\sigma$. However, $\widehat{S}_\sigma$ is much nicer than S_σ. Actually, we can give a formula to calculate all of the sets having the same shape as $\widehat{S}_\sigma$.

LEMMA 5.5. *Given real nonzero numbers* $b_0 = 1, b_1, b_2, \dots, b_d$, *let* $a'_k = b_k/b_{k-1}$ *and* $a_k = b_k/|b_{k-1}|, \forall k : 1 \le k \le d$. *Let* S *be the set defined by the following:*

$$s = (s_1, s_2, \dots, s_d) \in S \Leftrightarrow \forall 1 \le k \le d, s_k \in \Omega(\mathrm{conv}(0, a'_k s_{k-1})),$$

where s_0 *is set to* 1. *Then*

$$|\mathcal{L}(S)| = \sum_{s_1 \in \mathcal{L}(\Omega(\mathrm{conv}(0, a'_1)))} \sum_{s_2 \in \mathcal{L}(\Omega(\mathrm{conv}(0, a'_2 s_1)))} \cdots \sum_{s_d \in \mathcal{L}(\Omega(\mathrm{conv}(0, a'_d s_{d-1})))} 1.$$

In particular, if $b_d > 0$, *then*

$$|\mathcal{L}(S)| = \sum_{s_1=1}^{\overline{\lfloor a_1 \rfloor}} \sum_{s_2=1}^{\overline{\lfloor a_2 s_1 \rfloor}} \cdots \sum_{s_d=1}^{\overline{\lfloor a_d s_{d-1} \rfloor}} 1,$$

where for any real number x, $\lfloor x \rfloor$ *is the largest integer no greater than* x *and* $\overline{x}$ *is defined as*

$$\overline{x} = \begin{cases} x, & \text{if } x \ge 0 \\ -x-1, & \text{if } x < 0 \end{cases}.$$

PROOF. The first formula is straightforward. The second formula follows from the facts that for any real numbers x,

$$\mathcal{L}(\Omega(\mathrm{conv}(0, x))) = \begin{cases} \{z \in \mathbb{Z} \mid 1 \le z \le \overline{\lfloor x \rfloor}\} & \text{if } x \ge 0 \\ \{z \in \mathbb{Z} \mid -\overline{\lfloor x \rfloor} \le z \le 0\} & \text{if } x < 0 \end{cases},$$

the sign of s_i is the same as the sign of b_i, and because $b_d > 0$, all the s_i's are non-zero. □

However, for lattice-polytopes, we have another good property of the $z(\sigma, k)$'s.

LEMMA 5.6. *If* P *is a lattice-polytope* d*-simplex, then*

$$z(\sigma, k)/z(\sigma, k-1) \in \mathbb{Z}.$$

For any lattice-face d-simplex P, we can always find a way to order its vertices into $V = \{v_1, v_2, \dots, v_{d+1}\}$, so that the corresponding $\det(X(1,d))$ and $\det(Y(1,d))$ are positive, where 1 stands for the identity permutation in S_d. Note $z(\sigma, d)$ is independent of σ. So it is positive. Therefore, by Lemma 5.5 and Lemma 5.6, we have the following result.

PROPOSITION 5.7. *Let P be a lattice-face d-simplex with vertex set V, where the order of vertices makes both $\det(X(1,d))$ and $\det(Y(1,d))$ positive. Define*

$$a(\sigma, k) = \frac{z(\sigma,k)}{|z(\sigma,k-1)|}, \forall k : 1 \le k \le d.$$

Then

$$|\mathcal{L}(S_\sigma)| = \sum_{s_1=1}^{\overline{a(\sigma,1)}} \sum_{s_2=1}^{\overline{a(\sigma,2)s_1}} \cdots \sum_{s_d=1}^{\overline{a(\sigma,d)s_{d-1}}} 1. \tag{5.4}$$

Because of (5.4), it's natural for us to define

$$f_d(a_1, a_2, \dots, a_d) = \sum_{s_1=1}^{a_1} \sum_{s_2=1}^{a_2 s_1} \cdots \sum_{s_d=1}^{a_d s_{d-1}} 1, \tag{5.5}$$

for any positive integers $a_1, a_2, \dots, a_d$. However, since f_d is just a polynomial in the a_i's, we can extend the domain of f_d from $\mathbb{Z}^d_{>0}$ to $\mathbb{Z}^d$ or even $\mathbb{R}^d$. And for convenience, we still use the form of (5.5) to write $f_d(a_1, \dots, a_n)$ even when a_i's are not all positive integers.

Also, fixing $b_0 = 1$, we define

$$g_d(b_1, b_2, \dots, b_d) = f_d(b_1/b_0, b_2/b_1, \dots, b_d/b_{d-1}),$$

for any $(b_1, b_2, \dots, b_d) \in (\mathbb{R} \setminus \{0\})^d$.

f_d and g_d are closely related to formula (5.4). In next subsection, we will discuss Bernoulli polynomials and power sums, which are connected to f_d and g_d, and then rewrite (5.4) in terms of g_d. Please refer to [**3**, Section 2.4] for other examples about Bernoulli polynomials and their relation to lattice polytopes.

5.2. Power sums and Bernoulli polynomials. The *kth Bernoulli polynomials*, $B_k(x)$, is defined as [**1**, p. 804]

$$\frac{te^{tx}}{e^t - 1} = \sum_{k=0}^{\infty} B_k(x) \frac{t^k}{k!},$$

The Bernoulli polynomials satisfy [**5**]

$$B_k(1-x) = (-1)^k B_k(x), \forall k \ge 0, \tag{5.6}$$

as well as the relation [**8**, p. 127]

$$B_k(x+1) - B_k(x) = kx^{k-1}, \forall k \ge 1. \tag{5.7}$$

We call $B_k = B_k(0)$ a *Bernoulli number*. It satisfies [**7**] that

$$B_k(0) = 0, \text{for any odd number } k \ge 3. \tag{5.8}$$

For $k \ge 0$, let

$$S_k(x) = \frac{B_{k+1}(x+1) - B_{k+1}}{k+1}.$$

Given any n a nonnegative integer, by (5.7), we have that

$$S_k(n) = \sum_{i=0}^{n} i^k = \begin{cases} \sum_{i=1}^{n} i^k & \text{if } k \ge 1 \\ n+1 & \text{if } k = 0 \end{cases}.$$

Therefore, we call $S_k(x)$ the *kth power sum polynomial.*

LEMMA 5.8. *For any $k \ge 1$, the constant term of $S_k(x)$ is 0, i.e., x is a factor of $S_k(x)$, and*

$$S_k(x) = (-1)^{k+1} S_k(-x-1). \tag{5.9}$$

PROOF. The constant term of $S_k(x)$ is $S_k(0) = 0$. The formula follows from (5.6) and (5.8). ☐

Lemma 5.9. *$f_d(a_1,\dots,a_d)$ is a polynomial in a_1 of degree d. And $\prod_{i=1}^d a_i$ is a factor of it. In particular, f_d can be written as*

$$f_d(a_1,\dots,a_d)=\sum_{k=1}^{d} f_{d,k}(a_2,\dots,a_d)a_1^k, \tag{5.10}$$

where $f_{d,k}(a_2,\dots,a_d)$ is a polynomial in $a_2,\dots,a_d$ with a factor $\prod_{i=2}^d a_i$.

Proof. This can be proved by induction on d, using the fact that $S_k(x)$ has a factor x. $\square$

Proposition 5.10. *Given $s_0=1, a=(a_1,a_2,\dots,a_d)\in\mathbb{R}^d$, for any $j: 1\le j\le d-1$,*

$$f_d(a_1,a_2,\dots,a_d)=-\sum_{s_1=1}^{a_1s_0}\cdots\sum_{s_{j-1}=1}^{a_{j-1}s_{j-2}}\sum_{s_j=1}^{-a_js_{j-1}-1}\sum_{s_{j+1}=1}^{-a_{j+1}s_j}\sum_{s_{j+2}=1}^{a_{j+2}s_{j+1}}\cdots\sum_{s_d=1}^{a_ds_{d-1}}1.$$

Given $b=(b_1,b_2,\dots,b_d)\in(\mathbb{R}\setminus\{0\})^d$ with $b_d>0$, let $a_k=b_k/|b_{k-1}|$, then

$$g_d(b_1,b_2,\dots,b_d)=\operatorname{sign}\left(\prod_{i=1}^d b_i\right)\sum_{s_1=1}^{\overline{a_1}}\sum_{s_2=1}^{\overline{a_2s_1}}\cdots\sum_{s_d=1}^{\overline{a_ds_{d-1}}}1. \tag{5.11}$$

Proof. This follows from (5.9), (5.10) and an inductive argument. $\square$

Proposition 5.11. *Let P be a lattice-face d-simplex with vertex set V, where the order of vertices makes both $\det(X(1,d))$ and $\det(Y(1,d))$ positive. Then*

$$|\mathcal{L}(S_\sigma)|=\operatorname{sign}\left(\prod_{i=1}^d z(\sigma,i)\right)g_d(z(\sigma,1),z(\sigma,2),\dots,z(\sigma,d)). \tag{5.12}$$

Therefore,

$$|\mathcal{L}(\Omega(P))|=\sum_{\sigma\in S_d}\operatorname{sign}(\sigma)g_d(z(\sigma,1),z(\sigma,2),\dots,z(\sigma,d)). \tag{5.13}$$

Proof. We can get (5.12) by comparing (5.4) and (5.11). And (5.13) follows from (4.6), (4.4), (5.12) and the fact that $\det(X(\sigma,d))=\operatorname{sign}(\sigma)\det(X(1,d))$. $\square$

6. Proof of the Main Theorems

We now have all the ingredients but one to prove Theorem 3.5. The missing one is stated as the following proposition.

Proposition 6.1. *Let $V=\{v_1,v_2,\dots,v_{d+1}\}$ be the vertex set of a d-simplex in general position, where the coordinates of v_i are $x_i=(x_{i,1},x_{i,2},\dots,x_{i,d})$. Then*

$$\sum_{\sigma\in S_d}\operatorname{sign}(\sigma)g_d(z(\sigma,1),z(\sigma,2),\dots,z(\sigma,d))=\frac{1}{d!}\det(X(1,d)). \tag{6.1}$$

Given this proposition, we can prove Theorem 3.5.

Proof of Theorem 3.5. As we mentioned in Remark 4.3, to prove Theorem 3.5, it is sufficient to prove the case when P is a lattice-face simplex.

When P is a lattice-face d-simplex, we still assume that the order of the vertices of P makes both $\det(X(1,d))$ and $\det(Y(1,d))$ positive. Thus, (5.13), (6.1) and the fact that the volume of P is $\frac{1}{d!}|\det(X(1,d))|$ imply that

$$|\mathcal{L}(\Omega(P))|=\operatorname{Vol}(P).$$

$\square$

As we mentioned earlier, Theorem 3.4 follows from Theorem 3.5.

The proof of Proposition 6.1 is lengthy and self-contained, so we do not include it here.

7. Examples and Further discussion

7.1. Examples of lattice-face polytopes. In this subsection, we use a fixed family of lattice-face polytopes to illustrate our results. Let $d = 3$, and for any positive integer k, let P_k be the polytope with the vertex set $V = \{v_1 = (0,0,0), v_2 = (4,0,0), v_3 = (3,6,0), v_4 = (2,2,10k)\}$. One can check that P_k is a lattice-face polytope.

EXAMPLE 7.1 (Example of Theorem 3.4). The volume of P_k is $40k$, and

$$i(P_k, m) = 40km^3 + 12m^2 + 4m + 1.$$

$\pi(P_k) = \mathrm{conv}\{(0,0), (4,0), (3,6)\}$, where

$$i(\pi(P_k), m) = 12m^2 + 4m + 1.$$

So

$$i(P_k, m) = 40km^3 + i(\pi(P_k), m),$$

which agrees with Theorem 3.4.

EXAMPLE 7.2 (Example of Formula (4.1)). $F_4 = \mathrm{conv}(v_1, v_2, v_3)$ is a negative facet. The hyperplane determined by F_4 is $H = \{(x_1, x_2, x_3) \mid x_3 = 0\}$. Thus, $v_4' = \pi^{-1}(\pi(v_4)) \cap H = (2,2,0)$.

$F_3 = \mathrm{conv}(v_1, v_2, v_4)$ is a positive facet. $\pi(F_3) = \mathrm{conv}((0,0), (4,0), (2,2))$. $\Omega(\pi(F_3)) = \pi(F_3) \backslash \mathrm{conv}((0,0), (4,0))$. $F_3' = \pi^{-1}(\pi(F_3)) \cap H = \mathrm{conv}(v_1, v_2, v_4')$. So

$$Q_3 = \mathrm{conv}(F_3 \cup F_3') = \mathrm{conv}(v_1, v_2, v_4, v_4'),$$
$$\rho^+(\Omega(\pi(F_3)), Q_3) = Q_3 \setminus F_3'.$$

$F_2 = \mathrm{conv}(v_1, v_3, v_4)$ is a positive facet. $\pi(F_2) = \mathrm{conv}((0,0), (3,6), (2,2))$. $\Omega(\pi(F_2)) = \pi(F_2) \backslash (\mathrm{conv}((0,0), (2,2)) \cup \mathrm{conv}((2,2), (3,6)))$. $F_2' = \pi^{-1}(\pi(F_2)) \cap H = \mathrm{conv}(v_1, v_3, v_4')$. So

$$Q_2 = \mathrm{conv}(F_2 \cup F_2') = \mathrm{conv}(v_1, v_3, v_4, v_4'),$$
$$\rho^+(\Omega(\pi(F_2)), Q_2) = Q_2 \setminus (F_2' \cup \mathrm{conv}(v_1, v_4, v_4') \cup \mathrm{conv}(v_3, v_4, v_4')).$$

$F_1 = \mathrm{conv}(v_2, v_3, v_4)$ is a positive facet. $\pi(F_1) = \mathrm{conv}((4,0), (3,6), (2,2))$. $\Omega(\pi(F_1)) = \pi(F_1) \backslash \mathrm{conv}((4,0), (2,2))$. $F_1' = \pi^{-1}(\pi(F_1)) \cap H = \mathrm{conv}(v_2, v_3, v_4')$. So

$$Q_1 = \mathrm{conv}(F_1 \cup F_1') = \mathrm{conv}(v_2, v_3, v_4, v_4'),$$
$$\rho^+(\Omega(\pi(F_1)), Q_1) = Q_1 \setminus (F_1' \cup \mathrm{conv}(v_2, v_4, v_4')).$$

Therefore,

$$\Omega(P_k) = P_k \setminus F_4 = -\,\mathrm{sign}(F_4) \bigoplus_{i=1}^{3} \mathrm{sign}(F_i) \rho^+(\Omega(\pi(F_i)), Q_i),$$

which agrees with Proposition 4.4.

EXAMPLE 7.3 (Example of Decomposition). In this example, we decompose P_k into 3! sets, where 5 of them have positive signs and one has negative sign, which is different from the cases for cyclic polytopes, where half of the sets have positive signs and the other half have negative signs.

Recall that $v_{\sigma,3} = v_4 = (2,2,10k)$, for any $\sigma \in S_3$.

When $\sigma = 123 \in S_3$, $v_{123,2} = v_4' = (2,2,0)$, $v_{123,1} = (2,0,0)$ and $v_{123,0} = v_1 = (0,0,0)$. Then

$$S_{123} = \mathrm{conv}(\{v_{123,i}\}_{0 \le i \le 3}) \setminus \mathrm{conv}(\{v_{123,i}\}_{0 \le i \le 2}),$$

with $\mathrm{sign}(123, P_k) = +1$.

When $\sigma = 213 \in S_3$, $v_{213,2} = v_4' = (2,2,0)$, $v_{213,1} = (2,0,0)$ and $v_{213,0} = v_2 = (4,0,0)$. Then

$$S_{213} = \mathrm{conv}(\{v_{213,i}\}_{0 \le i \le 3}) \setminus (\mathrm{conv}(\{v_{213,i}\}_{0 \le i \le 2}) \cup \mathrm{conv}(\{v_{213,i}\}_{1 \le i \le 3})),$$

with $\mathrm{sign}(213, P_k) = +1$.

One can check that

$$S_{123} \oplus S_{213} = \rho^+(\Omega(\pi(F_3)), Q_3).$$

When $\sigma = 231 \in S_3$, $v_{231,2} = v_4' = (2,2,0)$, $v_{231,1} = (2,12,0)$ and $v_{231,0} = v_2 = (4,0,0)$. Then

$$S_{231} = \mathrm{conv}(\{v_{231,i}\}_{0 \le i \le 3}) \setminus (\mathrm{conv}(\{v_{231,i}\}_{0 \le i \le 2}) \cup \mathrm{conv}(\{v_{231,i}\}_{i=0,2,3} \cup \mathrm{conv}(\{v_{231,i}\}_{1 \le i \le 3})),$$

with $\mathrm{sign}(231, P_k) = +1$.

When $\sigma = 321 \in S_3$, $v_{321,2} = v'_4 = (2,2,0)$, $v_{321,1} = (2,12,0)$ and $v_{321,0} = v_3 = (3,6,0)$. Then

$$S_{321} = \operatorname{conv}(\{v_{321,i}\}_{0\leq i\leq 3}) \setminus (\operatorname{conv}(\{v_{321,i}\}_{0\leq i\leq 2}) \cup \operatorname{conv}(\{v_{321,i}\}_{i=0,2,3} \cup \operatorname{conv}(\{v_{321,i}\}_{1\leq i\leq 3})),$$

with $\operatorname{sign}(321, P_k) = -1$.

One can check that

$$S_{231} \ominus S_{321} = \rho^+(\Omega(\pi(F_1)), Q_1).$$

Similarly, we have that

$$S_{132} \oplus S_{312} = \rho^+(\Omega(\pi(F_2)), Q_2).$$

Therefore, $\Omega(P_k) = \bigoplus_{\sigma\in S_3} \operatorname{sign}(\sigma, P_k) S_\sigma$, which coincides with Theorem 4.6.

7.2. Further discussion. Recall that Remark 3.6 gives an alternative definition for lattice-face polytopes. Note in this definition, when $k = 0$, satisfying (3.2) is equivalent to say that P is an integral polytope, which implies that the last coefficient of the Ehrhart polynomial of P is 1. Therefore, one may ask

Question 7.4. If P is a polytope that satisfies (3.2) for all $k \in K$, where K is a fixed subset of $\{0, 1, \ldots, d-1\}$, can we say something about the Ehrhart polynomials of P?

A special set K can be chosen as the set of consecutive integers from 0 to d', where d' is an integer no greater than $d-1$. Based on some examples in this case, the Ehrhart polynomials seems to follow a certain pattern, so we conjecture the following:

Conjecture 7.5. *Given $d' \leq d-1$, if P is a d-polytope with vertex set V such that $\forall k : 0 \leq k \leq d'$, (3.2) is satisfied, then for $0 \leq k \leq d'$, the coefficient of m^k in $i(P, m)$ is the same as in $i(\pi^{d-d'}(P), m)$. In other words,*

$$i(P, m) = i(\pi^{d-d'}(P), m) + \sum_{i=d'+1}^{d} c_i m^i.$$

When $d' = 0$, the condition on P is simply that it is integral. And when $d' = d-1$, we are in the case that P is a lattice-face polytope. Therefore, for these two cases, this conjecture is true.

References

1. M. Abramowitz and I.A. Stegun, *Bernoulli and Eulear Polynomials and the Euler-Maclaurin Formula*, \$23.1 in Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. New York: Dover (1972), 804–806.
2. A. Barvinok, *Lattice points, polyhedra, and complexity*, Park City Math Institute Lecture Notes (Summer 2004), to appear.
3. M. Beck and S. Robins, *Computing the continuous discretely: Integer-point enumeration in polyhedra*, Springer, (to appear), Preprint at http://math.sfsu.edu/beck/papers/ccd.html.
4. E. Ehrhart, *Sur les polyèdres rationnels homothétiques à n dimensions*, C. R. Acad. Sci. Paris **254** (1962), 616–618.
5. D. H. Lehmer, *A new approach to Bernoulli polynomials*, Amer. Math. Monthly **95** (1988), 905–911.
6. F. Liu, *Ehrhart polynomials of cyclic polytopes*, Journal of Combinatorial Theory Ser. A **111** (2005), 111–127.
7. Eric W. Weisstein, *"Bernoulli number." from mathworld–a wolfram web resource*, http://mathworld.wolfram.com/BernoulliNumber.html.
8. E. T. Whittaker and G. N. Watson, *A course in modern analysis, 4th ed.*, Cambridge, England: Cambridge University Press, 1990.

Room 2-333, 77 Massachusetts Avenue, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

E-mail address: `fuliu@math.mit.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

On the chromatic symmetric function of a tree

Jeremy L. Martin and Jennifer D. Wagner

Abstract. Stanley defined the *chromatic symmetric function* $X(G)$ of a graph G as a sum of monomial symmetric functions corresponding to proper colorings of G, and asked whether a tree is determined up to isomorphism by its chromatic symmetric function. We approach Stanley's question by asking what invariants of a tree T can be recovered from its chromatic symmetric function $X(T)$. We prove that the *degree sequence* $(\delta_1, \dots)$, where δ_j is the number of vertices of T of degree j, and the *path sequence* $(\pi_1, \dots)$, where π_k is the number of k-edge paths in T, are given by explicit linear combinations of the coefficients of $X(T)$. These results are consistent with an affirmative answer to Stanley's question. We briefly present some applications of these results to classifying certain special classes of trees by their chromatic symmetric functions.

Résumé. Stanley a défini la *fonction symétrique chromatique* $X(G)$ d'un graphe G par une somme de fonctions symétriques monomials qui correspondent aux colorations propres de G, et il a demandé si un arbre est déterminé jusqu'à l'isomorphisme par sa fonction symétrique chromatique. Nous approchons la question de Stanley en demandant quels invariants d'un arbre T peut être récupéré de sa fonction symétrique chromatique $X(T)$. Nous prouvons que le *suite des degrés* $(\delta_1, \dots)$, où δ_j est le nombre des sommets de T de degré j, et le *suite des chemins* $(\pi_1, \dots)$, où π_k est le nombre de chemins de longueur k, sont données par des combinaisons lineaires explicites des coefficients $X(T)$. Ces résultats sont conformés à une réponse affirmative à la question de Stanley. Nous présentons brièvement quelques applications de ces résultats à classifier certaines classes spéciales des arbres par ses fonctions symétriques chromatiques.

Introduction

Let G be a simple graph with vertices $V(G)$ and edges $E(G)$, and let $n = \#V(G)$ (the *order* of G). We assume familiarity with standard facts about graphs and trees, as set forth in, e.g., [**11**, Chapters 1–2]. In particular, a *coloring* of G is a function $\kappa : V(G) \to \{1, 2, \dots\}$ such that $\kappa(v) \neq \kappa(w)$ whenever the vertices v, w are adjacent. Stanley [**7**] defined the *chromatic symmetric function* of G as

$$X(G) = X(G; x_1, x_2, \dots) \;=\; \sum_{\kappa} \prod_{v \in V(G)} x_{\kappa(v)},$$

the sum over all proper colorings κ, where $x_1, x_2, \dots$ are countably infinitely many commuting indeterminates. Note that $X(G)$ is homogeneous of degree n, and is invariant under permuting the x_i, so that $X(G)$ is a symmetric function. Moreover, the usual chromatic function $\chi(G; k)$, the number of colorings of G using at most k colors [**11**, Chapter 5], may be obtained from $X(G)$ by setting

$$x_i = \begin{cases} 1 & \text{for } i \leq k, \\ 0 & \text{for } i > k. \end{cases}$$

Our work is an attempt to resolve the following question.

Question (Stanley [**7**]): Is $X(G)$ a complete isomorphism invariant for trees? That is, must two nonisomorphic trees have different chromatic symmetric functions?

2000 *Mathematics Subject Classification.* Primary 05C05,05C60; Secondary 05E05.
Key words and phrases. tree, chromatic symmetric function, degree sequence, path sequence.

The answer to the question is "no" for arbitrary graphs; Stanley [**7**] exhibited two nonisomorphic graphs G, G' on 5 vertices such that $X(G) = X(G')$. For trees, however, the problem remains open. We note that Gebhard and Sagan [**2**] studied a chromatic symmetric function in *noncommuting* variables $x_1, x_2, \ldots$; this is easily seen to be a complete invariant of G. On the other hand, it is well-known (and elementary) that the chromatic function $\chi(G; k)$ is the same, namely $k(k-1)^{n-1}$, for all trees G on n vertices. Thus Stanley's question asks where $X(G)$ falls between these two extremes. Li-Yang Tan [**10**] has verified computationally[1] that $X(T)$ determines T up to isomorphism for all trees T of order ≤ 23.

Stanley showed that when $X(G)$ is expanded in the basis of power-sum symmetric functions p_λ (indexed by partitions λ), the coefficients c_λ enumerate the edge-selected subgraphs of G by the sizes of their components (see equations (1), (2) (3) below). With the additional assumption that G is a tree, this expansion is a powerful tool with which to recover the structure of G from $X(G)$. The first steps in this direction are due to Matthew Morin, who studied the chromatic symmetric functions of caterpillars (trees in which deleting all the leaves yields a path) in [**4, 5**].

We now summarize our results.

The *degree* $\deg_T(v)$ of a vertex v in a graph T is the number of edges having v as an endpoint, and the *degree sequence* of G is $(\delta_1, \delta_2, \ldots)$, where δ_j is the number of vertices having degree k. Our first main result is that the numbers δ_j are given by explicit linear combinations of the power-sum coefficients $c_\lambda(T)$.

Theorem 1. *For every tree T, we have $\delta_1(T) = c_{n-1}(T)$, and for all $j \geq 2$,*

$$\delta_j(T) \;=\; \sum_{\lambda \vdash n} \left(\ell(\tilde{\lambda}) \sum_{k \geq j} (-1)^{j+k-1} \binom{k}{j} \binom{\ell(\lambda) - 1}{k + \ell - n} \right) c_\lambda(T).$$

It is easier to compute directly the number s_k of subgraphs of T that are *k-edge stars*, or trees with one central vertex and k leaves. It is easily seen that the sequences $(s_1, s_2, \ldots)$ and $(\delta_1, \delta_2, \ldots)$ are linearly equivalent.

The *distance* between two vertices of T is the number of edges in the unique path connecting them. The *path sequence* of G is $(\pi_1, \pi_2, \ldots)$, where π_k is the number of vertex pairs at distance k, or equivalently the number of k-edge paths occurring as subgraphs of G. Our second main result, Theorem 2, asserts that the numbers π_k are again given by certain linear combinations of the coefficients $c_\lambda(T)$, as follows.

Theorem 2. *For every tree T, we have $\pi_1(T) = c_2(T)$ and $\pi_2(T) = c_3(T)$, and for all $k \geq 3$,*

$$\pi_k(T) \;=\; \sum_{\lambda \vdash n} \left((-1)^{n+k+1-\ell(\lambda)} \binom{\ell(\lambda) - 1}{k - n + \ell(\lambda)} m(\lambda) \right) c_\lambda(T),$$

where

$$m(\lambda) = \binom{n - \ell(\lambda)}{2} - \sum_{i=1}^{s} \binom{\lambda_i - 1}{2}.$$

To prove each of these theorems, we interpret the desired linear combination of the coefficients of $X(T)$ as generating functions for certain subgraphs of G, using Stanley's characterization of those coefficients. We then show that these labeled subgraphs admit a sign-reversing involution. The ensuing cancellation permits us to recognize the surviving terms as enumerating either stars or paths in G, as appropriate.

This extended abstract is organized as follows. Section 1 contains the elements of the theory of chromatic symmetric functions, as developed by Stanley in [**7**]. Sections 2 and 3 contain sketches of the proofs of the degree and path sequence theorems, respectively.

The final section contains some brief remarks about other isomorphism invariants that can be extracted from $X(T)$, and about some special classes of trees that can be distinguished up to isomorphism by their path and/or degree sequences (hence by their chromatic symmetric functions).

More details and applications will be found in a future paper written jointly by the present authors and Matthew Morin.

[1] In an earlier version of this extended abstract, it was mentioned that the present authors have checked this for trees of order ≤ 14, using the database of trees available online at `http://www.zis.agh.edu.pl/trees/`, generated by Piec, Malarz, and Kulakowski as described in [**6**]. Evidently Tan's result is a substantial improvement!

1. Basic properties of $X(G)$

We begin by reviewing some of the theory of chromatic symmetric functions developed by Stanley in [**7**]. A *partition* is a sequence of positive integers $\lambda = (\lambda_1, \dots, \lambda_s)$ with $\lambda_1 \geq \cdots \geq \lambda_s > 0$; the number $s = \ell(\lambda)$ is the *length* of λ. The corresponding *power-sum symmetric function* $p_\lambda = p_\lambda(x_1, x_2, \dots)$ is defined by $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_s}$, where $p_k = x_1^k + x_2^k + \cdots$.

One can obtain a family of useful invariants of G by expanding $X(G)$ in terms of the power sum symmetric functions p_λ. For each $S \subseteq E$, let $\lambda(S)$ be the partition of n whose parts are the orders of the components of the edge-induced subgraph $G|_S = (V, S)$. Stanley [**7**, Theorem 2.5] proved that

$$X(G) = \sum_{S \subset E} (-1)^{\# S} p_{\lambda(S)}. \tag{1}$$

In particular, the number of components of $G|_S$ is $\ell(\lambda(S))$. When $G = T$ is a tree, every subgraph S is a forest, so $\ell(\lambda(S)) = n - \# S$. Therefore, we may rewrite (1) as

$$X(T) = \sum_{\lambda \vdash n} (-1)^{n - \ell(\lambda)} c_\lambda p_\lambda, \tag{2}$$

where

$$c_\lambda = c_\lambda(T) = \#\{S \subset E \ : \ \lambda(S) = \lambda\}. \tag{3}$$

The coefficients $c_\lambda(T)$ are concrete combinatorial invariants of T that can be extracted from the chromatic symmetric function $X(T)$. Note that the c_λ are themselves not independent. For instance, it is immediate from (3) that

$$\sum_{\lambda:\ \ell(\lambda) = k} c_\lambda = \binom{n-1}{k}, \tag{4}$$

and there are several invariants of T that can be expressed in more than one distinct way in terms of the c_λ.

For notational simplicity, we shall often omit the parentheses and singleton parts when giving the index of one of these coefficients; for example, we abbreviate $c_{(h,1,1,\dots,1)}$ by c_h. (This raises the question of how we are going to denote the partition $\lambda = (1, 1, \dots, 1) = 1^n$. In fact, we won't need to do so, because 1^n is the only partition of length n, so (4) implies that $c_{1^n}(G) = 1$ for all G.)

For future reference, we list some properties of graphs and trees that can easily be read off its chromatic symmetric function. Several of these facts have already been noted by Morin [**4, 5**], and all of them are easy to deduce from (1) or (for trees) (2) and (3).

Proposition 3. *Let $G = (V, E)$ be a graph of order $n = \# V$.*

(i) *The number of vertices of G is the degree of $X(G)$.*
(ii) *The number of edges of G is c_2.*
(iii) *The number of components of G is* $\min\{\ell(\lambda) \mid c_\lambda(G) \neq 0\}$.
(iv) *If T is a tree, then the number of subtrees of T with k vertices is $c_k(T)$.*
(v) *If T is a tree, then the number of leaves (vertices of degree 1) in G is $c_{n-1}(T)$.*

Recall that a graph G is a tree if and only if it is connected and $\# E(G) = \# V(G) - 1$. Therefore, by (i), (ii) and (iii) of Proposition 3, the trees can be distinguished from other graphs by their chromatic symmetric functions. Moreover, part (v) implies that paths (trees with exactly two leaves) and stars (trees with exactly one nonleaf) are determined up to isomorphism by their chromatic symmetric functions.

2. The degree sequence

Let T be a tree with n vertices (and hence $n - 1$ edges). Recall that the degree $\deg_T(v)$ of a vertex $v \in V(T)$ is defined as the number of edges having v as an endpoint; a vertex of degree one is called a *leaf* of T.

Definition 4. The *degree sequence* of T is $(\delta_1, \delta_2, \dots, \delta_{n-1})$, where

$$\delta_j = \delta_j(T) = \#\{v \in V(T) \ : \ \deg_T(v) = j\}.$$

Notice that $\delta_j = 0$ whenever $j < 1$ or $j \geq n$. Moreover, it is a standard fact that $\sum \delta_j = 2n - 2$.

For $k \geq 1$, let S_k be the tree with vertices $\{0, 1, \ldots, k\}$ in which 0 is adjacent to every other vertex. Any graph that is isomorphic to S_k is called a *k-star*. If $k \geq 2$, then every k-star has a unique non-leaf vertex, called its *center*.

Definition 5. The *star sequence* of T is defined to be $(s_1, s_2, \ldots, s_{n-1})$, where

$$s_k = s_k(T) = \#\{U \subset T \;:\; U \equiv S_k\}.$$

Notice that $s_1 = n - 1$ (the number of edges of T), and that $s_k = 0$ whenever $k < 1$ or $k \geq n$.

Knowing the degree sequence of T is equivalent to knowing the number of substars of T of each possible order; it is straightforward to show that

$$s_k = \sum_{j \geq k} \binom{j}{k} \delta_j \tag{5a}$$

and that

$$\delta_j = \sum_{k \geq j} \binom{k}{j} (-1)^{j+k} s_k. \tag{5b}$$

It is more straightforward to recover the star sequence from the power-sum coefficients c_λ than it is to recover the degree sequence directly. For $\lambda \vdash n$, define $\ell(\lambda)$ to be the number of parts of λ, and let $\tilde{\lambda}$ be the partition obtained by deleting all the singleton parts of λ.

Theorem 6. *Let T be a tree with n vertices, and let $2 \leq k < n$. Then*

$$s_k(T) \;=\; -\sum_{\lambda \vdash n} \ell(\tilde{\lambda}) \binom{\ell(\lambda) - 1}{k + \ell(\lambda) - n} c_\lambda(T).$$

We sketch the proof, omitting many of the calculations and technical details. First, we obtain by straightforward calculation the identity

$$\sum_{\lambda \vdash n} \ell(\tilde{\lambda}) \binom{\ell(\lambda) - 1}{k + \ell(\lambda) - n} c_\lambda(T) \;=\; \sum_{\substack{F \subseteq T:\\ \#F = k}} \;\sum_{G \subseteq F} \;\sum_{\substack{\text{nontrivial}\\ \text{components}\\ C \text{ of } G}} (-1)^{\#G}. \tag{6}$$

For each subforest $F \subset T$, denote by $\Sigma(F)$ the summand indexed by F on the right-hand side of (6). The second step in the proof is to analyze $\Sigma(F)$. When F is a star, it is not hard to see that this summand reduces to

$$\left(\sum_{G \subseteq F} (-1)^{\#G} \right) - (-1)^{\#\emptyset} \;=\; -1.$$

Now, suppose that F is not a star; we wish to show that $\Sigma(F) = 0$. The expression $\Sigma(F)$ may be regarded as counting ordered pairs (G, C), where $G \subset F$ is a subforest and C is a nontrivial component of G, assigning to each such pair the weight $(-1)^{\#G}$. We construct an involution ψ on the set of such pairs (G, C). Whenever (G, C) and (G', C') are paired by ψ, we have $\#G' = \#G \pm 1$; in particular, the summands in $\Sigma(F)$ corresponding to (G, C) and (G', C') cancel. We conclude that $\Sigma(F) = 0$ as desired.

Theorem 6 now follows immediately from (6) together with the calculation of $\Sigma(F)$. The degree sequence formula, Theorem 1, follows in turn from Theorem 6 together with (5b).

3. The path sequence

Let $T = (V, E)$ be a tree with $\#V = n$. For any two vertices $v, w \in V$, their *distance* $d(v, w) = d_T(v, w)$ is defined as the number of edges in the unique path joining v and w. Define

$$\pi_k(T) \;:=\; \#\left\{\{v, w\} \subseteq V \;:\; d(v, w) = k\right\}.$$

Equivalently, $\pi_k(T)$ is the number of paths with exactly k edges that occur as subgraphs of T. It is easy to see that $\pi_k(T) = 0$ if $k \leq 0$ or $k \geq n$, and that $\sum_k \pi_k(T) = \binom{n}{2}$. Moreover, we have $\pi_1(T) = \#E = n - 1$ and $\pi_2(T) = s_2(T)$ (because a two-edge path is identical to a two-edge star). As we already know, these quantities can be recovered from $X(T)$.

Suppose that $k \geq 3$. We now recall Theorem 2, which describes the path numbers $\pi_k(T)$ as linear combinations of the coefficients of $X(T)$. For a partition $\lambda = (\lambda_1, \ldots, \lambda_\ell) \vdash n$, define

$$m(\lambda) = \binom{n - \ell(\lambda)}{2} - \sum_{i=1}^{s} \binom{\lambda_i - 1}{2}. \tag{7}$$

THEOREM 2. For every tree T, we have $\pi_1(T) = c_2(T)$ and $\pi_2(T) = c_3(T)$, and for all $k \geq 3$,

$$\pi_k(T) \;=\; \sum_{\lambda \vdash n} (-1)^{n+k+1-\ell(\lambda)} \binom{\ell(\lambda) - 1}{k - n + \ell(\lambda)} m(\lambda) c_\lambda(T).$$

Again, we give just a sketch of the proof. Using Stanley's interpretation for $c_\lambda(T)$, we can rewrite the right-hand side of the desired equality as

$$\sum_{A \subseteq E} (-1)^{k+1+\#A} \binom{n - \#A - 1}{k - \#A} m(\lambda(A)).$$

We interpret the binomial coefficient $\binom{n-a-1}{k-a}$ as counting the subsets of $E - A$ of cardinality $k - a$, and interpret $m(\lambda(A))$ as the number of pairs of distinct edges $e, f \in A$ that belong to different components of the induced subgraph (V, A); call such a pair of edges A-*okay*. Thus the last expression becomes

$$(-1)^{k+1} \sum_{A \subseteq E} \quad \sum_{\substack{B \subseteq E - A \\ \#B = k - \#A}} \quad \sum_{A\text{-okay pairs } e,f} (-1)^{\#A}.$$

For $e, f \in E$, let $P = P(e, f)$ be the unique shortest path between an endpoint of e and an endpoint of f. Then e, f is an A-okay pair if and only if $e, f \in A$ and $A \not\supseteq P$. In particular, e, f have no common endpoint (we abbreviate this condition as $e \cap f = \emptyset$), and $P \neq \emptyset$. Changing the order of summation and letting $A' = A - e - f$ and $C = A' \cup B$, we can rewrite the last expression as

$$(-1)^{k+1} \sum_{e \cap f = \emptyset} \quad \sum_{\substack{C \subseteq E - e - f \\ \#C = k-2}} \left(\sum_{\substack{A' \subseteq C \\ A' \not\supseteq P(e,f)}} (-1)^{\#A'} \right). \tag{8}$$

If we remove the condition $A' \not\supseteq P(e, f)$ from the last summation, then the parenthesized expression becomes zero (since $\#C = k - 2 > 0$). Therefore (8) can be rewritten as

$$(-1)^{k} \sum_{e \cap f = \emptyset} \quad \sum_{\substack{C \subseteq E - e - f \\ \#C = k-2}} \quad \sum_{\substack{A': \\ P(e,f) \subseteq A \subseteq C}} (-1)^{\#A'}. \tag{9}$$

The last sum is zero unless $C = P(e, f)$. So (9) collapses to

$$(-1)^k \sum_{e \cap f = \emptyset} \chi\Big[\#P(e,f) = k - 2\Big] (-1)^k \;=\; \sum_{e \cap f = \emptyset} \chi\Big[\#(e \cup f \cup P(e,f)) = k\Big] \;=\; \pi_k(T),$$

(where χ is the "Garsia chi": $\chi[S] = 1$ if the sentence S is true, or 0 if S is false). This completes the proof of Theorem 2.

4. Further remarks

4.1. Other invariants recoverable from $X(T)$. Theorems 1 and 2 imply that any isomorphism invariant of a tree T that can be derived from its path and degree sequences can be recovered from $X(T)$. Examples of such invariants include the *diameter* (the number of edges in a longest path) and the *Wiener index* (the quantity $\sigma(T) = \sum_{v,w} d(v, w)$, where v, w range over all pairs of vertices of T. The Wiener index can be obtained from the chromatic symmetric function in other ways. For example, when $X(T)$ is expanded as a sum of *elementary* symmetric functions, Stanley has interpreted the coefficients as counting sinks in acyclic orientations; this observation gives rise to a different expression for $\sigma(T)$. Note that the Wiener index is far from distinguishing trees up to isomorphism; see [**1**, §13].

One might ask whether the methods of Theorems 1 and 2 can be used to count other kinds of subtrees of a tree T (that is, other than stars and paths) by appropriate linear combinations of the coefficients of

$X(T)$. Such a class may be quite subtle; our empirical computations seem to rule out, for instance, spiders and double-stars (i.e., caterpillars with two branch vertices).

4.2. Spiders. Let T be a tree. A vertex $v \in V(T)$ is called a *branch vertex* if $\deg_T(v) \geq 3$. A *spider* (or *starlike tree*) is a tree with exactly one branch vertex (to avoid trivialities, we do not consider paths to be spiders). Since the definition of a spider relies only on the degree sequence, Theorem 1 implies that membership in the class of spiders can be deduced from $X(T)$. In fact, much more is true: one can show that every spider is determined up to isomorphism by its chromatic symmetric function.

We sketch the proof briefly. A spider may be regarded as a collection of edge-disjoint paths (the *legs*) joined at a common endpoint t (the *torso*). The torso is the unique branch vertex, and the lengths of the legs determine the spider up to isomorphism. That is, the isomorphism classes of spiders with n edges correspond to the partitions $\mu \vdash n$ with $\ell(\mu) \geq 3$. The partition μ can then be recovered from the coefficients $c_\lambda(T)$, where $\lambda \vdash n$ has exactly two parts. For example, when no single leg of the spider contains as many as half the edges, the sequence

$$(c_{1,n-1},\ c_{2,n-2},\ \dots);$$

is a partition whose conjugate is precisely μ. (The case of a spider with one "giant leg" is only slightly more complicated.)

4.3. Caterpillars. A *caterpillar* is a tree such that deleting all the leaves yields a path (called the *spine* of the caterpillar). It is not hard to see that this is equivalent to the condition that the diameter of T is one more than the number of nonleaf vertices; therefore, whether or not T is a caterpillar can be deduced from $X(T)$. When T is a *symmetric* caterpillar (i.e., it has an automorphism reversing the spine), it is determined up to isomorphism by $X(T)$. This fact was proved by Morin[**4**], and can also be recovered from Theorem 2. However, the corresponding statement for arbitrary caterpillars remains unknown. Gordon and McDonnell [**3**] showed that there exist arbitrarily large families of nonisomorphic caterpillars with the same path and degree sequences; however, we suspect that the additional information furnished by the chromatic symmetric function of a caterpillar T will be enough to reconstruct it up to isomorphism.

Acknowledgements

Chromatic symmetric functions came to the first author's attention thanks to Matthew Morin's talk at the Graduate Student Combinatorics Conference, held at the University of Minnesota on April 16 and 17, 2005. The authors are grateful to Matthew Morin, Victor Reiner, Bruce Sagan, and Stephanie van Willigenburg for numerous helpful discussions and suggestions.

References

[1] Andrey A. Dobrynin, Roger Entringer, and Ivan Gutman, Wiener index of trees: theory and applications, *Acta Appl. Math.* **66** (2001), no. 3, 211–249.

[2] David D. Gebhard and Bruce E. Sagan, A chromatic symmetric function in noncommuting variables, *J. Alg. Combin.* **2** (2001), 227–255.

[3] Gary Gordon and Eleanor McDonnell, Trees with the same degree sequence and path number, *Discrete Math.* **147** (1995), 297–300.

[4] Matthew Morin, Caterpillars, ribbons, and the chromatic symmetric function, M.S. thesis, University of British Columbia, 2005.

[5] Matthew Morin, The chromatic symmetric function of symmetric caterpillars, preprint, 2005.

[6] S. Piec, K. Malarz, and K. Kulakowski, How to count trees?, preprint, arXiv:cond-mat/0501594.

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

[8] Richard P. Stanley, *Enumerative Combinatorics*, Volume 2. Cambridge Univ. Press, Cambridge, 1997.

[9] Richard P. Stanley, Graph colorings and related symmetric functions: ideas and applications, *Discrete Math.* **193** (1998), 267–286.

[10] Li-Yang Tan, personal communication. Details of computation to be posted at `http://www.cs.wustl.edu/~lt1`.

[11] Douglas B. West, *Introduction to Graph Theory*, 2nd edn. Prentice Hall, Upper Saddle River, NJ, 2001.

Department of Mathematics, University of Kansas, Lawrence, KS 66045
E-mail address: `jmartin@math.ku.edu`
URL: `http://www.math.ku/edu/~jmartin`

E-mail address: `wagner@math.ku.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

An analogue of the Robinson-Schensted-Knuth Algorithm and its application to standard bases

Sarah Mason

Abstract. The Schur functions, $s_\lambda(x)$, form a basis for the vector space of symmetric functions. Recently Haglund, Haiman and Loehr derived a combinatorial formula for nonsymmetric Macdonald polynomials, which gives a new decomposition of the Macdonald polynomial into nonsymmetric components. Letting $q = t = 0$ in this identity implies $s_\lambda(x) = \sum_\mu NS_\mu(x)$, where the sum is over all rearrangements μ of the partition λ. We exhibit a bijection involving an analogue of Robinson-Schensted-Knuth Insertion between semi-standard Young tableaux and semi-standard skyline fillings to give a combinatorial proof of the formula. The insertion procedure led us to determine an analogue of the RSK Algorithm for semi-standard skyline fillings. This analogue is used to prove that the non-symmetric Schur functions equal the standard bases of Lascoux and Schützenberger.

Résumé. La fontion Schur forme une base pour l'espace de vecteur des fonctions symétriques. Récemment Haglund, Haiman et Loehr ont dérivé une formule combinatoire pour des polynômes nonsymmetric de Macdonald, qui donne une nouvelle décomposition du polynôme de Macdonald dans les composants nonsymmetric. Laisser $q = t = 0$ dans cette identité implique la fonction de Schur s_λ est la somme des fonctions nonsymmetric de Schur au-dessus de toutes les remises en ordre de la cloison λ. Nous exhibons un bijection impliquant un analogue d'insertion de Robinson-Schensted-Knuth entre de Young tableaux de semi-standard et remplissages d'horizon de semi-standard pour fournir des preuves combinatoires de la formule. Le procédé d'insertion nous a menés à déterminer un analogue de l'algorithme de RSK pour des remplissages d'horizon de semi-finale-standard. Cet analogue est employé pour montrer que les fonctions dissymétriques de Schur égalent les bases standard de Lascoux et de Schützenberger.

1. Introduction

A *symmetric function of degree* n over a commutative ring R (with identity) is a formal power series $f(x) = \sum_\alpha c_\alpha x^\alpha$, where α ranges over all weak compositions of n (of infinite length), $c_\alpha \in R$, x^α stands for the the monomial $x_1^{\alpha_1} x_2^{\alpha_2}\ldots$, and $f(x_{\omega(1)}, x_{\omega(2)}, \ldots) = f(x_1, x_2, \ldots)$ for every permutation ω of the positive integers, $\mathbb{P}$. Many different bases for the vector space of symmetric functions are known. One important basis is the Schur functions.

The *Schur function* $s_\lambda = s_\lambda(x)$ of shape λ in variables $x = (x_1, x_2, \ldots)$ is the formal power series $s_\lambda = \sum_T x^T$, summed over all semi-standard Young tableaux of shape λ. A *semi-standard Young tableau* is formed by first placing the parts of λ into rows of squares, where the i^{th} row has λ_i squares, called *cells.* This diagram, called the Young (or Ferrers) diagram, is drawn in the first quadrant, French style, as in [**3**]. Then each of these cells is assigned a positive integer in such a way that the row entries are weakly increasing and the column entries are strictly increasing. Thus, the values assigned to the cells of λ collectively form the multiset $\{1^{a_1}, 2^{a_2}, \ldots, n^{a_n}\}$, for some n, where a_i is the number of times i appears in T. Here, $x^T = \prod_{i=1}^n x_i^{a_i}$. See [**10**] for a more detailed discussion of symmetric functions and the Schur functions in particular.

The Macdonald polynomials $\tilde{H}_\mu(x; q, t)$ are a special class of symmetric functions which contain a vast array of information. Macdonald [**8**] introduced them and conjectured that their expansion in terms of Schur

2000 *Mathematics Subject Classification.* Primary 05E05; Secondary 05E10.
Key words and phrases. algebraic combinatorics, symmetric functions, Macdonald polynomials, compositions.

polynomials should have positive coefficients. A combinatorial formula for the Macdonald polynomials was recently proved by Haglund, Haiman, and Loehr [**5**].

Building on this work, Haglund, Haiman, and Loehr [**4**] derive a combinatorial formula for nonsymmetric Macdonald polynomials, which gives a new decomposition of the Macdonald polynomial into nonsymmetric components. The statistics involved in this formula can be used to define nonsymmetric Schur polynomials, NS_λ. Letting $q = t = 0$ in the identity implies $s_\mu(x) = \sum_\lambda NS_\lambda(x)$, where the sum is over all rearrangements λ of the partition μ. (A composition μ of n is called a *rearrangement* of a partition λ if it consists of n parts such that when the parts are arranged in descending order, the i^{th} part equals λ_i, for all i.) We give a bijective proof of this decomposition.

Theorem 1.1. $\sum_{\lambda'} NS_{\lambda'}(x_1, ..., x_n) = s_\lambda(x_1, ..., x_n)$, *where the sum is over all rearrangements* λ' *of* λ.

We exhibit a weight-preserving bijection between semi-standard Young tableaux and semi-standard skyline fillings to prove Theorem 1.1. The bijection involves an insertion procedure similar to Schensted insertion. This procedure is the fundamental operation in an analogue of the Robinson-Schensted-Knuth algorithm.

Theorem 1.2. *There exists a bijection between* $\mathbb{N}-$ *matrices with finite support and pairs* (F, G) *of semi-standard skyline fillings of compositions which rearrange the same partition.*

The Schur functions are alternatively defined as the irreducible characters of the linear group on $\mathbb{C}$. Demazure's "Formule des caractère" [**1**] [**2**] provides an interpolation between a dominant weight corresponding to a partition I and the Schur function of index I. For each permutation μ, he obtains a"partial" character whose interpretation involves the Schubert variety of index μ, best understood through the study of the "standard bases" of these spaces. Considering Young tableaux as words in the free algebra, Lascoux and Schützenberger [**6**] describe the standard bases $\mathfrak{U}(\mu, I)$ using symmetrizing operators on the free algebra which lift the operators used by Demazure. This description provides a recursive algorithm to determine the basis $\mathfrak{U}(\mu, I)$ given the basis $\mathfrak{U}(\lambda, I)$, where $\mu = \sigma_i\lambda$ for some i. The nonsymmetric Schur functions provide a non-recursive combinatorial description of $\mathfrak{U}(\mu, I)$ for arbitrary μ, I.

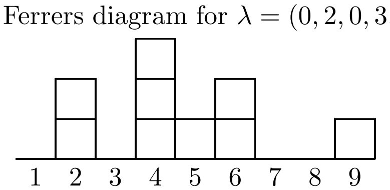
Theorem 1.3. $\mathfrak{U}(\mu, I) = NS_{\mu(I)}$, *where* $\mu(I)$ *denotes the action of* μ *on the parts of* I.

This theorem provides a mapping between the combinatorics of symmetrizing (or string) operators and nonsymmetric Schur functions.

2. Combinatorial description of the nonsymmetric Schur functions

2.1. Semi-standard skyline fillings. Let $\gamma = (\gamma_1, \gamma_2, ..., \gamma_n)$ be a composition of n into n parts, allowing zero as a part. (We will consider compositions of n into arbitrarily many parts in section 2.3.) The *composition Ferrers diagram of* γ is a figure consisting of n *cells* arranged in n columns. The i^{th} column contains γ_i cells, and the number of cells in a column is called the *height* of that column. This is an analogue of the Ferrers diagram of a partition λ, which consists of rows of cells such that the i^{th} row contains λ_i cells.

Example 2.1. The composition Ferrers diagram for $\lambda = (0, 2, 0, 3, 1, 2, 0, 0, 1)$



A *filling,* σ, of a composition Ferrers diagram, λ, is a function $\sigma : \lambda \to \mathbb{Z}_+$, which we picture as an assignment of positive integer entries to the cells of λ. We consider the 0^{th} row to consist of cells numbered from 1 to n in strictly increasing order. Let $\sigma(i)$ denote the entry in the i^{th} square of the composition Young diagram encountered if we read across rows from left to right, beginning at the highest row and working downwards.

To define the nonsymmetric Schur functions, we need the statistics $Des(\sigma)$ and $Inv(\sigma)$. As in [**3**], a *descent* of σ is a pair of entries $\sigma(a) > \sigma(b)$, where the cell a is directly above b. In other words, $b = (i, j)$ and $a = (i + 1, j)$, where the i^{th} coordinate denotes the height of cell b and the j^{th} coordinate denotes the column containing j. Define $Des(\sigma) = \{a \in \lambda : \sigma(a) > \sigma(b) \text{ is a descent}\}$.

Three cells $a, b, c \in \lambda$ form a *triple of type A* if they are situated as follows,

$$\begin{array}{ccc} \boxed{a} & \cdots & \boxed{c} \\ \boxed{b} & & \end{array}$$

where a and c are in the same row, possibly the first row, possibly with cells between them, and the column containing a and b has height greater than or equal to the height of the column containing c.

Define for $x, y \in \mathbb{Z}_+$

$$I(x,y) = \begin{cases} 1 & \text{if } x > y \\ 0 & \text{if } x \leq y \end{cases}$$

Let σ be a composition filling and let α, β, δ be the entries of σ in the cells of a type A triple (a, b, c):

$$\begin{array}{ccc} \boxed{\alpha} & \cdots & \boxed{\delta} \\ \boxed{\beta} & & \end{array}$$

Then the triple (a, b, c) is called an *inversion triple of type A* if and only if $I(\alpha,\delta) + I(\delta,\beta) - I(\alpha,\delta) = 1$.

The *reading order* of a filling is an ordering of its cells beginning with the top row and listing the cells from left to right, travelling down, row by row, to the bottom row. Define a filling σ to be *standard* if it is a bijection $\sigma : \mu \cong \{1,...,n\}$. The *standardization* of a composition filling is the unique standard filling ξ such that $\sigma \circ \xi^{-1}$ is weakly increasing, and for each α in the image of σ, the restriction of ξ to $\sigma^{-1}(\{\alpha\})$ is increasing with respect to the reading order. Therefore the triple (a, b, c) is an inversion triple of type A if and only if after standardization, the ordering from smallest to largest of the entries in cells a, b, c induces a counter-clockwise orientation.

Similarly, three cells $a, b, c \in \lambda$ form a *triple of type B* if they are situated as shown below,

$$\begin{array}{ccc} & & \boxed{b} \\ \boxed{a} & \cdots & \boxed{c} \end{array}$$

Here a and c are in the same row (possibly row 0) and the column containing b and c has greater height than the column containing a.

Let σ be a composition filling and let α, β, δ be the entries of σ in the cells of a type B triple (a, b, c).

$$\begin{array}{ccc} & & \boxed{\beta} \\ \boxed{\alpha} & \cdots & \boxed{\delta} \end{array}$$

Then the triple (a, b, c) is called an *inversion triple of type B* if and only if $I(\beta,\alpha) + I(\alpha,\delta) - I(\beta,\delta) = 1$. In other words, the triple (a, b, c) is an inversion triple of type B if and only if after standardization, the ordering from smallest to largest of the entries in cells a, b, c induces a clockwise orientation.

Denote by *semi-standard skyline filling* any composition filling F such that $Des(F) = \emptyset$ and every triple is an inversion triple. These conditions arise by setting $q = t = 0$ in the combinatorial formula for the nonsymmetric Macdonald polynomials [**4**].

Definition 2.2. Let λ be a composition of n into n parts, where some of the parts could be equal to zero. The nonsymmetric Schur function $NS_\lambda = NS_\lambda(x)$ in the variables $x = (x_1, x_2, ..., x_n)$ is the formal power series $NS_\lambda(x) = \sum_F x^F$ summed over all semi-standard skyline fillings F of composition λ. Here, $x^F = \prod_{i=1}^n x_i^{\sigma_i}$ is the *weight* of σ. (See Figure 2.1.)

The combinatorial formula for nonsymmetric Macdonald polynomials [**4**] contains an additional "non-attacking" condition. This condition states that for each pair of cells a and b with a to the left of b in the row directly below b, $\sigma(a) \neq \sigma(b)$. (If $\sigma(a) = \sigma(b)$, a and b are called *attacking* cells.)

Lemma 2.3. *The descent and inversion conditions used to describe the semi-standard skyline fillings guarantee that no two cells of a semi-standard skyline filling are attacking.*

Proof. Assume there exist two attacking cells a and b with $\sigma(a) = \sigma(b) = \alpha$ to get a contradiction. If the column containing a is taller than or equal to the column containing b, then a lies directly below a cell c which must have $\sigma(c) \leq \alpha$. When the values in these three cells are standardized, c, b, a form a non-inversion triple of type A. If the column containing b is taller than the column containing a, b is directly on top of a cell c which must have $\sigma(c) \geq \alpha$. The cells a, b, c form a type B non-inversion triple. $\square$

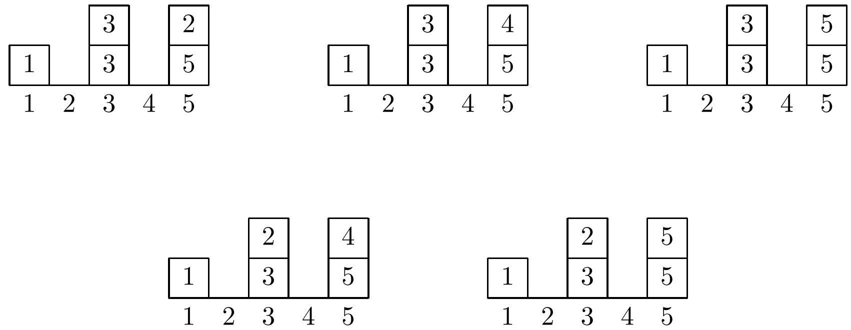
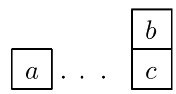


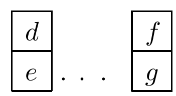
Figure 2.1. $NS_{(1,0,2,0,2)} = x_1x_2x_3^2x_5 + x_1x_3^2x_4x_5 + x_1x_3^2x_5^2 + x_1x_2x_3x_4x_5 + x_1x_2x_3x_5^2$.

Lemma 2.4. *If a, b, c is a type B triple with a and c on the same row and b directly above c, then $\sigma(a) < \sigma(c)$.*

Proof. Let a, b, c be a type B triple situated as pictured below.



To get a contradiction, first assume $\sigma(a) > \sigma(c)$. In the basement row, the column containing a has a value less than the value of the column containing c. So at some intermediate row we have



with $\sigma(d) > \sigma(f)$ and $\sigma(e) \leq \sigma(g)$. We must have $\sigma(d) \leq \sigma(e)$. Therefore, $\sigma(f) < \sigma(d) \leq \sigma(e) \leq \sigma(g)$. But then $\sigma(f) < \sigma(e) \leq \sigma(g)$ and this type B triple f, e, g is not an inversion triple.

Next assume $\sigma(a) = \sigma(c)$. If so, by standardization we may assume that $\sigma(a) < \sigma(c)$. To have an inversion triple, $\sigma(b)$ must be between $\sigma(a)$ and $\sigma(c)$. But then $\sigma(b)$ must equal $\sigma(a)$ and $\sigma(c)$, which implies that a and b are attacking. So $\sigma(a)$ cannot equal $\sigma(c)$.

$\square$

Lemmas 2.3 and 2.4 provide us with several conditions on the cells in our diagram. They will be useful in proving facts about the insertion process.

2.2. A basis for homogeneous polynomials of degree n in n variables. Several other bases for symmetric functions have nonsymmetric analogues. For instance, the nonsmmyetric monomial corresponding to a given composition γ of n into n parts is given by $NM_\gamma = x_1^{\gamma_1}x_2^{\gamma_2}...x_n^{\gamma_n}$. It is clear that the sum over all rearrangements of a given partition μ of the nonsymmetric monomials is equal to the monomial symmetric function m_μ. Every polynomial of degree n in n variables can be written as a sum of nonsymmetric monomials, so the nonsymmetric monomials form a basis for the algebra of homogeneous polynomials of degree n in n variables.

Definition 2.5. The **reverse dominance order** on compositions is defined as follows:
$\mu \leq \gamma \iff \sum_{i=k}^n \mu_i \leq \sum_{i=k}^n \gamma_i$ for $1 \leq i \leq n$.

A semi-standard skyline filling is said to have type μ if it contains μ_i copies of the number i for each i. If γ and μ are compositions of n into n parts, let $NK_{\gamma,\mu}$ denote the number of semi-standard skyline fillings of shape γ and type μ. $NK_{\gamma,\mu}$ is called a *nonsymmetric Kostka number*. The ordinary Kostka numbers are obtained as a sum of nonsymmetric Kostka numbers: $K_{\lambda,\mu} = \sum NK_{\gamma,\mu}$, where the sum is over all rearrangements γ of λ.

Theorem 2.6. *Suppose that γ and μ are both compositions of n into n parts and $NK_{\gamma,\mu} \neq 0$. Then $\gamma \geq \mu$ in the dominance order. Moreover, $NK_{\gamma,\gamma} = 1$.*

Proof. Assume that $NK_{\gamma,\mu} \neq 0$. By definition, there exists a semi-standard skyline filling of shape γ and type μ. Assume that an entry k appears in one of the first $k-1$ columns. Then this column would contain a descent, since there is an entry less than k in the column at a lower position than k, namely the basement entry. Therefore, the parts $k, k+1, ..., n$ all appear in the last $n-k+1$ columns. So $\mu_k + \mu_{k+1} + ... + \mu_n \leq \gamma_k + \gamma_{k+1} + ... + \gamma_n$ for each k, as desired. Moreover, if $\mu = \gamma$, then the i^{th} column must contain only entries with value i, so $NK_{\gamma,\gamma} = 1$. □

Corollary 2.7. *The nonsymmetric Schur functions form a basis for the algebra of homogeneous polynomials of degree n in n variables.*

Proof. Theorem 2 is equivalent to the assertion that the transition matrix from the nonsymmetric Schur functions to the nonsymmetric monomials (with respect to the reverse dominance order) is upper triangular with 1's on the main diagonal. Since this matrix is invertible, the nonsymmetric Schur functions of degree n are a basis for homogeneous polynomials of degree n in n variables. □

2.3. Nonsymmetric Schur functions in infinitely many variables. We may relax the restriction on the number of parts to obtain nonsymmetric Schur functions in infinitely many variables.

Definition 2.8. A **weak composition** of n is an infinite sequence of non-negative integers which sum to n.

Let γ be a weak composition of n. Its composition Ferrers diagram consists of infinitely many columns such that the i^{th} column contains γ_i cells. As above, fill this diagram with positive integers in such a way that there are no descents and every triple is an inversion triple to get a semi-standard skyline filling. Then $NS_\gamma(x) = \sum_F x^F$, where F ranges over all semi-standard skyline fillings of the composition Ferrers diagram of γ.

We may also define the nonsymmetric monomials in infinitely many variables. The nonsmmyetric monomial corresponding to a weak composition γ of n is given by $NM_\gamma = \prod_i x_i^{\gamma_i}$. It is clear that the sum over all rearrangements of a given partition μ of the nonsymmetric monomials is equal to the monomial symmetric function m_μ. Every polynomial can be written as a sum of nonsymmetric monomials, so the nonsymmetric monomials form a basis for all polynomials.

Definition 2.9. Let μ and γ be weak compositions of n. The **reverse dominance order** on weak compositions is defined as follows.

$$\mu \leq \gamma \iff \sum_{i=k}^{\infty} \mu_i \leq \sum_{i=k}^{\infty} \gamma_i \qquad \forall k, \quad k \geq 1$$

.

If γ and μ are weak compositions of n, $NK_{\gamma,\mu}$ denote the number of semi-standard skyline fillings of shape γ and type μ as above. Again, the ordinary Kostka numbers are obtained as a sum of nonsymmetric Kostka numbers. $K_{\lambda,\mu} = \sum NK_{\gamma,\mu}$, where the sum is over all rearrangements γ of λ.

Theorem 2.10. *Suppose that γ and μ are both weak compositions of n and $NK_{\gamma,\mu} \neq 0$. Then $\mu \leq \gamma$ in the dominance order. Moreover, $NK_{\gamma,\gamma} = 1$.*

Proof. Assume that $NK_{\gamma,\mu} \neq 0$. By definition, there exists a semi-standard skyline filling of shape γ and type μ. Assume that an entry k appears in one of the first $k-1$ columns. Then this column would contain a descent, since there is an entry less than k in the column at a lower position than k, namely the basement entry. Therefore, the entries greater than or equal to k all appear after the $(k-1)^{th}$ column. So $\sum_{i=k}^{\infty} \mu_i \leq \sum_{i=k}^{\infty} \gamma_i$ for each k, as desired. Moreover, if $\mu = \gamma$, then the i^{th} column must contain only entries with value i, so $NK_{\gamma,\gamma} = 1$. □

Corollary 2.11. *The nonsymmetric Schur functions form a basis for all polynomials.*

Proof. Theorem 2.10 is equivalent to the assertion that the transition matrix from the nonsymmetric Schur functions to the nonsymmetric monomials (with respect to the reverse dominance order) is upper triangular with 1's on the main diagonal. Since this matrix is invertible, the nonsymmetric Schur functions are a basis for all polynomials. □

3. Proof of Theorem 1.1

We will in fact prove a slightly more general statement

$$\sum_{\lambda'} NS_{\lambda'} = s_\lambda, \tag{3.1}$$

where the sum is over all weak compositions λ' which rearrange λ.

3.1. An analogue of Schensted insertion. Let F be a semi-standard skyline filling of a weak composition γ of n. Then $F = (F_j)$, where F_j is the j^{th} cell when the cells are in reading order, including the cells in the basement. We define the operation $F \leftarrow k$.

Let r be the smallest integer such that $\sigma(F_r) \geq k$ and there is no cell c with $\sigma(c) = k$ in the row directly above the row containing F_r. If there is no cell directly on top of F_r, then place k on top of F_r are the resulting figure is $F \leftarrow k$. Otherwise let a be the cell directly on top of F_r. If $\sigma(a) < k$ then k "bumps" $\sigma(a)$. In other words, k replaces $\sigma(a)$ and we now find the least r' such that $r' > r$ and $\sigma(F_{r'}) \geq a$ and repeat. If $\sigma(a) > k$ then continue to the next r' such that $r' > r$ and $\sigma(F_{r'}) \geq k$ and repeat. This procedure terminates, since there are infinitely many basement entries greater than k.

Lemma 3.1. *When restricted to n-compositions, this procedure terminates.*

Proof. Assume that the procedure does not terminate to get a contradiction. This could only occur if some letter α reaches the last cell in the basement without finding an r such that $\sigma(F_r) \geq \alpha$ and such that the cell b on top of F_r has $\sigma(b) \leq \alpha$. The value α is an entry in the basement, say $\sigma(F_j)$. The letter α which is unplaced could not have been bumped from a cell to the right of F_j in the row above F_j, for otherwise the cell containing α and F_j would be attacking. Since α was not inserted on top of F_j, the entry b on top of F_j must have $\sigma(b) \geq \alpha$. But since F has no descents, $\sigma(b) = \alpha$. So the leftover α must have come from a higher row. Continuing this line of reasoning, we see a column containing the value α at each row until a certain height h at which this column contains an entry strictly smaller than α. If α was bumped from row h, α must have been bumped from a cell to the right of the α^{th} column. However, then α and the α in row $h-1$ of column α would be entries in attacking cells in F. By Lemma 2.3, there are no attacking cells in F. Therefore we have a contradiction. ☐

The resulting diagram is $F \leftarrow k$.

Proposition 3.1. *If F is a semi-standard skyline filling, then $F \leftarrow k$ is a semi-standard skyline filling.*

Proof. It is clear by construction that $F \leftarrow k$ has no descents. We must prove that every triple is an inversion triple. We argue by contradiction. To get a contradiction, assume $F \leftarrow k$ contains a type A non-inversion triple, a, b, c situated as shown.

$$\begin{array}{ccc} \boxed{a} & & \boxed{b} \\ \boxed{c} & & \end{array}$$

Then we must have $\sigma(a) \leq \sigma(b) \leq \sigma(c)$. In F, we must have had different (possibly empty) entries in these cells. Because the insertion path moves along the reading word and its entries are decreasing, at most one of $\sigma(a)$, $\sigma(b)$, and $\sigma(c)$ is different from its value in F. Examine each cell individually to get a contradiction. For example, assume the cell a in $F \leftarrow k$ contained a different value, $\beta \neq \sigma(a)$, in F. Since $\beta, \sigma(b), \sigma(c)$ was an inversion triple in F, $\sigma(b) < \beta \leq \sigma(c)$. But since $\sigma(b)$ bumped β, $\sigma(a) > \beta$, so $\sigma(a) > \sigma(b)$ contradicts $\sigma(a) \leq \sigma(b)$.

Next assume that $F \leftarrow k$ contains a type B non-inversion triple, a, b, c situated as depicted.

$$\begin{array}{ccc} & & \boxed{b} \\ \boxed{a} & & \boxed{c} \end{array}$$

We must have $\sigma(b) < \sigma(a) \leq \sigma(c)$. Again, only one of the entries is different from its value in F. Examine each cell individually to derive a contradiction.

□

3.2. The bijection Ψ between SSYT and SSSF. Let T be a semi-standard young tableau. We may associate to T the word $col(T)$ which consists of the entries of each column of T, read top to bottom from columns left to right.

EXAMPLE 3.2. For T as below, $col(T) = 10\ 9\ 8\ 4\ 2\ 1 \cdot 11\ 10\ 7\ 5\ 2 \cdot 10\ 8\ 5\ 3 \cdot 5 \cdot 10$.

10				
9	11			
8	10	10		
4	7	8		
2	5	5		
1	2	3	5	10

Begin with an arbitrary SSYT T and the empty SSSF ϕ with the basement row containing all letters of $\mathbb{Z}_+$. Let k be the rightmost letter in $col(T)$. Insert k into ϕ to get the SSSF $F = \phi \leftarrow k$. Then let k' be the next letter in $col(T)$ reading right to left. Obtain the SSSF $F \leftarrow k'$. Continue in this manner until you have inserted all the letters of $col(T)$. The resulting diagram is the SSSF $\Psi(T)$.

LEMMA 3.3. *Let C_i be a column of $col(T)$. The placement of each letter of C_i terminates at a different column, with the smallest letter of C_i terminating at the top of the highest column, the second smallest letter terminating at the top of the second highest column, and so forth. (If there are two columns of the same height, the one farther left is the termination point of the smaller letter.)*

PROOF. The first letter of C_i is smaller than or equal to all letters which came before it, so it is placed onto the top of the tallest column. To argue inductively, assume that Lemma 3.3 is true after the first j letters of C_i have been placed. The next letter α is greater than each of the other letters, therefore its inertion path lies below that of the other letters, so the first place it can terminate is on top of the tallest column which has not yet been a termination point for a letter of C_i. The highest entry in this column is greater than or equal to the letter β which has been most recently bumped, so β is placed on top of this column and the proof is complete. □

PROPOSITION 3.2. *The shape of $\Psi(T)$ is a rearrangement of the shape of T.*

PROOF. Argue by induction on the number of columns of T. First assume that T contains only one column. The shape of T is 1^n. Then $col(T)$ is a strictly decreasing word. Therefore each letter maps to the bottom row of the semi-standard skyline filling. The resulting shape is an arrangement of zeros and ones, a rearrangement of 1^n.

Next, assume that if T contains j columns then the shape of $\Psi(T)$ is a rearrangement of the shape of T. Let T be an SSYT of shape λ which contains $j+1$ columns. After mapping the first j columns of T, the shape of the resulting figure is a rearrangement of $(\lambda_1 - 1, \lambda_2 - 1, ...)$. By Lemma 3.3, mapping the next column into the shape adds one cell to each existing column, plus possibly several new cells if the $(j+1)^{th}$ column is taller than the j^{th} column. Therefore the resulting shape is a rearrangement of $(\lambda_1, \lambda_2, ...) = \lambda$. □

PROPOSITION 3.3. *The map Ψ is invertible.*

PROOF. Consider the set S of cells which are in the top row of some column. Of these, begin with the cell c which is farthest right in the reading order. This was the last cell to be bumped into place. Scan backwards through the reading order to find the next cell, d, such that $\sigma(d) > \sigma(c)$ and d lies directly below an entry less than or equal to c. This entry $\sigma(d)$ bumped $\sigma(c)$. Repeat with $\sigma(d)$. Continue this scanning procedure until there are no cells farther back in the reading order which could have bumped the selected entry, e. This entry is the first letter in $col(T)$.

Choose the next element of S to appear in the backwards reading order. (If there are no other cells in S, create a new set S' consisting of all the cells which are in the top row of some column.) Move backwards from this element through the reading word to determine the initial element whose placement terminated with this particular element. Continue this procedure until the entire word $col(T)$ has been determined. This procedure inverts the map Ψ. □

The map Ψ : SSYT $\rightarrow$ SSSF is a weight-preserving invertible map between semi-standard Young tableaux and semi-standard skyline fillings. In particular, this means that the number of SSYT of shape λ with weight $\prod x_i^{a_i}$ is equal to the number of SSSF with weight $\prod x_i^{a_i}$ whose shape rearranges λ. Thus the coefficient of $\prod x_i^{a_i}$ in $\sum_{\lambda'} NS_{\lambda'}$ is equal to the coefficient of $\prod x_i^{a_i}$ in s_λ. This completes the proof of Equation 3.1.

4. An analogue of the Robinson-Schensted-Knuth Algorithm

The insertion process utilized in the above bijection is reminiscent of Schensted insertion, the fundamental operation of the Robinson-Schensted-Knuth Algorithm.

Theorem 4.1. *(Robinson-Schensted-Knuth* [**9**]*) There exists a bijection between $\mathbb{N}$-matrices of finite support and pairs of semi-standard Young tableaux of the same shape.*

We apply the same procedure to arrive at an analogue of the RSK Algorithm for semi-standard skyline fillings. Recall that Theorem 1.2 states that there exists a bijection between $\mathbb{N}$-matrices of finite support and pairs of semi-standard skyline fillings whose shapes are rearrangements of the same partition.

4.1. ρ **:** $\mathbb{N}$**-matrices** $\longrightarrow$ **SSSF** $\times$ **SSSF.** Let $A = (a_{i,j})$ be an $\mathbb{N}$-matrix with finite support. There exists a unique two-line array corresponding to A which is defined by the non-zero entries in A. Beginning at the upper lefthand corner and reading left to right, top to bottom, find the first non-zero entry $a_{i,j}$. Place an i in the top line and a j in the bottom line $a_{i,j}$ times. When this has been done for each non-zero entry, we have the resulting array

$$w_A = \begin{pmatrix} i_1 & i_2 & \dots \\ j_1 & j_2 & \dots \end{pmatrix}$$

Begin with an empty semi-standard skyline filling F. Read the bottom row from right to left, inserting the entries into F according to the map Ψ described above as they are read. Each time an entry from the bottom line is placed, send the entry directly above it into an SSSF G which records the place where a cell is added. If the cell j_k is added to the bottom row of F, the corresponding entry i_k is placed on the bottom row in the i_k^{th} column of G. If there is ambiguity about which column of G an entry is placed on, it is always placed on the leftmost possible column of the same height as the column in F on which its counterpart was placed. In this way the shape of G becomes a rearrangement of the shape of F. When the process is complete, the result is a pair (F, G) of SSSF whose shapes are rearrangements of the same partition.

4.2. The inverse of the map ρ. Given (F, G), a pair of semi-standard skyline fillings whose shapes are rearrangements of the same partition μ, let G_{rs} be the highest occurrence of the smallest entry of G. Here G_{rs} is the element of G in row r and column s. Since equal elements of G are inserted bottom to top, it follows that $G_{rs} = i_1$ and $F_{rs'}$ was the last element of F to be bumped into place after inserting j_1. (If s is the i^{th} column of height r in G, then s' is the i^{th} column of height r in F.)

Delete $F_{rs'}$ from F and G_{rs} from G. Scan right to left, bottom to top (backwards through the reading word) starting with the cell directly to the left of $F_{rs'}$ to determine which (if any) cell bumped $F_{rs'}$. If there exists a cell k before $F_{rs'}$ in the reading word such that $\sigma(k) > \sigma(F_{rs'})$ and the cell directly on top of k has value less than or equal to $\sigma(F_{rs'})$, this k bumped $F_{rs'}$. Replace $\sigma(k)$ by $\sigma(F_{rs'})$ and repeat the procedure with $\sigma(k)$. Continue working backward through the reading word until there are no more letters. The remaining entry is the letter i_1.

Next find the highest occurrence of the smallest entry j_2 of G. Repeat the procedure to find i_2. Continue until there are no more entries in F and G. Then all of the i and j values of the array w_A have been determined, and the process is inverted.

5. The standard bases of Lascoux and Schützenberger

The Schubert polynomials were introduced by Lascoux and Schützenberger [**7**] as a combinatorial tool for attacking problems in algebraic geometry. The Schubert polynomials can be described as a sum of *standard bases*, $\mathfrak{U}(\mu, I)$, where μ is a permutation and I is a partition. Lascoux and Schützenberger [**6**] define an action of the symmetric group on the free algebra and this action is used to build the standard bases inductively.

5.1. Constructing the standard bases. Each permutation in the symmetric group can be decomposed into a series of elementary transpositions, so it is enough to define the action for a simple transposition, σ_i, which permutes i and $i+1$. The operator $\overline{\pi}_i = \overline{\pi}_{\sigma_i}$ is

$$f \longrightarrow (f^{\sigma_i} - f)/(1 - x_i/x_{i+1}) = f\overline{\pi}_i,$$

where f^{σ_i} denotes the transposition action of σ_i on the indices of the variables in f.

The operators $\overline{\pi}_i$ satisfy the Coxeter relations $\overline{\pi}_i\overline{\pi}_{i+1}\overline{\pi}_i = \overline{\pi}_{i+1}\overline{\pi}_i\overline{\pi}_{i+1}$ and $\overline{\pi}_i\overline{\pi}_j = \overline{\pi}_j\overline{\pi}_i$ for $\|j-i\| > 1$. We can lift the operator $\overline{\pi}_i$ into an operator θ_i on the free algebra by the following process. Given i and a word w in the alphabet X, let m be the number of times the letter x_{i+1} occurs in w and let $m+k$ be the number of times the letter x_i occurs in w. Then if $k \geq 0$, w and w^{σ_i} differ by the exchange of a subword x_i^k into x_{i+1}^k. If $k < 0$, then w and w^{σ_i} differ by the exchange of x_{i+1}^{-k} into x_i^{-k}.

When $k \geq 0$, define $w\theta_i$ to be the sum of all words in which the subword x_i^k has been changed respectively into $x_i^{k-1}x_{i+1}, x_i^{k-2}x_{i+1}^2, ..., x_{i+1}^k$. When $k < 0$, define $w\theta_i$ to be $-(w^{\sigma_i})\theta_i$. (This second case will not be needed in this paper.)

Every partition $I = (I_1, I_2, ...)$ has a corresponding *dominant monomial* $x^I = (x_{I_1}...x_2x_1)(x_{I_2}...x_2x_1)...$, which equals the weight of the *super tableau*, which is the SSYT with is in the i^{th} row. We take the following theorem to be the definition of the *standard basis* $\mathfrak{U}(\mu, I)$ associated to the pair μ, I (where μ is a permutation and I is a partition).

Theorem 5.1. *(Lascoux-Schützenberger [**6**]) Let x^I be a dominant monomial and $\sigma_i\sigma_j...\sigma_k$ be any reduced decomposition of a permutation μ. Then $\mathfrak{U}(\mu, I) = x^I\theta_i\theta_j...\theta_k$.*

Theorem 5.1 provides an inductive method for constructing the standard basis $\mathfrak{U}(\mu, I)$. Begin with $\mathfrak{U}(id, I)$ and apply θ_i to determine $\mathfrak{U}(\sigma_i, I)$. Then apply θ_j to $\mathfrak{U}(\sigma_i, I)$ to get $\mathfrak{U}(\sigma_i\sigma_j, I)$. Continue this process until the desired standard basis is obtained. Figure 5.1 depicts all the standard bases for the partition $(2,1)$.

5.2. A non-inductive construction of the standard bases. The standard bases with partition λ can be considered as a decomposition of the Schur function s_λ. For any partition λ of n, we have

$$\sum_{\sigma \in S_n} \mathfrak{U}(\sigma, \lambda) = s_\lambda.$$

Since the nonsymmetric Schur functions are also a decomposition of the Schur functions, it is natural to determine their relationship to the standard bases. Theorem 1.3 states that $\mathfrak{U}(\mu, I) = NS_{\mu(I)}$, where $\mu(I)$ denotes the action of μ on the parts of I. To prove Theorem 1.3, we need a few lemmas.

Lemma 5.2. *Let col(T) be the column reading word of T. Label the occurrences of the entry α in col(T) in increasing order starting from the right. Then $i < j \Rightarrow$ the i^{th} occurrence of α (denoted α_i) in $\Psi(T)$ is in a lower row than the j^{th} occurrence of α, denoted α_j.*

Proof. Consider the step during which α_j is being placed. At this step, the α_i is already placed in some row of the partial semi-standard skyline filling. If α_j reaches a cell a with $\sigma(a) = \alpha$ without being placed, the cell b on top of α must have $\sigma(b) < \alpha$. Therefore, α_j will bump $\sigma(b)$ and be placed on top of a. Therefore, α_j will always remain in a higher row than α_i. □

Lemma 5.3. *Given an arbitrary semi-standard skyline filling F with row entries $R_1, R_2, ...R_k$, where $k = max_i\{\gamma_i\}$, F is the only SSSF with these row entries.*

Proof. Given the row entries $R_1, R_2, ..., R_k$, map them into a semi-standard skyline filling as follows. Let α_1 be the largest entry in R_1. Place α_1 as far left as possible in the first row of an empty SSSF. Next place the second largest entry of R_1 as far left as possible in the first row of the SSSF. Continue placing the elements of R_1 in this manner. Next, choose the largest entry of R_2. Place it as far left as possible in the second row of the partially constructed SSSF. Continue this procedure until the smallest entry of R_2 has been placed. Do this for each of the k rows. Once R_k has been placed, the resulting figure is indeed a semi-standard skyline filling, and the only SSSF with row entries $R_1, R_2, ..., R_k$. □

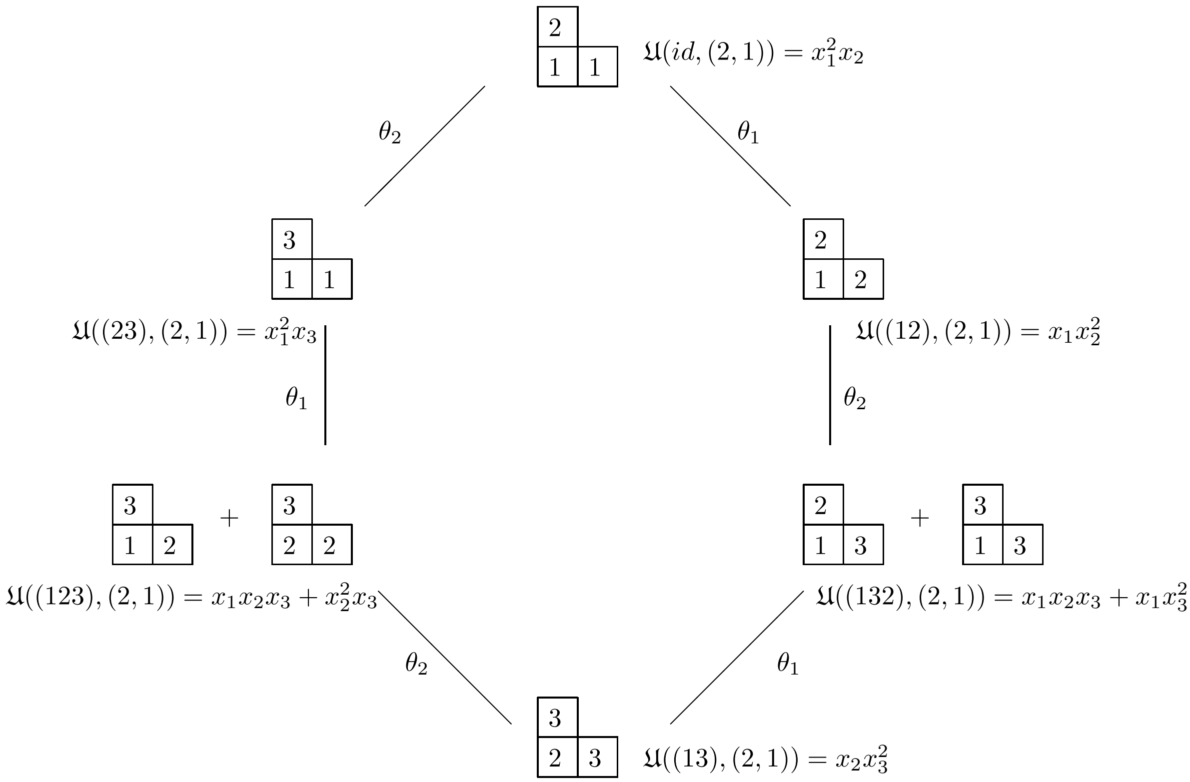


FIGURE 5.1. The standard bases for the partition $(2,1)$ and all permutations in S_3.

Let $\tilde{\theta}_i$ be the action of θ_i on an individual semi-standard Young tableau. This action is described by a matching procedure. Let $col(T)$ be the column word of T and let $(i+1)_1$ be the leftmost occurrence of $i+1$ in $col(T)$. Match $(i+1)_1$ with the leftmost occurrence of i which lies to the right of $(i+1)_1$ in $col(T)$. If there is no such i, the matching procedure is complete. Otherwise, continue with the next $i+1$ until there are no more occurrences of $i+1$.

When the matching procedure is complete, send the rightmost occurrence of i to $i+1$. The resulting word is $\tilde{\theta}_i(T) = T'$. If $T \in \mathfrak{U}(\mu, I)$ then either $\tilde{\theta}_i(T) \in \mathfrak{U}(\mu, I)$ or $\tilde{\theta}_i(T) \in \mathfrak{U}(\sigma_i\mu, I)$.

LEMMA 5.4. *There exists a map Θ_i :SSSF$\longrightarrow$SSSF such that for $F \in NS_\mu$, either $\Theta_i(F) \in NS_\mu$ or $\Theta_i(F) \in NS_{\sigma_i\mu}$ and the following diagram commutes.*

$$\begin{tikzcd} T \arrow[r, "\tilde{\theta}_i"] \arrow[d, "\Psi"] & T' \arrow[d, "\Psi"] \\ F \arrow[r, "\Theta_i"] & F' \end{tikzcd}$$

PROOF. Let F be an arbitrary semi-standard skyline filling and let $leftread(F)$ be the reading word obtained by reading F right to left, top to bottom, keeping track of the rows. Find the first entry a of this word such that $\sigma(a) = i+1$. Match this entry $i+1$ to the first $\sigma(b)$ which lies to the right of a in $leftread(F)$ such that $\sigma(b) = i$. If there is no such b, $\sigma(a)$ is unmatched and the matching process is complete. Continue this matching until an unmatched $i+1$ is reached.

Pick the rightmost unmatched i. Change it to $i+1$. (If there is none, then $\Theta_i(F) = F$.) The result is a collection of rows which differ from $leftread(F)$ in precisely one entry. Lemma 5.3 provides a procedure for mapping this collection of rows to a unique SSSF. This SSSF is $\Theta_i(F) = F'$, and $\Theta_i(\Psi(T)) = \Psi(\tilde{\theta}_i(T))$. So the diagram commutes.

Assume $F \in NS_\mu$. When $\Theta_i(F) = F$, $\Theta_i(F) \in NS_\mu$. We must show that in the case where an unmatched i is mapped to $i+1$, the resulting semi-standard skyline filling is either in NS_μ or in $NS_{\sigma_i\mu}$. But the map Θ_i shifts the highest unmatched i in F. If there are no occurrences of the letter i in the row directly below the shifted i, this $i+1$ is mapped to the same position, as are all the cells above it. So the shape of the diagram remains the same. Otherwise, there is a column consisting only of is and a column consisting only of $(i+1)$s below the first unmatched i. Sending this i to $i+1$ moves it into the $(i+1)^{th}$ column and therefore permutes the i^{th} and $(i+1)^{th}$ column, resulting in the shape $\sigma_i\mu$. So our proof is complete.

□

5.3. Proof of Theorem 1.3. We fix a partition I and argue by induction on the length of the permutation μ in $\mathfrak{U}(\mu, I)$. Let μ be the identity. Then $\mathfrak{U}(\mu, I)$ is the dominant monomial. Consider I as a composition of n into n parts by adding zeros to the right if necessary. Each entry a in I_1 must have $\sigma(a) = 1$, for otherwise there would be a descent. If the second column contained an entry b such that $\sigma(b) = 1$, this cell and the cell in the row directly below of the first column would be attacking. Continuing in this manner, we see that the i^{th} column must have $\sigma(c) = i$ for each cell c. Therefore, the $NS_I = \mathfrak{U}(\mu, I)$.

Next assume that $\mathfrak{U}(\mu, I) = NS_{\mu(I)}$, where $\mu(I)$ is the permutation μ applied to the columns of I when I is considered as a composition of n into n parts. The monomials in $\mathfrak{U}(\sigma_i\mu, I)$ are the monomials of $\mathfrak{U}(\mu, I)$ whose image under (possibly multiple applications of) $\tilde{\theta}_i$ is not a monomial of $\mathfrak{U}(\mu, I)$. Pick some such monomial, represented by the SSYT T. By Lemma 5.4 $\Psi(\tilde{\theta}_i(T)) = \Theta_i(\Psi(T))$. Since $\Psi(T) \in NS_\mu$, $\Theta_i(\Psi(T)) \in NS_\mu$ or $\Theta_i(\Psi(T)) \in NS_{\sigma_i\mu}$. If $\Theta_i(\Psi(T)) \in NS_\mu$, then $\tilde{\theta}_i(T) \in \mathfrak{U}(\mu, I)$ by assumption. But this is a contradiction, so $\Theta_i(\Psi(T)) \in NS_{\sigma_i\mu}$. Therefore, $\mathfrak{U}(\sigma_i\mu, I) \subseteq NS_{\sigma_i\mu}$. If F is a monomial in $NS_{\sigma_i\mu}$, one can determine an element of NS_μ which, after possibly multiple applications of Θ_i maps to F. Therefore $NS_{\sigma_i\mu} \subseteq \mathfrak{U}(\mu, I)$. So $NS_{\sigma_i\mu} = \mathfrak{U}(\mu, I)$.

6. Applications of Theorems 1.2 and 1.3

The analogue of the Robinson-Schensted-Knuth algorithm (Theorem 1.2) can be used to extend results about plane partitions and permutation enumeration. The non-inductive description of standard bases provided in Theorem 1.3 facilitates our understanding of the representation theory of Schubert polynomials and nonsymmetric Schur functions.

References

[1] Demazure, M., Désingularisation des variétés de Schubert, Ann. E. N. S., 6 (1974), 163-172.

[2] Demazure, M., Une formule des caractères, Bull. Sc. M., 98 (1974), 163-172.

[3] Haglund, J., Haiman, M., and Loehr, N., A combinatorial formula for Macdonald polynomials, Jour. Amer. Math. Soc., 18 (2005), 735-761.

[4] Haglund, J., Haiman, M., and Loehr, N., A combinatorial formula for Nonsymmetric Macdonald polynomials, pre-print (2005).

[5] Haglund, J., Haiman, M., and Loehr, N., Combinatorial theory of Macdonald polynomials I: Proof of Haglund's formula, Proc. Nat. Acad. Sci. U.S.A., 102: 8 (2005), 2690-2696.

[6] Lascoux, A., and Schützenberger, M.-P., Keys and Standard Bases, Invariant Theory and Tableaux, IMA Volumes in Math and its Applications (D. Stanton, ED.), Southend on Sea, UK, 19 (1990), 125–144.

[7] Lascoux, A., and Schützenberger, M.-P., Polynômes de Schubert C. R. Acad. Sci. Paris Sér. I Math. 294:13 (1982), 447-450.

[8] Macdonald, I. G., A new class of symmetric polynomials, Actes du 20^e Séminaire Lotharingien, Publ. Inst. Rech. Math. Av., 372 (1988).

[9] Schensted, C. E., Longest increasing and decreasing subsequences, Canad. J. Math., 13 (1961), 179-191.

[10] Stanley, R. P., Enumerative Combinatorics, Cambridge University Press, Cambridge, United Kingdom, 2 (1999).

Department of Mathematics, University of Pennsylvania, Philadelphia, PA, 19104

E-mail address: sarahm2@math.upenn.edu

URL: http://www.math.upenn.edu/ saramh2

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Matrix compositions

Emanuele Munarini, Maddalena Poneti, and Simone Rinaldi

Abstract. In this paper we study the class of *m-row matrix compositions* (briefly, m-compositions), i.e. matrices with nonnegative integer entries, having m rows, and whose columns are different from the zero vector. We provide enumeration results, combinatorial identities, and various combinatorial interpretations. In particular we extend to the m-dimensional case most of the combinatorial properties of ordinary compositions.

Résumé. Dans cet article nous étudions la classe des *compositions de matrices de m-lignes* (appelées simplement m-compositions), dont les éléments sont des entiers positifs ou nuls, et sans vecteur colonne nul. Nous présentons, outre des interprétations combinatoires, leur énumération ainsi que des identités combinatoires. En particulier nous étendons au cas m-dimensionnel la plupart des propriétés combinatoires des compositions usuelles d'entiers.

1. Introduction

A *composition* (sometimes called *ordered partition*) of a natural number n is any k-tuple $\gamma = (x_1, \dots, x_k)$ of positive integers such that $x_1 + \cdots + x_k = n$. The elements x_i, k and n are the parts, the length and the order of γ, respectively. It is well known that there are $\binom{n-1}{k-1}$ compositions of length k of n and 2^{n-1} compositions of n, when $n \geq 1$. Compositions are very well known combinatorial objects [**1, 9, 13**] and several of their properties have been studied in some recent papers, as in [**7, 10, 14, 15, 17, 18, 23**].

In [**12**] the authors extended the definition of ordinary compositions introducing 2-*compositions* in order to have a bijection between this class and the class of L-convex polyominoes. Such an extension to the bidimensional case can be immediately generalized to the m-dimensional case. Indeed, for any positive integer m, an *m-row matrix composition*, or *m-composition* for short, is an $m \times k$ matrix with nonnegative integer entries

$$M = \begin{bmatrix} x_{11} & \dots & x_{1k} \\ \vdots & & \vdots \\ x_{m1} & \dots & x_{mk} \end{bmatrix}$$

whose columns are different from the zero vector. We say that the number k of columns is the *length* of the composition. Moreover we say that M is an *m-composition* of a nonnegative integer n if the sum of all its elements is exactly n. We will write $\sigma(M)$ for the sum of all the elements of the matrix M. For instance, there are seven 2-compositions of 2:

$$\begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

The aim of this paper is to study the class of m-compositions by several points of view, and to extend to the m-dimensional case most of the combinatorial properties of ordinary compositions.

2000 *Mathematics Subject Classification.* Primary 05A17; Secondary 05A15.
Key words and phrases. compositions, integer partitions, enumeration.

We also remark that our matrix compositions are very similar to the *vector compositions* [**1**, p.57] defined by P. A. MacMahon [**20, 21, 22**] and studied for instance in [**2, 3, 4**]. Another extension of ordinary compositions is described in [**19**].

2. Combinatorial identities

As a first step we present several identities about m-compositions, obtained using elementary combinatorial arguments. Since some of the proofs in this section are rather simple, sometimes they will only be sketched.

Let us start by recalling some basic definitions and properties of multisets. A *multiset* on a set X is a function $\mu : X \to \mathbb{N}$. The *multiplicity* of an element $x \in X$ is $\mu(x)$. The *order* of μ is the sum $\operatorname{ord}(\mu)$ of the multiplicities of the elements of X, i.e. $\operatorname{ord}(\mu) = \sum_{x\in X} \mu(x)$. The number of all multisets of order k on a set of size n is the *multiset coefficient*

$$\left(\!\!\binom{n}{k}\!\!\right) = \frac{n^{\overline{k}}}{k!} = \frac{n(n+1)\dots(n+k-1)}{k!}\,.$$

Let $\mathcal{C}^{(m)}_{n,k}$ be the set of all m-compositions of n of length k and let $c^{(m)}_{n,k} = |\mathcal{C}^{(m)}_{n,k}|$. Similarly let $\mathcal{C}^{(m)}_{n}$ be the set of all m-compositions of n and let $c^{(m)}_{n} = |\mathcal{C}^{(m)}_{n}|$. Let us observe that any $M \in \mathcal{C}^{(m)}_{n+m,k+1}$ can be decomposed into two parts: the first column, equivalent to a multiset of $[m] = \{1,\dots,m\}$ of nonzero order i and the rest of the matrix, that is any m-composition of $n+m-i$ of length k. Hence it follows the recurrence:

$$c^{(m)}_{n+m,k+1} = \sum_{i=1}^{n+m-k} \left(\!\!\binom{m}{i}\!\!\right) c^{(m)}_{n+m-i,k}\,. \tag{2.1}$$

The same argument yields the identity

$$c^{(m)}_{n+m} = \sum_{i=1}^{n+m} \left(\!\!\binom{m}{i}\!\!\right) c^{(m)}_{n+m-i}\,. \tag{2.2}$$

Now we will use some arguments based on the Principle of Inclusion-Exclusion. Let A_i be the set of all m-compositions M of $n+m$ with a positive entry in position $i1$. Then, since the first column of M is different from the zero vector, it follows that $\mathcal{C}^{(m)}_{n+m} = A_1 \cup \dots \cup A_m$ and from the Principle of Inclusion-Exclusion

$$c^{(m)}_{n+m} = |A_1 \cup \dots \cup A_m| = \sum_{\substack{S\subseteq[m]\\ S\neq\emptyset}} (-1)^{|S|-1} \left|\bigcap_{i\in S} A_i\right|\,.$$

The set $\bigcap_{i\in S} A_i$ is formed of all the m-compositions $M = [x_{i,j}]$ of $n+m$ having positive entries in the first column in the positions indexed by S. If we replace each element $x_{i,1}$, $i \in S$, with $x_{i,1}-1$, we have two cases: the first column of M is the zero vector or it is not. In the first case removing the first column we have an m-compositions of $n+m-|S|$, while in the second case we just have an m-composition of $n+m-|S|$. Hence

$$\left|\bigcap_{i\in S} A_i\right| = 2c^{(m)}_{n+m-|S|}\,.$$

Since this result depends only on the size of S it follows that

$$c^{(m)}_{n+m} = 2\sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} c^{(m)}_{n+m-i}\,. \tag{2.3}$$

For instance for $m = 2,3,4$ we have the recurrences

$$c^{(2)}_{n+2} = 4c^{(2)}_{n+1} - 2c^{(2)}_{n}, \qquad c^{(3)}_{n+3} = 6c^{(3)}_{n+2} - 6c^{(3)}_{n+1} + 2c^{(3)}_{n}, \qquad c^{(4)}_{n+4} = 8c^{(4)}_{n+3} - 12c^{(4)}_{n+2} + 8c^{(4)}_{n+1} - 2c^{(4)}_{n}\,.$$

We remark that the recurrence $c^{(2)}_{n+2}$ was first obtained in [**12**]. Exactly with the same argument we can obtain the following recurrence

$$c^{(m)}_{n+m,k+1} = \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} c^{(m)}_{n+m-i,k} + \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} c^{(m)}_{n+m-i,k+1}\,. \tag{2.4}$$

m/n	0	1	2	3	4	5	6	7	8	9	10	11
0	1	0	0	0	0	0	0	0	0	0	0	0
1	1	1	2	4	8	16	32	64	128	256	512	1024
2	1	2	7	24	82	280	956	3264	11144	38048	129904	443520
3	1	3	15	73	354	1716	8318	40320	195444	947380	4592256	22260144
4	1	4	26	164	1031	6480	40728	255984	1608914	10112368	63558392	399478064
5	1	5	40	310	2395	18501	142920	1104060	8528890	65885880	508970002	3931805460
6	1	6	57	524	4803	44022	403495	3698352	33898338	310705224	2847860436	26102905368

FIGURE 1. Table of the numbers $c_n^{(m)}$, with $m = 0, \ldots, 6$.

Let A_i be the set of all matrices $M \in \mathcal{M}_{m,k}(\mathbb{N})$ with the i-th column is equal to the zero vector such that $\sigma(M) = n$. Then $\mathcal{C}_{n,k}^{(m)} = A_1' \cap \ldots \cap A_k'$ and from the Principle of Inclusion-Exclusion

$$c_{n,k}^{(m)} = |A_1' \cap \ldots \cap A_k'| = \sum_{S \subseteq [k]} (-1)^{|S|} \left| \bigcap_{i \in S} A_i \right| .$$

The set $\bigcap_{i \in S} A_i$ is formed of all matrices $M \in \mathcal{M}_{m,k}(\mathbb{N})$ with the zero vector in each column indexed by the elements of S. It corresponds to the set of all multisets of order n on a set of size $mk - m|S|$ and so

$$\left| \bigcap_{i \in S} A_i \right| = \left(\!\!\binom{m(k-|S|)}{n}\!\!\right) .$$

Since this result depends only on the size of S it follows that

$$c_{n,k}^{(m)} = \sum_{i=0}^{k} \binom{k}{i} \left(\!\!\binom{m(k-i)}{n}\!\!\right) (-1)^i . \tag{2.5}$$

Moreover

$$c_n^{(m)} = \sum_{k=0}^{n} c_{n,k}^{(m)} = \sum_{k=0}^{n} \sum_{i=0}^{k} \binom{k}{i} \left(\!\!\binom{m(k-i)}{n}\!\!\right) (-1)^i . \tag{2.6}$$

This argument can be easily generalized as follows. Consider the set $\mathcal{C}_k^{(m)}(r_1, \ldots, r_m)$ of all m-compositions of length k where the i-th row has sum equal to r_i, for each $i = 1, \ldots, m$, and let $c_k^{(m)}(r_1, \ldots, r_m)$ be its cardinality. Now let A_i denote the set of all matrices $M \in \mathcal{M}_{m,k}(\mathbb{N})$ having the i-th column equal to the zero vector, and row-sums $r_1, \ldots, r_m$. Then $\mathcal{C}_k^{(m)}(r_1, \ldots, r_m) = A_1' \cap \ldots \cap A_k'$, and from the Principle of Inclusion-Exclusion

$$c_k^{(m)}(r_1, \ldots, r_m) = |A_1' \cap \ldots \cap A_k'| = \sum_{S \subseteq [k]} (-1)^{|S|} \left| \bigcap_{i \in S} A_i \right| .$$

The set $\bigcap_{i \in S} A_i$ is formed of all matrices $M \in \mathcal{M}_{m,k}(\mathbb{N})$ with the zero vector in each column indexed by the elements of S. The i-th row of such a matrix M corresponds to a multiset of order r_i on a set of size $k - |S|$. Hence it follows that

$$\left| \bigcap_{i \in S} A_i \right| = \left(\!\!\binom{k-|S|}{r_1}\!\!\right) \cdots \left(\!\!\binom{k-|S|}{r_m}\!\!\right) .$$

Since the result depends again only on the size of S it follows that

$$c_k^{(m)}(r_1, \ldots, r_m) = \sum_{i=0}^{k} \binom{k}{i} \left(\!\!\binom{k-i}{r_1}\!\!\right) \cdots \left(\!\!\binom{k-i}{r_m}\!\!\right) (-1)^i . \tag{2.7}$$

The Table in Fig. 1 reports the first terms of the sequences $c_n^{(m)}$, with $m = 0, 1, \ldots, 6$. We remark that for $m \geq 3$ the sequence $c_n^{(m)}$ is not present in [**27**].

3. Enumeration of m-compositions trough formal languages

A large amount of combinatorial properties of m-compositions can simply be derived by encoding them as words on an infinite alphabet. In fact, an m-composition can be viewed as the concatenation of its columns. This implies that the set $\mathcal{C}^{(m)}$ of all m-composition is equivalent to the free language $\mathcal{A}^*$ on the infinite alphabet $\mathcal{A}^{(m)} = \{a_\mu \ : \ \mu \in \mathcal{M}^{(m)}_{\neq 0}\}$, where $\mathcal{M}^{(m)}_{\neq 0}$ is the set of all multisets $\mu : [m] \to \mathbb{N}$ with positive order and the letter a_μ corresponds to the column $\begin{bmatrix}\mu(1) & \ldots & \mu(m)\end{bmatrix}^T$. Substituting each letter a_μ with an indeterminate x_μ, it follows immediately that the generating series of $\mathcal{C}^{(m)}$ is

$$c(X) = \frac{1}{1 - \sum_{\mu \in \mathcal{M}^{(m)}_{\neq 0}} x_\mu} \tag{3.1}$$

where $X = \{x_\mu \ : \ \mu \in \mathcal{M}^{(m)}_{\neq 0}\}$. In particular, for $x_\mu = x^{\operatorname{ord}(\mu)}$ we get the generating series

$$c^{(m)}(x) = \sum_{n \geq 0} c^{(m)}_n \, x^n = \frac{1}{1 - h(x)} \quad \text{where} \quad h(x) = \sum_{k \geq 1} \left(\!\!\binom{m}{k}\!\!\right) x^k = \frac{1}{(1-x)^m} - 1\,.$$

Hence

$$c^{(m)}(x) = \frac{(1-x)^m}{2(1-x)^m - 1}\,, \tag{3.2}$$

from which we can derive the recurrence (2.3) already obtained in the previous section. Similarly, for $x_\mu = x^{\operatorname{ord}(\mu)} y$ we get the generating series

$$c(x,y)^{(m)} = \sum_{n,k \geq 0} c^{(m)}_{n,k} \, x^n y^k = \frac{1}{1 - h(x) y}. \tag{3.3}$$

By the series (3.2), and making some easy computations, we obtain the following results:

(1) a recurrence relation for the numbers $c^{(m)}_{n+1}$,

$$c^{(m)}_{n+1} = -\delta_{n,0} + 2c^{(m)}_n + \sum_{k=0}^{n} \binom{m+k-1}{k+1} c^{(m)}_{n-k}$$

which generalizes the following identity satisfied by the number $c^{(2)}_n$ of 2-compositions [**12**]: $c^{(2)}_{n+2} = 3c^{(2)}_{n+1} + c^{(2)}_n + \ldots + c^{(2)}_0$.

(2) the following Binet-like formula:

$$c^{(m)}_n = \frac{1}{2}\left[\delta_{n,0} + \frac{1}{m\sqrt[m]{2}} \sum_{k=0}^{m-1} \frac{\omega^k_m}{(x_k)^{n+1}}\right]$$

where $x_k = 1 - \frac{1}{\sqrt[m]{2}}\,\omega^k_m$, $k = 0, 1, \ldots, m-1$, and $\omega_m = e^{2\pi i/m}$ is a primitive root of the unity. From this expression we obtain an asymptotic expansion for the coefficients $c^{(m)}_n$,

$$c^{(m)}_n \sim -\frac{A_0}{2x_0^{n+1}} = \frac{1}{2m(\sqrt[m]{2} - 1)} \left(\frac{\sqrt[m]{2}}{\sqrt[m]{2} - 1}\right)^n \qquad \text{as} \quad n \to \infty\,.$$

In particular we have

$$c^{(m)}_{n+1} \sim \frac{\sqrt[m]{2}}{\sqrt[m]{2} - 1}\, c^{(m)}_n \qquad \text{as} \quad n \to \infty\,.$$

A regular language for m-compositions. Extending the encoding used in [**7**] for the ordinary compositions, we are able to prove that m-compositions can be encoded as words on the alphabet $\mathcal{A}_m = \{a_1, \cdots, a_m, b_1, \ldots, b_m\}$. Let us define a map $\ell : \mathcal{C}^{(m)} \to \mathcal{A}^*_m$ setting

$$\begin{matrix}1\\0\\ \vdots\\0\end{matrix} \stackrel{\ell}{\longmapsto} a_1\,, \quad \ldots \quad \begin{matrix}0\\ \vdots\\0\\1\end{matrix} \stackrel{\ell}{\longmapsto} a_m\,, \qquad +\begin{matrix}1\\0\\ \vdots\\0\end{matrix} \stackrel{\ell}{\longmapsto} b_1\,, \quad \ldots \quad +\begin{matrix}0\\ \vdots\\0\\1\end{matrix} \stackrel{\ell}{\longmapsto} b_m \tag{3.4}$$

and proceeding as follows. First of all write an m-composition M as the formal sum (i.e. juxtaposition) of its columns (as in the previous case). Then write each column as juxtaposition of simple columns where a simple column is a column in which all the entries except one are zero. We stipulate to order the simple columns according to the position of the nonzero entry. At this point write each simple column as juxtaposition of elementary columns, where an elementary column is a column in which all the entries are zero except one equal to 1. Hence, if the nonzero entry of a simple column is k then it will be written as the juxtaposition of k elementary columns. Finally substitute each elementary column with the corresponding letter according to the encoding in (3.4). An example will explain better the correspondence. Consider the 3-composition

$$M = \begin{bmatrix} 2 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix}.$$

Following the described procedure we have

$$M = \begin{matrix} 2 \\ 0 \\ 1 \end{matrix} + \begin{matrix} 0 \\ 1 \\ 0 \end{matrix} + \begin{matrix} 1 \\ 0 \\ 1 \end{matrix} + \begin{matrix} 2 \\ 1 \\ 2 \end{matrix} = \begin{matrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{matrix} + \begin{matrix} 0 \\ 1 \\ 0 \end{matrix} + \begin{matrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{matrix} + \begin{matrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{matrix}$$

and hence $\ell(M) = a_1 a_1 a_3 b_2 b_1 a_3 b_1 a_1 a_2 a_3 a_3$.

Let $\mathcal{L}_m = \ell(\mathcal{C}^{(m)})$ be the language on the alphabet $\mathcal{A}_m$ corresponding to the m-compositions. The words of $\mathcal{L}_m$ are characterized by the following conditions: i) each word begins with one letter $a_1, \ldots, a_m$; ii) each letter a_i or b_i can be followed by any b_j, while it can be followed by a letter a_j only when $i \leq j$. This implies that these words have a unique factorization of the form xy where:

(1) x is a non-empty word of the form $a_1^{i_1} \ldots a_m^{i_m}$, with $i_1, \ldots, i_m \geq 0$;
(2) y is a (possibly empty) word $y = y_1 \ \ldots \ y_k$, with $y_r = b_j \, a_j^{q_j} \ldots a_m^{q_m}$, with $q_j, \ldots, q_m \geq 0$.

According to such a characterization $\mathcal{L}_m$ is a regular language defined by the unambiguous regular expression:

$$\varepsilon + \left(a_1^+ a_2^* \ldots a_m^* + a_2^+ a_3^* \ldots a_m^* + \ldots + a_m^+ \right) \left(b_1 a_1^* a_2^* \ldots a_m^* + b_2 a_2^* \ldots a_m^* + \ldots + b_m a_m^* \right)^*$$

where, as usual, ε denotes the empty word.

4. Combinatorial interpretations

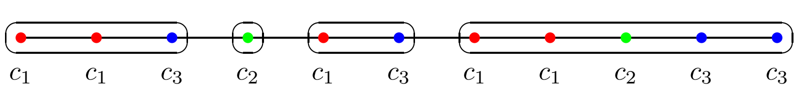
In this section we present three combinatorial interpretations for m-compositions. Here, for brevity's sake, we will give only the basic definitions, even though the relations between the structural properties of the different classes deserve a further investigation.

4.1. Colored linear partitions. m-compositions can be interpreted in terms of linear species [**5, 16**] as follows. Let $C = \{c_1, \ldots, c_m\}$ be a set of colors totally ordered in the natural way $c_1 < \cdots < c_m$. We say that the linearly ordered set $[n] = \{1, 2, \ldots, n\}$ is *m-colored* when each element is colored with one color in C respecting the following condition: if c_i and c_j are the respective colors of two elements x and y, with $x \leq y$, then $i \leq j$. In other words, an m-coloring of $[n]$ is an order preserving map $\gamma : [n] \to C$. We define an m-colored linear partition of $[n]$ as a linear partition in which each block is m-colored.

The m-compositions of length k of n are equivalent to the m-colored linear partitions of $[n]$ with k blocks. Indeed any $M \in \mathcal{C}_{n,k}^{(m)}$ corresponds to the m-colored linear partition π of $[n]$ obtained transforming the i-th column $(h_1, \ldots, h_m)$ of M into the i-th block of π of size $h_1 + \cdots + h_m$ with the first h_1 elements colored with c_1, …, the last h_m elements colored with c_m, for every $1 \leq i \leq k$. For instance, the 3-composition

$$M = \begin{bmatrix} 2 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \end{bmatrix}$$

corresponds to the following 3-colored partition of the set $\{1, \ldots, 11\}$:



Let $\mathbf{Comp}_m$ be linear species of the m-compositions, i.e. the linear species of m-colored linear partitions. To give a structure of this species on a linearly ordered set L is equivalent to assign a linear partition π on L and then an m-coloring, that is an order preserving map in C, on each block of π. Then, if $\mathbf{G}$ denotes the uniform linear species and $\mathbf{Map}_{\neq\emptyset}^{(m)}$ denotes the linear species of the order preserving maps from a nonempty linear order to the set of colors C, we have that

$$\mathbf{Comp}_m = \mathbf{G} \circ \mathbf{Map}_{\neq\emptyset}^{(m)} .$$

An order preserving map $f : [k] \to [m]$ is equivalent to a multiset of order k on the set $[m]$. Hence it follows that

$$\mathbf{Card}(\mathbf{Map}_{\neq\emptyset}^{(m)}; x) = \sum_{k\geq 1} \left(\!\!\binom{m}{k}\!\!\right) x^k = \frac{1}{(1-x)^m} - 1$$

and consequently $\mathbf{Card}(\mathbf{Comp}_m; x) = \mathbf{Card}(\mathbf{G}; x) \circ \mathbf{Card}(\mathbf{Map}_{\neq\emptyset}^{(m)}; x) = c^{(m)}(x)$.

Using this interpretation we can obtain some useful identities. Let $\pi \in \mathbf{Comp}_m[L]$, where $L = \{x_1, \ldots, x_{i+1}, \ldots, x_{i+j+1}\}$ has size $i+j+1$. The element x_{i+1} belongs to a block of the form $\{x_{i-h+1}, \ldots, x_i, x_{i+1}, x_{i+2}, \ldots, x_{i+k+2}\}$ where $h, k \in \mathbb{N}$. Removing such a block, π splits into two m-colored linear partitions of a linear order of size $i-h$ and a linear order of size $j-k$, respectively. Then it follows that

$$c_{i+j+1}^{(m)} = \sum_{h,k\geq 0} \left(\!\!\binom{m}{h+k+1}\!\!\right) c_{i-h}^{(m)} c_{j-k}^{(m)} . \tag{4.1}$$

Recall that $\left(\!\binom{m}{i+j+1}\!\right)$ gives the number of all the order maps $f : [i+j+1] \to [m]$. Suppose that $f(i+1) = k$, with $k \in [m]$. Since f is order preserving, it follows that $f(x) \in [k]$ for every $x \in [i]$ and $f(x) \in \{k, \ldots, m\}$ for every $x \in \{i+2, \ldots, i+j+1\}$. Then

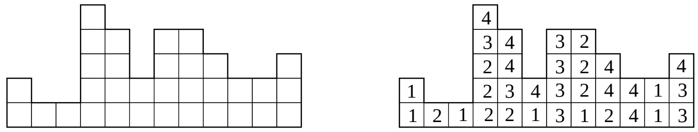
$$\left(\!\!\binom{m}{i+j+1}\!\!\right) = \sum_{k=1}^{m} \left(\!\!\binom{k}{i}\!\!\right) \left(\!\!\binom{m-k+1}{j}\!\!\right) = \sum_{k=0}^{m-1} \binom{i+k}{i} \left(\!\!\binom{m-k}{j}\!\!\right) . \tag{4.2}$$

4.2. Surjective families. Let P_1, ..., P_m and Q be linearly ordered sets. Consider a family $\{f_i : P_i \to Q\}_{i=1}^m$ of order preserving maps with the following property: for every element $q \in Q$ there exists at least one index i and one element $p \in P_i$ such that $q = f_i(p)$. The single maps are not necessarily surjective but every element of the codomain admits at least one preimage along one of the maps of the family. Hence we call *surjective family* any family with such a property.

Now we can ask how many surjective families are there, when $|P_1| = r_1$, ..., $|P_m| = r_m$ and $|Q| = k$. The answer is: $c_k^{(m)}(r_1, \ldots, r_m)$. Indeed given a surjective family $\{f_i : P_i \to Q\}_{i=1}^m$ we can build up an m-composition M of length k as follows. The i-row of M is generated by the map $f_i : P_i \to Q$ taking as entries the numbers of the preimages of the elements of Q along f_i, that is defining it as $[\ |f_i^\bullet(1)| \ldots |f_i^\bullet(k)|\ ]$, where $f_i^\bullet(y)$ denotes the set of all preimages of y along f_i. Clearly the sum of this row is $|P_i| = r_i$. Moreover any column of M is different from the zero vector for the characterizing property of the surjective families. So, finally, we have that M is an m-composition of length k with row-sum vector $(r_1, \ldots, r_k)$.

4.3. Labelled bargraphs. A *bargraph* is a column-convex polyomino, such that the lower edge lies on the horizontal axis. It is uniquely defined by the heights of its columns, see Figure 2 (a). The enumeration of bargraphs according to perimeter, area, and site-perimeter has been treated in [**25, 26**], related to the study of percolation models, and more recently, by an analytical point of view, in [**8**]. For basic definitions on polyominoes we refer to [**6**].

Here we deal with *labelled bargraphs*, i.e. bargraphs whose cells are all labelled with positive integer numbers, and such that, for each column, the label of a cell is less then or equal to the label of the cell immediately above (if any), see Figure 2 (b). The degree is the maximal label of the bargraph. For any given $m \geq 1$, every m-composition of an integer n can be represented as a labelled bargraph of degree $j \leq m$ having n cells, as follows. Let M be an m-composition of $n \geq 0$, having length k, and let $\mathbf{c}_i^T = (a_{1j}, \ldots, a_{mj})$ be the j-th column of M. We build a bargraph made of k columns, of degree m at most, where the j-th column has exactly $a_{1j} + \ldots + a_{mj}$ cells, and a_{ij} is the number of cells with label i in the j-th column, which are placed, according to the definition of labelled bargraph, just above the cells



(a) (b)

FIGURE 2. (a) a bargraph; (b) a labelled bargraph of degree 4.

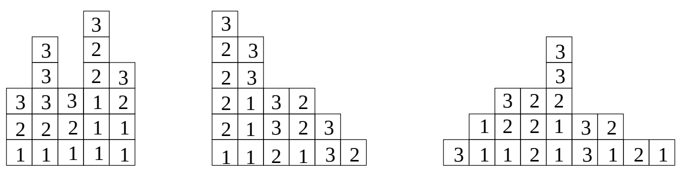
with label $i-1$ (if any). For instance, the bargraph in Fig. 2 (b) is associated with the 4-composition of 33:

$$\begin{bmatrix} 2 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 3 & 1 & 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 4 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 2 & 1 & 0 & 0 & 2 & 2 & 0 & 1 \end{bmatrix}.$$

Of course, ordinary compositions (i.e. 1-compositions) are represented as bargraphs of degree 1, i.e. the usual bargraphs, as already pointed out in [**23**].

Some subclasses of m-compositions. The simple correspondence between m-compositions and labelled bargraphs can be applied to determine bijections between particular subclasses. So, for instance we can consider:

(1) the set of bargraphs having all the m labels in each column (Fig. 3 (a)); it corresponds to the set of m-compositions containing no 0s. The generating function of such objects is $1+\frac{\frac{x^m}{(1-x)^m}}{1-\frac{x^m}{(1-x)^m}} = \frac{1}{1-\left(\frac{x}{1-x}\right)^m}$.



(a) (b) (c)

FIGURE 3. Labelled bargraphs of degree 3: (a) having all the labels in each of its columns; (b) a 3-partition; (c) a labelled stack of degree 3.

(2) the set of *labelled Ferrers diagrams*, i.e. those labelled bargraphs for which each column has height greater than or equal to the height of the column on its right, see Fig. 3 (b). A labelled Ferrers diagram of degree m corresponds to an m-composition such that the sum of the entries of each column is greater than or equal to the sum of the entries of column on its right. We call these objects m-partitions. This definition is motivated by the fact that the ordinary partitions correspond to Ferrers diagrams, i.e. labelled Ferrers diagrams of degree 1. For instance, the bargraph in Fig. 3 (b) corresponds to the 3-partition of 20:

$$\begin{bmatrix} 1 & 3 & 0 & 1 & 0 & 0 \\ 4 & 0 & 1 & 2 & 0 & 1 \\ 1 & 2 & 2 & 0 & 2 & 0 \end{bmatrix}.$$

(3) the set of *labelled stacks*, i.e. of those labelled bargraphs for which each row is connected; these objects have indeed the shape of stack polyominoes, see Fig. 3 (c). A labelled stack of degree m corresponds to an m-composition such that the sequence $c_1, \ldots, c_k$ is unimodal, being c_i the sum of the entries of the i-th column.

The problem of enumerating labelled Ferrers diagrams and labelled stacks has been solved in [**24**] in a more general context.

5. Combinatorial properties of m-compositions

5.1. Cassini-like identities. In [**12**] it has been proved that the numbers $c_n^{(2)}$ of all 2-compositions of n satisfy the Cassini-like identity: $c_n^{(2)} c_{n+2}^{(2)} - (c_{n+1}^{(2)})^2 = -2^{n-1}$, for every $n \geq 1$. Here we prove that such an identity can be generalized to the numbers $c_n^{(m)}$. Specifically we prove that

$$\begin{vmatrix} c_n^{(m)} & c_{n+1}^{(m)} & \cdots & c_{n+m-1}^{(m)} \\ c_{n+1}^{(m)} & c_{n+2}^{(m)} & \cdots & c_{n+m}^{(m)} \\ \vdots & \vdots & & \vdots \\ c_{n+m-1}^{(m)} & c_{n+m}^{(m)} & \cdots & c_{n+2m-2}^{(m)} \end{vmatrix} = (-1)^{\lfloor m/2 \rfloor}\, 2^{n-1} \tag{5.1}$$

for every $m, n \geq 1$. Let $C_n^{(m)} = [\, c_{n+i+j}^{(m)} \,]_{i,j=0}^{m-1}$. Its i-th row is $\mathbf{r}_i = [\, c_{n+i+j}^{(m)} \,]_{j=0}^{m-1}$. In particular, by recurrence (2.3), the last row is

$$\mathbf{r}_m = 2 \sum_{k=1}^{m} \binom{m}{k} (-1)^{k-1} \mathbf{r}_{m-k} = 2 \sum_{k=1}^{m-1} \binom{m}{k} (-1)^{k-1} \mathbf{r}_{m-k} + (-1)^{m-1} 2\, \mathbf{r}_0$$

where $\mathbf{r}_0 = [\, c_{n-1+j}^{(m)} \,]_{j=0}^{m-1}$. Then subtracting to the last row the following linear combination of the previous rows

$$2 \sum_{k=1}^{m-1} \binom{m}{k} (-1)^{k-1} \mathbf{r}_{m-k}$$

the last row of $\det C_n^{(m)}$ becomes $(-1)^{m-1} 2\, \mathbf{r}_0$. Extracting $(-1)^{m-1} 2$ from the last line and then shifting cyclically all rows downward we obtain that

$$\det C_n^{(m)} = 2 \det C_{n-1}^{(m)} .$$

Then, for every $n \geq 1$, it follows that: $\det C_n^{(m)} = 2^{n-1} \det C_1^{(m)}$. So we have to compute only the determinant of the matrix $C_1^{(m)} = [\, c_{i+j+1}^{(m)} \,]_{i,j=0}^{m-1}$. By identity (4.1) we have the decomposition $C_1^{(m)} = L_m M_m L_m^T$ where $L_m = [\, c_{i-j}^{(m)} \,]_{i,j=0}^{m-1}$, $M_m = \left[\left(\!\binom{m}{i+j+1}\!\right) \right]_{i,j=0}^{m-1}$ and L_m^T is the transpose of L_m. Since L_m is a triangular matrix with unitary diagonal elements, it follows that $\det C_1^{(m)} = \det M_m$. Now identity (4.2) implies that $M_m = B_m \widetilde{B}_m$ where $B_m = [\, \binom{i+j}{i} \,]_{i,j=0}^{m-1}$ and $\widetilde{B}_m = \left[\left(\!\binom{m-i}{j}\!\right) \right]_{i,j=0}^{m-1}$. Being $\widetilde{B}_m = J_m B_m$ where $J_m = [\delta_{i+j,m-1}]_{i,j=0}^{m-1}$, it is $M_m = B_m J_m B_m$ and $\det M_m = \det J_m (\det B_m)^2$. Since, as very well known, $\det J_m = (-1)^{\lfloor m/2 \rfloor}$ and $\det B_m = 1$, it follows that $\det M_m = (-1)^{\lfloor m/2 \rfloor}$ and consequently $\det C_1^{(m)} = (-1)^{\lfloor m/2 \rfloor}$. Finally we have: $\det C_n^{(m)} = (-1)^{\lfloor m/2 \rfloor} 2^{n-1}$, for every $n \geq 1$.

5.2. m-compositions without zero rows. In this section we will study the m-compositions in which every row is different from the zero vector. We begin by determining an expression for the number $f_n^{(m)}$ of all such m-compositions of n. Let A_i be the set of all m-compositions $M \in \mathcal{C}_n^{(m)}$ where the i-th row is zero. Then

$$f_n^{(m)} = |A_1' \cap \cdots \cap A_m'| = \sum_{S \subseteq [m]} (-1)^{|S|} \left| \bigcap_{i \in S} A_i \right| .$$

Since $\bigcap_{i \in S} A_i$ is clearly in a bijective correspondence with the set of all $(m-|S|)$-compositions of n, it follows that

$$f_n^{(m)} = \sum_{k=0}^{m} \binom{m}{k} (-1)^k c_n^{(m-k)} = \sum_{k=0}^{m} \binom{m}{k} (-1)^{n-k} c_n^{(k)} . \tag{5.2}$$

On the other hand, the set $\mathcal{C}_n^{(m)}$ can be partitioned according to the number of zero rows and this yields the following identity:

$$c_n^{(m)} = \sum_{k=0}^{m} \binom{m}{k} f_n^{(k)} . \tag{5.3}$$

Clearly this formula can be also obtained formally by inverting (5.2). From (5.2) also follows that the generating series for the numbers $f_n^{(m)}$ is

$$f^{(m)}(x) = \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} c^{(k)}(x) = \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} \frac{(1-x)^k}{2(1-x)^k - 1} .$$

Then this series has the form

$$f^{(m)}(x) = \frac{x^m p_m(x)}{(1-2x)(1-4x+2x^2)\cdots(2(1-x)^m-1)} \tag{5.4}$$

where $p_m(x)$ is a polynomial with degree (less than or) equal to $\binom{m}{2}$. This implies that, for $n \geq 1$, the numbers $f_n^{(m)}$ satisfy a homogeneous linear recurrence with constant coefficients of order $\binom{m+1}{2}$, which can be deduced from the denominator of the series (5.4).

Now we will establish an explicit formula for the numbers $f_n^{(m)}$. Since $f_n^{(m)}$ counts all m-compositions in which every row-sum is nonzero, it immediately follows that

$$f_n^{(m)} = \sum_{k \geq 0} \sum_{\substack{(r_1,\ldots,r_m) \in \mathbb{Z}_+{}^m \\ r_1+\cdots+r_m=n}} c_k^{(m)}(r_1,\ldots,r_m) = \sum_{k \geq 0} \sum_{\substack{\rho \in \mathbb{Z}_+{}^m \\ |\rho|=n}} c_k^{(m)}(\rho)$$

where $\rho = (r_1, \ldots, r_m)$ and $|\rho| = r_1 + \cdots + r_m$. Then, using (2.7), we have the formula

$$f_n^{(m)} = \sum_{\substack{\rho \in \mathbb{Z}_+{}^m \\ |\rho|=n}} \sum_{k \geq 0} \sum_{i=0}^{k} \binom{k}{i} \left(\!\!\binom{k-i}{r_1}\!\!\right) \cdots \left(\!\!\binom{k-i}{r_m}\!\!\right) (-1)^i . \tag{5.5}$$

Clearly $f_n^{(m)} = 0$ whenever $n < m$. Consider now the case $m = n$. In this case we have only the vector $\rho = (1, \ldots, 1)$ and the identity (5.5) becomes

$$f_n^{(n)} = \sum_{k \geq 0} \left[\sum_{i=0}^{k} \binom{k}{i} (k-i)^n (-1)^i \right] .$$

The sum in the brackets is very well known and gives the number of surjective functions from a set of size n to a set of size k. Moreover it can be expressed it terms of the Stirling numbers of the second kind, and precisely it is equal to $\left\{ n \atop k \right\} k!$. Then

$$f_n^{(n)} = \sum_{k=0}^{n} \left\{ n \atop k \right\} k! .$$

But also this sum is very well known, and gives the number t_n of all *preferential arrangements* on a set of size n (sequence A000670 in [**27**]). So, in conclusion, we have that $f_n^{(n)} = t_n$.

This result can be generalized. Indeed in the formula for $f_{n+1}^{(n)}$ we have only the n vectors $\rho = (1, \ldots, 1, 2, 1, \ldots, 1)$. Hence (5.5) becomes

$$f_{n+1}^{(n)} = \frac{n}{2} \sum_{k \geq 0} \sum_{i=0}^{k} \binom{k}{i} (k-i)^n (k-i+1)(-1)^i = \frac{n}{2} \left(\sum_{k \geq 0} \left\{ n+1 \atop k \right\} k! - \sum_{k \geq 0} \left\{ n \atop k \right\} k! \right)$$

that is

$$f_{n+1}^{(n)} = \frac{n}{2} \left(t_{n+1} + t_n \right) .$$

Similarly, when we consider $f_{n+2}^{(n)}$, we have only the n vectors $\rho = (1, \ldots, 1, 3, 1, \ldots, 1)$ and the $\binom{n}{2}$ vectors $\rho = (1, \ldots, 1, 2, 1, \ldots, 1, 2, 1, \ldots, 1)$. Hence (5.5), after simplification, becomes

$$f_{n+2}^{(n)} = \frac{n}{24} \left[(3n+1) t_{n+2} + 6(n+1) t_{n+1} + (3n+5) t_n \right] .$$

All these results suggest that there exist polynomials $p_i^{(k)}(x)$ such that

$$f_{n+k}^{(n)} = \sum_{i=0}^{k} p_i^{(k)}(n) \, t_{n+i} .$$

The nature of such polynomials needs some further investigations.

m/n	0	1	2	3	4	5	6	7	8	9	10	11
0	1	0	0	0	0	0	0	0	0	0	0	0
1	1	1	2	2	4	4	8	8	16	16	32	32
2	1	2	5	8	18	28	62	96	212	328	724	1120
3	1	3	9	19	48	96	236	468	1146	2270	5556	11004
4	1	4	14	36	101	240	648	1520	4082	9560	25660	60088
5	1	5	20	60	185	501	1470	3910	11390	30230	88002	233530
6	1	6	27	92	309	930	2939	8640	27048	79280	247968	726672

Figure 4. Table of the numbers $p_n^{(m)}$.

5.3. m-compositions with palindromic rows. An ordinary composition is *palindromic* when its elements are the same in the given or in the reverse order. In the literature palindromic compositions have been studied by various authors [**10, 11, 23**]. Here we generalize this definition to the m-compositions saying that an m-composition is palindromic when all its rows are palindromic. For instance the following is a palindromic 4-composition of length 5 of 24:

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 2 & 0 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 3 & 1 & 1 & 1 & 3 \end{bmatrix}.$$

Clearly every m-compositions with palindromic rows has the form $[M|M_s]$ when its length is even and the form $[M|\mathbf{v}|M_s]$ when its length is odd, where M is an arbitrary m-composition, M_s is the specular m-composition obtained from M by reversing every row and $\mathbf{v}$ is an arbitrary column vector. Hence the generating series for the m-compositions with palindromic rows is given by

$$p^{(m)}(x) = \sum_{n\geq 0} p_n^{(m)}\, x^n = c^{(m)}(x^2) + \left(\frac{1}{(1-x)^m} - 1\right) c^{(m)}(x^2) = \frac{1}{(1-x)^m}\; c^{(m)}(x^2) = \frac{(1+x)^m}{2(1-x^2)^m - 1}\,.$$

From this identity it immediately follows that

$$p_n^{(m)} = \sum_{k=0}^{\lfloor n/2 \rfloor} \left(\!\!\binom{m}{n-2k}\!\!\right) c_k^{(m)}\,.$$

The first terms of $p_n^{(m)}$ are reported in Fig. 4. Let now $q_n^{(m)}$ be the number of all m-compositions of n with palindromic non zero rows. With arguments completely similar to the ones used in the case of ordinary m-compositions we have that

$$p_n^{(m)} = \sum_{k=0}^{m} \binom{m}{k} q_n^{(k)}, \qquad\qquad q_n^{(m)} = \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} p_n^{(k)}\,.$$

Notice that when $n = m$ there is just one n-composition with palindromic rows, given by the column vector with all entries equal to 1. Hence $q_n^{(n)} = 1$.

5.4. m-compositions of Carlitz type. We say that an m-composition is of *Carlitz type* when no two adjacent columns are equal. When $m = 1$ we obtain the ordinary *Carlitz compositions* [**9**]. As in Section 2, also m-compositions of Carlitz type can be viewed as words on the infinite alphabet $\mathcal{A}^{(m)} = \{a_\mu \;:\; \mu \in \mathcal{M}_{\neq 0}^{(m)}\}$. Let $\mathcal{Z}$ be the set of all words corresponding to the m-composition of Carlitz type and let $\mathcal{Z}_\mu$ be the subset of $\mathcal{Z}$ formed exactly by the words ending with a_μ, for every $\mu \in \mathcal{M}_{\neq 0}^{(m)}$. It immediately follows that

$$\mathcal{Z} = 1 + \sum_{\mu \in \mathcal{M}_{\neq 0}^{(m)}} \mathcal{Z}_\mu \qquad \text{and} \qquad \mathcal{Z}_\mu = (\mathcal{Z} - \mathcal{Z}_\mu) a_\mu \qquad \forall\, \mu \in \mathcal{M}_{\neq 0}^{(m)}\,.$$

m/n	0	1	2	3	4	5	6	7	8	9	10	11
0	1	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	3	4	7	14	23	39	71	124	214
2	1	2	5	18	53	162	505	1548	4756	14650	45065	138622
3	1	3	12	58	255	1137	5095	22749	101625	454116	2028939	9065145
4	1	4	22	136	793	4660	27434	161308	948641	5579224	32811986	192971168
5	1	5	35	265	1925	14056	102720	750255	5480235	40031030	292408771	2135917405
6	1	6	51	458	3984	34788	303902	2654064	23179743	202445610	1768099107	15442052496

FIGURE 5. Table of the numbers $z_n^{(m)}$.

In order to obtain the generating series associated with the languages $\mathcal{Z}$ and $\mathcal{Z}_\mu$ it is sufficient to replace the letter a_μ with the indeterminate x_μ, thus obtaining the linear system

$$z(X) = 1 + \sum_{\mu \in \mathcal{M}^{(m)}_{\neq 0}} z_\mu(X) \qquad \text{and} \qquad z_\mu(X) = (z(X) - z_\mu(X))x_\mu \qquad \forall\, \mu \in \mathcal{M}^{(m)}_{\neq 0}$$

from which

$$z_\mu(X) = \frac{x_\mu}{1+x_\mu}\, z(X) \qquad \text{and then} \qquad z(X) = \frac{1}{1 - \displaystyle\sum_{\mu \in \mathcal{M}^{(m)}_{\neq 0}} \frac{x_\mu}{1+x_\mu}}\,.$$

Setting $x_\mu = x^{\operatorname{ord}(\mu)}$, we obtain the generating series for the coefficients $z_n^{(m)}$ giving the number of all m-compositions of Carlitz type of n. Specifically we have

$$z^{(m)}(x) = \sum_{n\geq 0} z_n^{(m)}\, x^n = \frac{1}{1 - \displaystyle\sum_{k\geq 1} \left(\!\!\binom{m}{k}\!\!\right) \frac{x^k}{1+x^k}}\,. \tag{5.6}$$

For $m = 1$ we reobtain the generating series for the ordinary Carlitz compositions. The sequence $z_n^{(1)}$ appears in [**27**] as the sequence #A003242, while for $m \geq 2$ the corresponding sequences are absent. The first terms of $z_n^{(m)}$ are reported in Fig. 5.

From series (5.6) it is possible to obtain the following explicit formula for the numbers $z_n^{(m)}$. Indeed

$$\begin{aligned}
z^{(m)}(x) &= \sum_{k\geq 0} \left(\sum_{n\geq 1} \left(\!\!\binom{m}{n}\!\!\right) \frac{x^n}{1+x^n} \right)^k = \sum_{k\geq 0} \sum_{a_1 \geq 1} \left(\!\!\binom{m}{a_1}\!\!\right) \frac{x^{a_1}}{1+x^{a_1}} \cdots \sum_{a_k\geq 1} \left(\!\!\binom{m}{a_k}\!\!\right) \frac{x^{a_k}}{1+x^{a_k}} \\
&= \sum_{k\geq 0} \sum_{a_1,\dots,a_k\geq 1} \left(\!\!\binom{m}{a_1}\!\!\right) \cdots \left(\!\!\binom{m}{a_k}\!\!\right) \frac{x^{a_1}}{1+x^{a_1}} \cdots \frac{x^{a_k}}{1+x^{a_k}} \\
&= \sum_{k\geq 0} \sum_{\substack{a_1,\dots,a_k\geq 1 \\ b_1,\dots,b_k\geq 1}} \left(\!\!\binom{m}{a_1}\!\!\right) \cdots \left(\!\!\binom{m}{a_k}\!\!\right) (-1)^{b_1+\cdots+b_k-k} x^{a_1b_1+\cdots+a_kb_k}\,.
\end{aligned}$$

Then

$$z^{(m)}(x) = \sum_{n\geq 0} \left(\sum_{k\geq 0} \sum_{\substack{\alpha,\beta\in\mathbb{N}_0^k \\ \alpha\cdot\beta = n}} \left(\!\!\binom{m}{\alpha}\!\!\right) (-1)^{|\beta|-k} \right) x^n$$

where if $\alpha = (a_1, \dots, a_k)$ and $\beta = (b_1, \dots, b_k)$ then $\alpha \cdot \beta = a_1b_1 + \cdots + a_kb_k$, $|\beta| = b_1 + \cdots + b_k$ and $\left(\!\binom{m}{\alpha}\!\right) = \left(\!\binom{m}{a_1}\!\right) \dots \left(\!\binom{m}{a_k}\!\right)$. Finally, we have the following expression

$$z_n^{(m)} = \sum_{k\geq 0} \sum_{\substack{\alpha,\beta\in\mathbb{N}_0^k \\ \alpha\cdot\beta = n}} \left(\!\!\binom{m}{\alpha}\!\!\right) (-1)^{|\beta|-k}\,.$$

With the same argument used in [**9**] by Carlitz it is possible to obtain the following expression for the series $z^{(m)}(x)$:

$$z^{(m)}(x) = \frac{1}{1 + \sum_{k\geq 1} (-1)^k \frac{1-(1-x^k)^m}{(1-x^k)^m}} \,. \tag{5.7}$$

Let now $g_n^{(m)}$ be the number of all m-compositions of Carlitz type of n without zero rows. With arguments completely similar to the ones used in the case of ordinary m-compositions we have that

$$z_n^{(m)} = \sum_{k=0}^{m} \binom{m}{k} g_n^{(k)}, \qquad g_n^{(m)} = \sum_{k=0}^{m} \binom{m}{k} (-1)^{m-k} z_n^{(k)} \,.$$

References

[1] G. E. Andrews, *The Theory of Partitions*, Encyclopedia of Mathemetics and its Applications, Addison-Wesley Publishing Company, Reading, Massachusetts, 1976.

[2] G. E. Andrews, *The Theory of Compositions, II: Simon Newcom's problem*, Utilitas Math. **7** (1975), 33–54.

[3] G. E. Andrews, *The Theory of Compositions, I: The ordered factorizations of n and a conjecture of C. Long*, Canadian Math. Bull. **18** (1976), 479–484.

[4] G. E. Andrews, *The Theory of Compositions, III: The MacMahon formula and the Stanton-Cowan numbers*, Utilitas Math. **9** (1976), 283–290.

[5] F. Bergeron, G. Labelle, P. Leroux, *Combinatorial Species and Tree-like Structures*, Encyclopedia of Mathematics and Its Applications **67**, G.-C. Rota editor, Cambridge University Press, Cambridge, 1998.

[6] M. Bousquet-Mélou, *A method for the enumeration of various classes of column-convex polygons*, Discrete Math. **154** (1996), 1–25.

[7] A. Björner, R. Stanley, *An analogue of Young's lattice for compositions*, preprint 2005.

[8] M. Bousquet-Mèlou, A. Rechnitzer, *The site-perimeter of bargraphs*, Advances in Applied Mathematics **31** (2003), 86–112.

[9] L. Carlitz, *Restricted compositions*, The Fibonacci Quarterly **14** (1976), 254–264.

[10] P. Chinn, S. Heubach, *Compositions of n with no occurrence of k*, Proceedings of the Thirty-Fourth Southeastern International Conference on Combinatorics, Graph Theory and Computing. Congr. Numer. **164** (2003), 33–51.

[11] P. Chinn, S. Heubach *Integer sequences related to compositions without 2's*, Journal of Integer Sequences, 6, 2003.

[12] G. Castiglione, A. Frosini, E. Munarini, A. Restivo, S. Rinaldi, *Enumeration of L-convex polyominoes II. Bijection and area*, Proceedings of FPSAC'05, Formal Power Series and Algebraic Combinatorics 2005, Taormina, #49.

[13] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht-Holland, Boston, 1974.

[14] S. Heubach, T. Mansour, *Compositions of n with parts in a set*, Proceedings of the Thirty-Fifth Southeastern International Conference on Combinatorics, Graph Theory and Computing. Congr. Numer. **168** (2004), 127–143.

[15] S. Heubach, T. Mansour, *Counting Rises, Levels, and Drops in Compositions*, Integers **A11**, 2005.

[16] A. Joyal, *Une théorie combinatoire des séries formelles*, Adv. in Math. **42** (1981), 1–82.

[17] A. Knopfmacher, H. Prodinger, *On Carlitz compositions*, European J. Combin. **19** (1998), 579–589.

[18] A. Knopfmacher, N. Robbins, *Compositions with parts constrained by the leading summand*, to appear in Ars Combinatorica.

[19] Z. Lin, H. Rui, *Cyclotomic Q-Schur algebras and Schur-Weyl duality*, preprint.

[20] P. A. MacMahon, *Memoir on the theory of the compositions of numbers*, Phil. Trans. Royal Soc. **184** (1894), 835–901.

[21] P. A. MacMahon, *Second memoir on the compositions of numbers*, Phil. Trans. Royal Soc. **207** (1908), 65–134.

[22] P. A. MacMahon, *"Combinatory Analysis"*, vol. 1, Cambridge Univerity Press, Cambridge 1915 (reprinted: Chelsea, New York 1960).

[23] D. Merlini, F. Uncini, C. Verri, *A unified approach to the study of general and palindromic compositions*, Integers **4** (2004), A23.

[24] E. Munarini, *Culnum enriched stacks*, preprint.

[25] A. L. Owczarek, T. Prellberg, *Exact Solution of the Discrete (1+1)-dimensional SOS Model with Field and Surface Interactions*, J. Stat. Phys. **70** (1993), 1175–1194.

[26] T. Prellberg, R. Brak, *Critical Exponents from Non-Linear Functional Equations for Partially Directed Cluster Models*, J. Stat. Phys. **78** (1995), 701–730.

[27] N.J.A. Sloane, *On-Line Encyclopedia of Integer Sequences*, `http://www.research.att.com/~njas/sequences/`.

POLITECNICO DI MILANO, DIPARTIMENTO DI MATEMATICA, PIAZZA LEONARDO DA VINCI 32, 20133 MILANO, ITALY.
E-mail address: `munarini@mate.polimi.it`

UNIVERSITÀ DI FIRENZE, DIPARTIMENTO DI MATEMATICA, VIALE MORGAGNI, 65, 50134 FIRENZE, ITALY.
E-mail address: `poneti@unifi.it`

UNIVERSITÀ DI SIENA, DIPARTIMENTO DI MATEMATICA, VIA DEL CAPITANO 15, 53100, SIENA, ITALY.
E-mail address: `rinaldi@unisi.it`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Combinatorial aspects of elliptic curves

Gregg Musiker

ABSTRACT. Given an elliptic curve C, we study here the number $N_k = \#C(\mathbb{F}_{q^k})$ of points of C over the finite field $\mathbb{F}_{q^k}$. We obtain two combinatorial formulas for N_k. In particular we prove that $N_k = -\mathcal{W}_k(q,t)|_{t=-N_1}$ where $\mathcal{W}_k(q,t)$ is a (q,t)-analogue for the number of spanning trees of the wheel graph.

RÉSUMÉ. Étant donnée une courbe elliptique C on étudie le nombre $N_k = \#C(\mathbb{F}_{q^k})$ de points de C dans le corps fini $\mathbb{F}_{q^k}$. On obtient deux formules combinatoires pour N_k. En particulier on démontre que $N_k = -\mathcal{W}_k(q,t)|_{t=-N_1}$ oú $\mathcal{W}_k(q,t)$ est une (q,t)-extension du nombre des arbres recouvrants du graphe roue.

1. Introduction

An interesting problem at the cross-roads between combinatorics, number theory, and algebraic geometry, is that of counting the number of points on an algebraic curve over a finite field. Over a finite field, the locus of solutions of an algebraic equation is a discrete subset, but since they satisfy a certain type of algebraic equation this imposes a lot of extra structure below the surface. One of the ways to detect this additional structure is by looking at field extensions: the infinite sequence of cardinalities is only dependent on a finite set of data. Specifically the number of points over $\mathbb{F}_q$, $\mathbb{F}_{q^2}$, $\ldots$, and $\mathbb{F}_{q^g}$ will be sufficient data to determine the number of points on a genus g algebraic curve over any other algebraic field extension. This observation begs the question of how the points over higher field extensions correspond to points over the first g extensions. To see this more clearly, we specialize to the case of elliptic curves, where $g = 1$. Letting N_k equal the number of points on C over $\mathbb{F}_{q^k}$, we find a polynomial formula for N_k in terms of q and N_1. Moreover, the coefficients in our formula have a combinatorial interpretation related to spanning trees of the wheel graph.

2. The Zeta Function of a Curve

The zeta function of a curve C is defined to be the exponential generating function

$$Z(C,T) = \exp\left(\sum_{k\geq 1} N_k \frac{T^k}{k}\right).$$

A result due to Weil [**7**] is that the zeta function of an elliptic curve, in fact for any curve, $Z(C,T)$ is rational, and moreover can be expressed as

$$Z(C,T) = \frac{(1-\alpha_1 T)(1-\alpha_2 T)}{(1-T)(1-qT)} = \frac{1-(\alpha_1+\alpha_2)T + \alpha_1\alpha_2 T^2}{(1-T)(1-qT)}.$$

The inverse roots α_1 and α_2 satisfy a functional equation which reduces to

$$\alpha_1\alpha_2 = q$$

in the elliptic curve case.

2000 *Mathematics Subject Classification.* Primary 11G07; Secondary 05C05.
Key words and phrases. elliptic curves, finite fields, zeta functions, spanning trees, Lucas numbers, symmetric functions.
The author thanks the National Science Foundation for their generous support.

Among other things, rationality and the functional equation implies that $N_k = 1 + q^k - \alpha_1^k - \alpha_2^k$, which can be written in plethystic notation as $p_k[1 + q - \alpha_1 - \alpha_2]$. As a special case,

$$\alpha_1 + \alpha_2 = 1 + q - N_1.$$

Hence we can rewrite the Zeta function $Z(C, T)$ totally in terms of q and N_1, hence all the N_k's are actually dependent on these two quantities. The first few formulas are given below.

$$\begin{aligned}
N_2 &= (2+2q)N_1 - N_1^2 \\
N_3 &= (3+3q+3q^2)N_1 - (3+3q)N_1^2 + N_1^3 \\
N_4 &= (4+4q+4q^2+4q^3)N_1 - (6+8q+6q^2)N_1^2 + (4+4q)N_1^3 - N_1^4 \\
N_5 &= (5+5q+5q^2+5q^3+5q^4)N_1 - (10+15q+15q^2+10q^3)N_1^2 \\
&+ (10+15q+10q^2)N_1^3 - (5+5q)N_1^4 + N_1^5
\end{aligned}$$

This data gives rise to our first observation.

Theorem 2.1.

$$N_k = \sum_{i=1}^{k} (-1)^{i+1} P_{i,k}(q) N_1^i$$

where the $P_{i,k}$'s are polynomials with positive integer *coefficients.*

We will prove this in the course of the deriviations in Section 3. Also see [**3**] for a direct proof. This result motivates the combinatorial question: what are the objects that the family of polynomials, $\{P_{i,k}\}$ enumerate?

3. The Lucas Numbers and a (q,t)-analogue

Definition 3.1. We define the $(q,t)-$**Lucas numbers** to be a sequence of polynomials in variables q and t such that $L_n(q,t)$ is defined as

$$L_n(q,t) = \sum_{S \subseteq \{1,2,\ldots,n\} \;:\; S \cap S_1^{(n)} = \phi} q^{\#\text{ even elements in } S}\, t^{\lfloor \frac{n}{2} \rfloor - \#S}. \tag{3.1}$$

Here $S_1^{(n)}$ is the circular shift of set S modulo n, i.e. element $x \in S_1$ if and only if $x - 1 \pmod{n} \in S$. In other words, the sum is over subsets S with no two numbers circularly consecutive.

These polynomials are a generalization of the sequence of Lucas numbers L_n which have the initial conditions $L_1 = 1$, $L_2 = 3$ (or $L_0 = 2$ and $L_1 = 1$) and satisfy the Fibonacci recurrence $L_n = L_{n-1} + L_{n-2}$. The first few Lucas numbers are

$$1, 3, 4, 7, 11, 18, 29, 47, 76, 123, \ldots$$

As described in numerous sources, e.g. [**1**], L_n is equal to the number of ways to color an $n-$beaded necklace black and white so that no two black beads are consecutive. You can also think of this as choosing a subset of $\{1, 2, \ldots, n\}$ with no consecutive elements, nor the pair $1, n$. (We call this circularly consecutive.) Thus letting q and t both equal one, we get by definition that $L_n(1, 1,) = L_n$.

We will prove the following theorem, which relates our newly defined $(q,t)-$Lucas numbers to the polynomials of interest, namely the N_k's.

Theorem 3.2.

$$1 + q^k - N_k = L_{2k}(q,t)\Big|_{t=-N_1} \tag{3.2}$$

for all $k \geq 1$.

To prove this result it suffices to prove that both sides are equal for $k \in \{1, 2\}$, and that both sides satisfy the same three-term recurrence relation. Since

$$\begin{aligned}
L_2(q,t) &= 1 + q + t \quad \text{and} \\
L_4(q,t) &= 1 + q^2 + (2q+2)t + t^2
\end{aligned}$$

we have proven that the initial conditions agree. Note that the sets of (3.1) yielding the terms of these sums are respectively

$$\{1\},\ \{2\},\ \{\ \} \quad \text{and} \quad \{1,3\},\ \{2,4\},\ \{1\},\ \{2\},\ \{3\},\ \{4\},\ \{\ \}.$$

It remains to prove that both sides of (3.2) satisfy the recursion

$$G_{k+1} = (1+q-N_1)G_k - qG_{k-1}$$

for $k \geq 1$.

Proposition 3.1. *For the $(q,t)-$Lucas Numbers $L_k(q,t)$ defined as above,*

$$L_{2k+2}(q,t) = (1+q+t)L_{2k}(q,t) - qL_{2k-2}(q,t). \tag{3.3}$$

Proof. To prove this we actually define an auxiliary set of polynomials, $\{\tilde{L}_{2k}\}$, such that

$$L_{2k}(q,t) = t^k \tilde{L}_{2k}(q,t^{-1}).$$

Thus recurrence (3.3) for the L_{2k}'s translates into

$$\tilde{L}_{2k+2}(q,t) = (1+t+qt)\tilde{L}_{2k}(q,t) - qt^2\tilde{L}_{2k-2}(q,t) \tag{3.4}$$

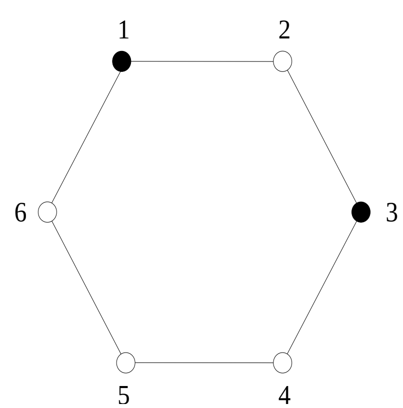
for the $\tilde{L}_{2k}$'s. The $\tilde{L}_{2k}$'s happen to have a nice combinatorial interpretation also, namely

$$\tilde{L}_{2k}(q,t) = \sum_{S \subseteq \{1,2,\ldots,2k\} \ : \ S \cap S_1^{(2k)} = \phi} q^{\#\text{ even elements in } S}\, t^{\#S}.$$

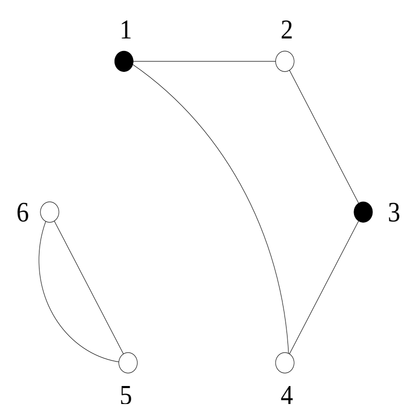
Recall our slightly different description which considers these as the generating function of 2-colored, labeled necklaces. We will find this terminology slightly easier to work with. We can think of the beads labeled 1 through $2k+2$ to be constructed from a pair of necklaces; one of length $2k$ with beads labeled 1 through $2k$, and one of length 2 with beads labeled $2k+1$ and $2k+2$.

Almost all possible necklaces of length $2k+2$ can be decomposed in such a way since the coloring requirements of the $2k+2$ necklace are more stringent than those of the pairs. However not all necklaces can be decomposed this way, nor can all pairs be pulled apart and reformed as a $(2k+2)$-necklace. For example, if $k=2$:

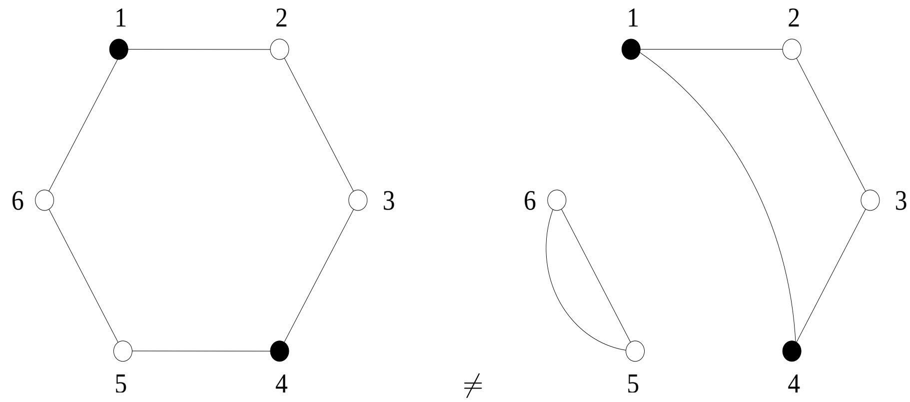
Decomposable



=



Not Decomposable



It is clear enough that the number of pairs is $\tilde{L}_2(q,t)\tilde{L}_{2k}(q,t) = (1+t+qt)\tilde{L}_{2k}(q,t)$. To get the third term of the recurrence, i.e. $qt^2\tilde{L}_{2k-2}$, we must define linear analogues, $\tilde{F}_n(q,t)$'s, of the previous generating function. Just as the $\tilde{L}_n(1,1)$'s were Lucas numbers, the $\tilde{F}_n(1,1)$'s will be Fibonacci numbers.

Definition 3.3. The (twisted) $(q,t)-$Fibonacci polynomials, denoted as $\tilde{F}_n(q,t)$, will be defined as

$$\tilde{F}_k(q,t) = \sum_{S\subseteq\{1,2,\ldots,k-1\} \;:\; S\cap(S_1^{(k-1)}-\{1\})=\phi} q^{\#\text{ even elements in } S}\, t^{\#S}.$$

The summands here are subsets of $\{1,2,\ldots,k-1\}$ such that no two elements are **linearly** consecutive, i.e. we now allow a subset with both the first and last elements. An alternate description of the objects involved are as (linear) chains of $k-1$ beads which are black or white with no two consecutive black beads. With these new polynomials at our disposal, we can calculate the third term of the recurrence, which is the difference between the number of pairs that cannot be recombined and the number of necklaces that cannot be decomposed.

Lemma 3.4. *The number of pairs that cannot be recombined into a longer necklace is* $2qt^2\tilde{F}_{2k-2}(q,t)$.

Proof. We have two cases: either both 1 and $2k+2$ are black, or both $2k$ and $2k+1$ are black. These contribute a factor of qt^2, and imply that beads 2, $2k$, and $2k+1$ are white, or that 1, $2k-1$, and $2k+2$ are white, respectively. In either case, we are left counting chains of length $2k-3$, which have no consecutive black beads. In one case we start at an odd-labeled bead and go to an evenly labeled one, and the other case is the reverse, thus summing over all possibilities yields the same generating function in both cases. ☐

Lemma 3.5. *The number of* $(2k+2)$*-necklaces that cannot be decomposed into a* 2*-necklace and a* $2k$*-necklace is* $qt^2\tilde{F}_{2k-3}(q,t)$.

Proof. The only ones the cannot be decomposed are those which have beads 1 and $2k$ both black. Since such a necklace would have no consecutive black beads, this implies that beads 2, $2k-1$, $2k+1$, and $2k+2$ are all white. Thus we are reduced to looking at chains of length $2k-4$, starting at an odd, 3, which have no consecutive black beads. ☐

Lemma 3.6. *The difference of the quantity referred to in Lemma 3.5 from the quantity in Lemma 3.4 is exactly* $qt^2\tilde{L}_{2k-2}(q,t)$.

Proof. It suffices to prove the relation

$$qt^2\tilde{L}_{2k-2}(q,t) = 2qt^2\tilde{F}_{2k-2}(q,t) - qt^2\tilde{F}_{2k-3}(q,t)$$

which is equivalent to

$$qt^2\tilde{L}_{2k-2}(q,t) = qt^2\tilde{F}_{2k-2}(q,t) + q^2t^3\tilde{F}_{2k-4}(q,t) \tag{3.5}$$

since

$$\tilde{F}_{2k-2}(q,t) = qt\tilde{F}_{2k-4}(q,t) + \tilde{F}_{2k-3}(q,t). \tag{3.6}$$

Note that identity (3.6) simply comes from the fact that the $(2k-2)$nd bead can be black or white. Finally we prove (3.5) by dividing by qt^2, and then breaking it into the cases where bead 1 is white or black. If bead 1 is white, we remove that bead and cut the necklace accordingly. If bead 1 is black, then beads 2 and $2k+2$ must be white, and we remove all three of the beads.

□

With this Lemma proven, the recursion for the $\tilde{L}_{2k}$'s, hence the L_{2k}'s follows immediately. □

Proposition 3.2. *For an elliptic curve C with N_k points over $\mathbb{F}_{q^k}$ we have that*

$$1 + q^{k+1} - N_{k+1} = (1 + q - N_1)(1 + q^k - N_k) - q(1 + q^{k-1} - N_{k-1}).$$

Proof. Recalling that for an elliptic curve C we have the identity $N_k = 1 + q^k - \alpha_1^k - \alpha_2^k$, we can rewrite the statement of this Proposition as

$$\alpha_1^{k+1} + \alpha_2^{k+1} = (\alpha_1 + \alpha_2)(\alpha_1^k + \alpha_2^k) - q(\alpha_1^{k-1} + \alpha_2^{k-1}). \tag{3.7}$$

Noting that $q = \alpha_1\alpha_2$ we obtain this Proposition after expanding out algebraically the right-hand-side of (3.7). □

With the proof of Proposition 3.1 and 3.2, we have proven Theorem 3.2.

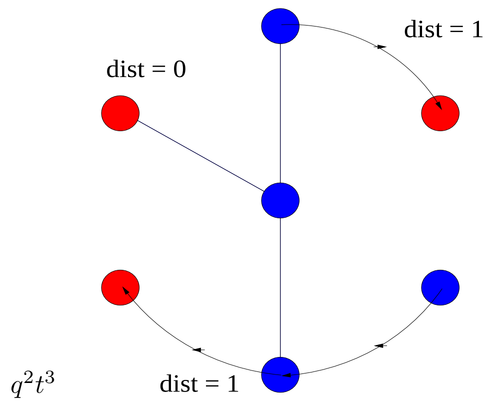
4. $(q,t)-$Wheel Numbers

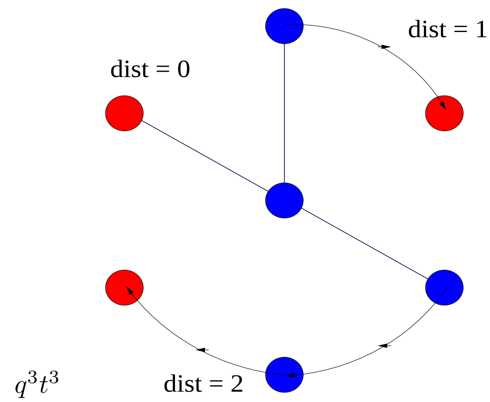
Given that we found the Lucas numbers are related to the polynomial formulas $N_k(q, N_1)$, a natural question concerns how alternative interpretations of the Lucas numbers can help us better understand N_k. As noted in [**1**], [**4**], and [**5**, Seq. A004146], the sequence $\{L_{2n} - 2\}$ counts the number of spanning trees in the wheel graph W_n; a graph which consists of $n+1$ vertices, n of which lie on a circle and one vertex in the center, a hub, which is connected to all the other vertices. This combinatorial interpretation motivates the following definition.

Definition 4.1.

$$\mathcal{W}_n(q,t) = \sum_{T \text{ a spanning tree of } W_n} q^{\text{sum of arc tail distance in } T}\, t^{\#\text{ spokes of } T}.$$

Here the exponent of t counts the number of edges emanating from the central vertex, and the exponent of q requires further explanation. We note that a spanning tree T of W_n consists of spokes and a collection of disconnected arcs on the rim. Further, since there are no cycles, each spoke will intersect exactly one arc. (An isolated vertex is considered to be an arc of length 1.) We imagine the circle being oriented clockwise, and imagine the tail of each arc being the vertex which is the sink for that arc. In the case of an isolated vertex, the lone vertex is the tail of that arc. Since the spoke intersects each arc exactly once, if an arc has length k, meaning that it contains k vertices, there will be k choices of where the spoke and the arc meet. We define the $q-$weight of an arc to be $q^{\text{ number of edges between the spoke and the tail}}$. We define the $q-$weight of the tree to be the product of the $q-$weights for all arcs on the rim of the tree.





This definition actually provides exactly the generating function that we desired, namely we have

Theorem 4.2 (Main Theorem).

$$N_k = -\mathcal{W}_k(q,t)\big|_{t=-N_1}$$

for all $k \geq 1$.

Notice that this yields an exact interpretation of the $P_{i,k}$ polynomials as follows:

$$P_{i,k}(q) = \sum_{T \text{ a spanning tree of } W_n \text{ with exactly } i \text{ spokes}} q^{\text{sum of arc tail distance in } T}.$$

We will prove this Theorem in two different ways. The first method will utilize Theorem 3.2 and an analogue of the bijection given in [**1**] which relates perfect and imperfect matchings of the circle of length $2k$ and spanning trees of W_k. Our second proof will use the observation that we can categorize the spanning trees bases on the sizes of the various connected arcs on the rims. Since this categorization will correspond to partitions, this method will exploit formulas for decomposing p_k into a linear combination of h_λ's, as described in Section 6.

5. First Proof: Bijective

There is a simple bijection between subsets of $[2n]$ with no two elements circularly consecutive and spanning trees of the wheel graph W_n. We will use this bijection to give our first proof of Theorem 4.2. The bijection is as follows:

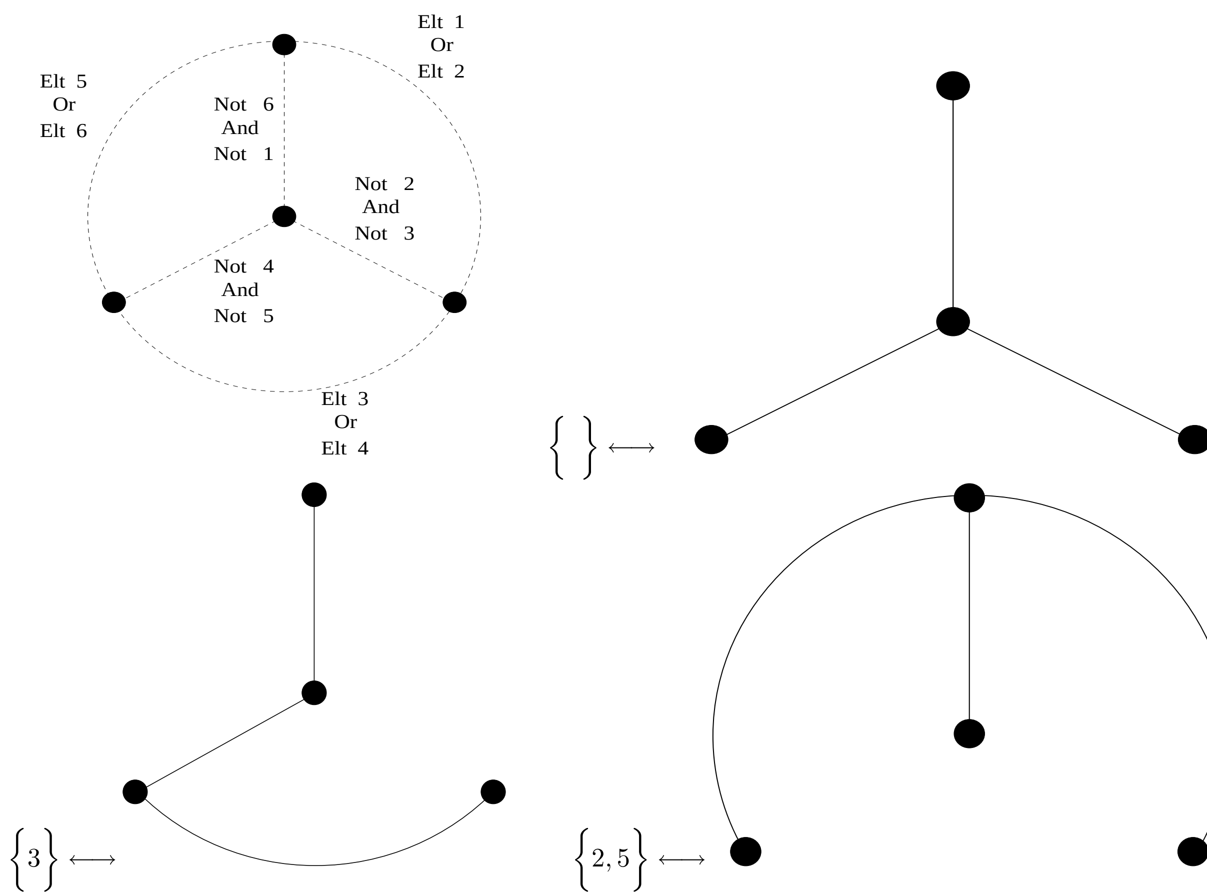
Given a subset S of the set $\{1, 2, \ldots, 2n-1, 2n\}$ with no circularly consecutive elements, we define the corresponding spanning tree T_S of W_n (with the correct q and t weight) in the following way:

1) We will use the convention that the vertices of the graph W_n are labeled so that the vertices on the rim are w_1 through w_n, and the central vertex is w_0.

2) We will exclude the two subsets which consist of all the odds or all the evens from this bijection. Thus we will only be looking at subsets which contain $n-1$ or fewer elements.

3) For $1 \leq i \leq n$, an edge exists from w_0 to w_i if and only if neither $2i-2$ nor $2i-1$ (element 0 is identified with element $2n$) is contained in S.

4) For $1 \leq i \leq n$, an edge exists from w_i to w_{i+1} (w_{n+1} is identified with w_1) if and only if element $2i-1$ or element $2i$ is contained in S.



Proposition 5.1. *Given this construction, T_S is in fact a spanning tree of W_n and further, tree T_S has the same $q-$ and $t-$weights as set S.*

Proof. Suppose that set S contains k elements. From our above restriction, we have that $0 \leq k \leq n-1$. Since S is a k-subset of a $2n$ element set with no circularly consecutive elements, there will be $n-k$ pairs $\{2i-2, 2i-1\}$ with neither element in set S, and k pairs $\{2i-1, 2i\}$ with one element in set S. Consequently, subgraph T_S will consist of exactly $(n-k)+k = n$ edges. Since $n = (\#\text{ vertices of } W_n) - 1$, to prove T_S is a spanning tree, it suffices to show that each vertex of W_n is included. For every oddly-labeled element of $\{1, 2, \ldots, 2n\}$, i.e. $2i-1$ for $1 \leq i \leq n$, we have the following rubric:

1) If $(2i-1) \in S$ then the subgraph T_S contains the edge from w_i to w_{i+1}.
2) If $(2i-1) \notin S$ and additionally $(2i-2) \notin S$, then T_S contains the spoke from w_0 to w_i.
3) If $(2i-1) \notin S$ and additionally $(2i-2) \in S$, then T_S contains the edge from w_{i-1} to w_i.

Since one of these three cases will happen for all $1 \leq i \leq n$, vertex w_i is incident to an edge in T_S. Also, the central vertex, w_0, has to be included since by our restriction, $0 \leq k \leq n-1$ and thus there are $n-k \geq 1$ pairs $\{2i-2, 2i-1\}$ which contain no elements of S.

The number of spokes in T_S is $n-k$ which agrees with the $t-$weight of a set S with k elements. Finally, we prove that the q-weight is preserved by induction on the number of elements in the set S. If set S has no elements, the $q-$weight should be q^0, and spanning tree T_S will consist of n spokes which also has $q-$weight q^0.

Now given a k element subset S ($0 \leq k \leq n-2$), it is only possible to adjoin an odd number if there is a sequence of three consecutive numbers starting with an even, i.e. $\{2i-2, 2i-1, 2i\}$, which is disjoint from S. Such a sequence of S corresponds to a segment of T_S where a spoke and tail of an arc intersect. (Note this includes the case of vertex w_i being an isolated vertex.)

In this case, subset $S' = S \cup \{\text{odd}\}$ corresponds to $T_{S'}$, which is equivalent to spanning tree T_S except that one of the spokes w_0 to w_i has been deleted and replaced with an edge from w_i to w_{i+1}. The arc corresponding to the spoke from w_i will now be connected to the next arc, clockwise. Thus the distance between the spoke and the tail of this arc will not have changed, hence the $q-$weight of $T_{S'}$ will be the same as the $q-$weight of T_S.

Alternatively, it is only possible to adjoin an even number to S if there is a sequence $\{2i-1, 2i, 2i+1\}$ which is disjoint from S. Such a sequence of S corresponds to a segment of T_S where a spoke meets the *end* of an arc. (Note this includes the case of vertex w_i being an isolated vertex.)

Here, subset $S'' = S \cup \{\text{even}\}$ corresponds to $T_{S''}$, which is equivalent to spanning tree T_S except that one of the spokes w_0 to w_{i+1} has been deleted and replaced with an edge from w_i to w_{i+1}. The arc corresponding to the spoke from w_{i+1} will now be connected to the *previous* arc, clockwise. Thus the cumulative change to the total distance between spokes and the tails of arcs will be an increase of one, hence the $q-$weight of $T_{S''}$ will be q^1 times the $q-$weight of T_S.

Since any subset S can be built up this way from the empty set, our proof is complete via this induction. □

Since the two sets we excluded, of size k had $(q,t)-$weights q^0t^0 and q^kt^0 respectively, we have proven Theorem 4.2.

6. Brick Tabloids and Symmetric Function Expansions

Recall that we wrote N_k plethystically as $p_k[1+q-\alpha_1-\alpha_2]$. One advantage of plethystic notation is that we can exploit the following symmetric function identity [**6**, pg. 21]:

$$\sum_{n=0}^{\infty} h_n T^n = \prod_{k\in\mathcal{I}} \frac{1}{1-t_kT} = \exp\left(\sum_{n=1} p_n \frac{T^n}{n}\right) \tag{6.1}$$

where h_n and p_n are symmetric functions in the variables in $\mathcal{I}$. We note that $Z(C,T)$ resembles the right-hand-side of this identity, and consequently, if we had written $Z(C,T)$ as an ordinary power series

$$Z(C,T) = \sum_{k\geq 0} H_k T^k$$

we obtain that $H_k = h_k[1+q-\alpha_1-\alpha_2]$, where h_k denotes the kth homogeneous symmetric function.

Remark 6.1. In fact H_k has an algebraic geometric interpretation also, just as the N_k's did. H_k equals the number of positive divisors of degree k on curve C.

For a general curve we can thus, by plethysm, write cardinalities N_k in terms of H_1 through H_k, using the same coefficients as those that appear in the expansion of p_k in terms of h_1 through h_k:

$$N_k = \sum_{\lambda\vdash k} c_\lambda H_{\lambda_1} H_{\lambda_2} \cdots H_{\lambda_{|\lambda|}} \tag{6.2}$$

where the c_λ can be written down concisely as

$$c_\lambda = (-1)^{l(\lambda)-1} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1, d_2, \ldots, d_k} \tag{6.3}$$

where $l(\lambda)$ denotes the length of λ, which is a partition of k with type $1^{d_1}2^{d_2}\cdots k^{d_k}$.

We give one proof of this using Egecioglu and Remmel's interpretation involving weighted brick tabloids [**2**]. We will give another proof of this, involving a *possibly new* combinatorial interpretation for these coefficients, further on, in Section 7.

A brick tabloid [**2**] of type $\lambda = 1^{d_1}2^{d_2}\cdots k^{d_k}$ and shape μ is a filling of the Ferrers' Diagram μ with bricks of various sizes, d_1 which are 1×1, d_2 which are 2×1, d_3 which are 3×1, etc. The weight of a brick tabloid is the product of the lengths of all bricks at the end of the rows of the Ferrers' Diagram. We let $w(B_{\lambda,\mu})$ denote the weighted-number of brick tabloids of type λ and shape μ, where each tabloid is counted with multiplicity according to its weight.

Proposition 6.1 (Egecioglu-Remmel 1991, [**2**]).

$$p_\mu = \sum_\lambda (-1)^{l(\lambda)-l(\mu)} w(B_{\lambda,\mu})$$

and in particular

$$p_k = \sum_{\lambda} (-1)^{l(\lambda)-1} w(B_{\lambda,(k)}).$$

Brick tabloids of type λ and shape (k) are simply fillings of the $k \times 1$ board with bricks as specified by λ. Thus if divide these tabloids into classes based on the size of the last brick, we obtain, by counting the number of rearrangements, that there are

$$\binom{l(\lambda)-1}{d_1, \ldots, d_i - 1, \ldots, d_k}$$

brick tabloids of type (k) and shape $\lambda = 1^{d_1} 2^{d_2} \cdots k^{d_k}$ which have a last brick of length i.

Since each of these tabloids has weight i, summing up over all possible i, we get that (by abusing multinomial notation slightly)

$$\begin{aligned} w(B_{\lambda,(k)}) &= \sum_{i=0}^{k} i \cdot \binom{l(\lambda)-1}{d_1, \ldots, d_i - 1, \ldots, d_k} \\ &= \left(\sum_{i=0}^{k} i d_i\right) \cdot \binom{l(\lambda)-1}{d_1, \ldots, d_i, \ldots, d_k} \\ &= k \cdot \binom{l(\lambda)-1}{d_1, d_2, \ldots, d_k} = \frac{k}{l(\lambda)} \cdot \binom{l(\lambda)}{d_1, d_2, \ldots, d_k} \end{aligned}$$

Thus after comparing signs, we obtain that c_λ equals exactly the desired expression.

We now specialize to the case of $g = 1$. Here we can write H_k in terms of N_1 and q. We expand the series

$$Z(C,T) = \frac{1 - (1 + q - N_1)T + qT^2}{(1-T)(1-qT)}$$

with respect to T, and obtain $H_0 = 1$ and $H_k = N_1(1 + q + q^2 + \cdots + q^{k-1})$ for $k \geq 1$. Plugging these into formula (6.2), and using (6.3), we get polynomial formulas for N_k in terms of q and N_1, which in fact are an alternative expression for the formulas found in section 2.

$$N_k = \sum_{\lambda \vdash k} (-1)^{l(\lambda)-1} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_k} \left(\prod_{i=1}^{l(\lambda)} (1 + q + q^2 + \cdots + q^{\lambda_i - 1}) \right) N_1^{l(\lambda)}.$$

Thus using these alternative expressions for N_k, we have that Theorem 4.2 is equivalent to the statement

$$\mathcal{W}_k = \sum_{\lambda \vdash k} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_k} \left(\prod_{i=1}^{l(\lambda)} (1 + q + q^2 + \cdots + q^{\lambda_i - 1}) \right) t^{l(\lambda)}.$$

7. Second Proof: Via Symmetric Functions

For our second proof of Theorem 4.2, we start with the observation that the sequence of lengths of all disjoint arcs on the rim of W_n corresponds to a partition of n. We will construct a spanning tree of W_n from the following choices:

First we choose a partition $\lambda = 1^{d_1} 2^{d_2} \cdots k^{d_k}$ of n. We let this dictate how many arcs of each length occur, i.e. we have d_1 isolated vertices, d_2 arcs of length 2, etc. Note that this choice also dictates the number of spokes, which is equal to the number of arcs, i.e. the length of the partition.

Second, we pick an arrangement of $l(\lambda)$ arcs on the circle. After picking one to start with, without loss of generality since we are on a circle, we have

$$\frac{1}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_k}$$

choices for such an arrangement.

Third, we pick which vertex w_i of the rim to start with. There are n such choices.

Fourth, we pick where the $l(\lambda)$ spokes actually intersect the arcs. There will be $|\mathrm{arc}|$ choices for each arc, and the $q-$weight of this sum will be $(1 + q + q^2 + \cdots + q^{|\mathrm{arc}|})$ for each arc.

Summing up all the possibilities yields

$$\mathcal{W}_n = \sum_{\lambda \vdash n} \frac{n}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_k} \left(\prod_{i=1}^{l(\lambda)} (1 + q + q^2 + \cdots + q^{\lambda_i - 1}) \right) t^{l(\lambda)}.$$

As noted in Section 6, these coefficients are exactly the correct expansion coefficients by identities (6.1), (6.3), and plethysm. Thus we have given a second proof of Theorem 4.2.

Remark 7.1. We note that in the course of this second proof we have obtained a combinatorial interpretation for the c_λ's that is distinct from the one given in Egecioglu and Remmel's paper [**2**]. In particular this interpretation does not require weighted counting, only signed counting. Instead of defining c_λ as $(-1)^{l(\lambda)-1} w(B_{\lambda,(k)})$, we could define it as

$$(-1)^{l(\lambda)-1} |CB_{\lambda,(k)}|$$

where we define a new combinatorial class of **circular brick tabloids** which we denote as $CB_{\lambda,\mu}$. We define this for the case of $\mu = (k)$ just as we defined the usual brick tabloids, except we are not filling a $k \times 1$ rectangle, but are filling an annulus of circumference k and width 1 with curved bricks of sizes designated by λ. In this way we mimic our construction of the spanning trees.

Additionally, by using the fact that the power symmetric functions are multiplicative, i.e. $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_r}$, we are able to generalize our definition of circular brick tabloids to allow μ to be any partition. We simply let λ designate what collection of bricks we have to use, and μ determines the filling: we are trying to fill $l(\mu)$ concentric circles where each circle has μ_i spaces. To summarize,

$$p_\mu = \sum_\lambda (-1)^{l(\lambda)-l(\mu)} |CB_{\lambda,\mu}| h_\lambda.$$

Consequently, all identities of [**2**] now involve *cardinalities* of $B_{\lambda,\mu}$, $OB_{\lambda,\mu}$ (Ordered Brick Tabloids), or $CB_{\lambda,\mu}$ and signs depending on $l(\lambda)$ and $l(\mu)$, with no additional weightings needed.

8. Conclusion

The new combinatorial formula for N_k presented in this write-up appears fruitful. It leads one to ask how spanning trees of the wheel graph are related to points on elliptic curves. For instance, is there an involution on (weighted) spanning trees whose fixed points enumerate points on $C(\mathbb{F}_{q^k})$? The fact that the Lucas numbers also enter the picture is also exciting since the Fibonacci numbers and Lucas numbers have so many different combinatorial interpretations, and there is such an extensive literature about them. Perhaps these combinatorial interpretations will lend insight into why N_k depends only on the finite data of N_1 and q for an elliptic curve, and how we can associate points over higher extension fields to points on $C(\mathbb{F}_q)$.

9. Ackowledgements

The author would like to thank Adriano Garsia for his support and many helpful conversations.

References

[1] A. Benjamin and C. Yerger, Combinatorial Interpretations of Spanning Tree Identities, *Bulletin of the Institute for Combinatorics and its Applications*, to appear.

[2] O. Egecioglu and J. Remmel, Brick Tabloids and the Connection Matrices Between Bases of Symmetric Functions, *Disc. Appl. Math.*, **34** (1991), 107-120.

[3] A. Garsia and G. Musiker, *Basics on Hyperelliptic Curves over Finite Fields*, in progress.

[4] B. R. Myers, Number of Spanning Trees in a Wheel, *IEEE Trans. Circuit Theory*, **18** (1971), 280-282.

[5] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, http://www.research.att.com/~njas/sequences/index.html.

[6] R. P. Stanley, *Enumerative Combinatorics Vol. 2, volume 62 of Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge (1999).

[7] A. Weil, *Sur les Courbes Algébriques et les Variétés qui s'en Déduisent*, Hermann, Paris (1948).

Department of Mathematics, UCSD, San Diego, USA, 92037
E-mail address: gmusiker@math.ucsd.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Polynomial realizations of some trialgebras

Jean-Christophe Novelli and Jean-Yves Thibon

Abstract. We realize several combinatorial Hopf algebras based on set compositions, plane trees and segmented compositions in terms of noncommutative polynomials in infinitely many variables. For each of them, we describe a trialgebra structure, an internal product, and several bases.

Résumé. Nous réalisons plusieurs algèbres de Hopf combinatoires dont les bases sont indexées par les partitions d'ensembles ordonnées, les arbres plans et les compositions segmentées en termes de polynômes non-commutatifs en une infinité de variables. Pour chacune d'elles, nous décrivons sa structure de trigèbre, un produit intérieur et plusieurs bases.

1. Introduction

The aim of this note is to construct and analyze several combinatorial Hopf algebras arising in the theory of operads from the point of view of the theory of noncommutative symmetric functions. Our starting point will be the algebra of noncommutative polynomial invariants

$$\mathbf{WQSym}(A) = \mathbb{K}\langle A\rangle^{\mathfrak{S}(A)_{QS}}$$

of Hivert's quasi-symmetrizing action [**8**]. It is known that, when the alphabet A is infinite, $\mathbf{WQSym}(A)$ acquires the structure of a graded Hopf algebra whose bases are parametrized by ordered set partitions (also called set compositions) [**8, 20, 2**]. Set compositions are in one-to-one correspondence with faces of permutohedra, and actually, **WQSym** turns out to be isomorphic to one of the Hopf algebras introduced by Chapoton in [**4**]. From this algebra, Chapoton obtained graded Hopf algebras based on the faces of the associahedra (corresponding to plane trees counted by the little Schröder numbers) and on faces of the hypercubes (counted by powers of 3). Since then, Loday and Ronco have introduced the operads of dendriform trialgebras and of tricubical algebras [**15**], in which the free algebras on one generator are respectively based on faces of associahedras and hypercubes, and are isomorphic (as Hopf algebras) to the corresponding algebras of Chapoton. More recently, we have introduced a Hopf algebra **PQSym**, based on parking functions [**17, 18, 19**], and derived from it a series of Hopf subalgebras or quotients, some of which being isomorphic to the above mentioned ones as associative algebras, but not as Hopf algebras.

In the following, we will show that applying the same techniques, starting from **WQSym** instead of **PQSym**, allows one to recover all of these algebras, together with their original Hopf structure, in a very natural way. This provides in particular for each of them an explicit realization in terms of noncommutative polynomials. The Hopf structures can be analyzed very efficiently by means of Foissy's theory of bidendriform bialgebras [**6**]. A natural embedding of **WQSym** in $\mathbf{PQSym}^*$ implies that **WQSym** is bidendriform, hence, free and self-dual. These properties are inherited by $\mathfrak{TD}$, the free dendriform trialgebra on one generator, and some of them by $\mathfrak{TC}$, the free cubical trialgebra on one generator. A lattice structure on the set of faces of the permutohedron (introduced in [**12**] under the name "pseudo-permutohedron" and rediscovered in [**21**]) leads to the construction of various bases of these algebras. Finally, the natural identification of the homogeneous components of the dual $\mathbf{WQSym}_n^*$ (endowed with the internal product induced by **PQSym**)

2000 *Mathematics Subject Classification.* Primary 05E99, Secondary 16W30, 18D50.
Key words and phrases. Algebraic combinatorics, symmetric functions, dendriform structures, lattice theory.

with the Solomon-Tits algebras (that is, the face algebras of the braid arrangements of hyperplanes) implies that all three algebras admit an internal product.

Notations – We assume that the reader is familiar with the standard notations of the theory of noncommutative symmetric functions [**7, 5**] and with the Hopf algebra of parking functions [**17, 18, 19**]. We shall need an infinite totally ordered alphabet $A = \{a_1 < a_2 < \cdots < a_n < \cdots\}$, generally assumed to be the set of positive integers. We denote by $\mathbb{K}$ a field of characteristic 0, and by $\mathbb{K}\langle A\rangle$ the free associative algebra over A when A is finite, and the projective limit $\operatorname{proj\,lim}_B \mathbb{K}\langle B\rangle$, where B runs over finite subsets of A, when A is infinite. The *evaluation* of a word w is the sequence whose i-th term is the number of times the letter a_i occurs in w. The *standardized word* $\mathrm{Std}(w)$ of a word $w \in A^*$ is the permutation obtained by iteratively scanning w from left to right, and labelling $1, 2, \ldots$ the occurrences of its smallest letter, then numbering the occurrences of the next one, and so on. For example, $\mathrm{Std}(bbacab) = 341624$. For a word w on the alphabet $\{1, 2, \ldots\}$, we denote by $w[k]$ the word obtained by replacing each letter i by the integer $i+k$. If u and v are two words, with u of length k, one defines the *shifted concatenation* $u \bullet v = u \cdot (v[k])$ and the *shifted shuffle* $u \Cup v = u ⧢ (v[k])$, where ⧢ is the usual shuffle product.

2. The Hopf algebra WQSym

2.1. Noncommutative quasi-symmetric invariants. The *packed word* $u = \mathrm{pack}(w)$ associated with a word $w \in A^*$ is obtained by the following process. If $b_1 < b_2 < \ldots < b_r$ are the letters occuring in w, u is the image of w by the homomorphism $b_i \mapsto a_i$. A word u is said to be *packed* if $\mathrm{pack}(u) = u$. We denote by PW the set of packed words. With such a word, we associate the polynomial

$$\mathbf{M}_u := \sum_{\mathrm{pack}(w)=u} w\,. \tag{1}$$

For example, restricting A to the first five integers,

$$\mathbf{M}_{13132} = 13132 + 14142 + 14143 + 24243 + 15152 + 15153 + 25253 + 15154 + 25254 + 35354. \tag{2}$$

Under the abelianization $\chi :\ \mathbb{K}\langle A\rangle \to \mathbb{K}[X]$, the $\mathbf{M}_u$ are mapped to the monomial quasi-symmetric functions M_I ($I = (|u|_a)_{a\in A}$ being the evaluation vector of u).

These polynomials span a subalgebra of $\mathbb{K}\langle A\rangle$, called **WQSym** for Word Quasi-Symmetric functions [**8**] (and called **NCQSym** in [**2**]), consisting in the invariants of the noncommutative version of Hivert's quasi-symmetrizing action [**9**], which is defined by $\sigma \cdot w = w'$ where w' is such that $\mathrm{Std}(w') = \mathrm{Std}(w)$ and $\chi(w') = \sigma \cdot \chi(w)$. Hence, two words are in the same $\mathfrak{S}(A)$-orbit iff they have the same packed word.

WQSym can be embedded in **MQSym** [**8, 5**], by $\mathbf{M}_u \mapsto \mathbf{MS}_M$, where M is the packed $(0,1)$-matrix whose jth column contains exactly one 1 at row i whenever the jth letter of u is a_i. Since the duality in **MQSym** consists in tranposing the matrices, one can also embed **WQSym*** in **MQSym**. The multiplication formula for the basis $\mathbf{M}_u$ follows from that of $\mathbf{MS}_M$ in **MQSym**:

PROPOSITION 2.1. *The product on* **WQSym** *is given by*

$$\mathbf{M}_{u'}\mathbf{M}_{u''} = \sum_{u\in u'*_W u''} \mathbf{M}_u\,, \tag{3}$$

where the convolution $u' *_W u''$ *of two packed words is defined as*

$$u' *_W u'' = \sum_{v,w;u=v\cdot w\,\in\,\mathrm{PW},\mathrm{pack}(v)=u',\mathrm{pack}(w)=u''} u\,. \tag{4}$$

For example,

$$\mathbf{M}_{11}\mathbf{M}_{21} = \mathbf{M}_{1121} + \mathbf{M}_{1132} + \mathbf{M}_{2221} + \mathbf{M}_{2231} + \mathbf{M}_{3321}. \tag{5}$$

Similarly, the embedding in **MQSym** implies immediately that **WQSym** is a Hopf subalgebra of **MQSym**. However, the coproduct can also be defined directly by the usual trick of noncommutative symmetric functions, considering the alphabet A as an ordered sum of two mutually commuting alphabets $A'\hat{+}A''$. First, by direct inspection, one finds that

$$\mathbf{M}_u(A'\hat{+}A'') = \sum_{0\le k\le \max(u)} \mathbf{M}_{(u|_{[1,k]})}(A')\mathbf{M}_{\mathrm{pack}(u|_{[k+1,\max(u)]})}(A''), \tag{6}$$

where $u|_B$ denote the subword obtained by restricting u to the subset B of the alphabet, and now, the coproduct Δ defined by

$$\Delta \mathbf{M}_u(A) = \sum_{0\le k\le \max(u)} \mathbf{M}_{(u|_{[1,k]})} \otimes \mathbf{M}_{\mathrm{pack}(u|_{[k+1,\max(u)})}, \tag{7}$$

is then clearly a morphism for the concatenation product, hence defines a bialgebra structure.

Given two packed words u and v, define the *packed shifted shuffle* $u \Cup_W v$ as the shuffle product of u and $v[\max(u)]$. One then easily sees that

$$\Delta \mathbf{M}_w(A) = \sum_{u,v;\, w\in u\Cup_W v} \mathbf{M}_u \otimes \mathbf{M}_v. \tag{8}$$

For example,

$$\Delta \mathbf{M}_{32121} = 1\otimes \mathbf{M}_{32121} + \mathbf{M}_{11}\otimes \mathbf{M}_{211} + \mathbf{M}_{2121}\otimes \mathbf{M}_1 + \mathbf{M}_{32121}\otimes 1. \tag{9}$$

Packed words can be naturally identified with *ordered set partitions*, the letter a_i at the jth position meaning that j belongs to block i. For example,

$$u = 313144132 \ \leftrightarrow\ \Pi = (\{2,4,7\},\{9\},\{1,3,8\},\{5,6\})\,. \tag{10}$$

To improve the readability of the formulas, we write instead of Π a *segmented permutation*, that is, the permutation obtained by reading the blocks of Π in increasing order and inserting bars | between blocks.

For example,

$$\Pi = (\{2,4,7\},\{9\},\{1,3,8\},\{5,6\}) \ \leftrightarrow 247|9|138|56. \tag{11}$$

On this representation, the coproduct amounts to deconcatenate the blocks, and then standardize the factors. For example, in terms of segmented permutations, Equation (9) reads

$$\Delta \mathbf{M}_{35|24|1} = 1\otimes \mathbf{M}_{35|24|1} + \mathbf{M}_{12}\otimes \mathbf{M}_{23|1} + \mathbf{M}_{24|13}\otimes \mathbf{M}_1 + \mathbf{M}_{35|24|1}\otimes 1. \tag{12}$$

The dimensions of the homogeneous components of **WQSym** are the ordered Bell numbers 1, 1, 3, 13, 75, 541, ... (sequence A000670, [**22**]) so that

$$\dim \mathbf{WQSym}_n = \sum_{k=1}^{n} S(n,k)k! = A_n(2)\,, \tag{13}$$

where $A_n(q)$ are the Eulerian polynomials.

2.2. The trialgebra structure of WQSym. A *dendriform trialgebra* [**15**] is an associative algebra whose multiplication $\odot$ splits into three pieces

$$x\odot y = x\prec y + x\circ y + x\succ y\,, \tag{14}$$

where $\circ$ is associative, and

$$(x\prec y)\prec z = x\prec (y\odot z)\,,\quad (x\succ y)\prec z = x\succ (y\prec z)\,,\quad (x\odot y)\succ z = x\succ (y\succ z)\,, \tag{15}$$

$$(x\succ y)\circ z = x\succ (y\circ z)\,,\quad (x\prec y)\circ z = x\circ (y\succ z)\,,\quad (x\circ y)\prec z = x\circ (y\prec z)\,. \tag{16}$$

It has been shown in [**19**] that the augmentation ideal $\mathbb{K}\langle A_n\rangle^+$ has a natural structure of dendriform trialgebra: for two non empty words $u,v\in A^*$, we set

$$u\prec v = \begin{cases} uv & \text{if } \max(u) > \max(v)\\ 0 & \text{otherwise,}\end{cases} \tag{17}$$

$$u\circ v = \begin{cases} uv & \text{if } \max(u) = \max(v)\\ 0 & \text{otherwise,}\end{cases} \tag{18}$$

$$u\succ v = \begin{cases} uv & \text{if } \max(u) < \max(v)\\ 0 & \text{otherwise.}\end{cases} \tag{19}$$

Theorem 2.2. **WQSym**$^+$ *is a sub-dendriform trialgebra of* $\mathbb{K}\langle A\rangle^+$*, the partial products being given by*

$$\mathbf{M}_{w'} \prec \mathbf{M}_{w''} = \sum_{w=u.v\in w' *_W w'', |u|=|w'|; \max(v)<\max(u)} \mathbf{M}_w, \tag{20}$$

$$\mathbf{M}_{w'} \circ \mathbf{M}_{w''} = \sum_{w=u.v\in w' *_W w'', |u|=|w'|; \max(v)=\max(u)} \mathbf{M}_w, \tag{21}$$

$$\mathbf{M}_{w'} \succ \mathbf{M}_{w''} = \sum_{w=u.v\in w' *_W w'', |u|=|w'|; \max(v)>\max(u)} \mathbf{M}_w, \tag{22}$$

It is known [**15**] that the free dendriform trialgebra on one generator, denoted here by $\mathfrak{TD}$, is a free associative algebra with Hilbert series

$$\sum_{n\geq 0} s_n t^n = \frac{1+t-\sqrt{1-6t+t^2}}{4t} = 1+t+3t^2+11t^3+45t^4+197t^5+\cdots, \tag{23}$$

the generating function of the *super-Catalan*, or *little Schröder* numbers, counting *plane trees.* The previous considerations allow us to give a simple polynomial realization of $\mathfrak{TD}$. Consider the polynomial

$$\mathbf{M}_1 = \sum_{i\geq 1} a_i \ \in \mathbf{WQSym}, \tag{24}$$

Theorem 2.3 ([**19**]). *The sub-trialgebra* $\mathfrak{TD}$ *of* **WQSym**$^+$ *generated by* $\mathbf{M}_1$ *is free as a dendriform trialgebra.*

Based on numerical evidence, we conjecture the following result:

Conjecture 2.4. **WQSym** *is a free dendriform trialgebra.*

The number g'_n of generators in degree n of **WQSym** as a free dendriform trialgebra would then be

$$\sum_{n\geq 0} g'_n t^n = \frac{OB(t)-1}{2OB(t)^2-OB(t)} = t+2\,t^3+18\,t^4+170\,t^5+1\,794\,t^6+21\,082\,t^7+O(t^8). \tag{25}$$

where $OB(t)$ is the generating series of the ordered Bell numbers.

2.3. Bidendriform structure of WQSym. A *dendriform dialgebra*, as defined by Loday [**13**], is an associative algebra D whose multiplication $\odot$ splits into two binary operations

$$x\odot y = x \ll y + x \gg y, \tag{26}$$

called left and right, satisfying the following three compatibility relations for all a, b, and c different from 1 in D:

$$(a\ll b)\ll c = a\ll (b\odot c), \quad (a\gg b)\ll c = a\gg (b\ll c), \quad (a\odot b)\gg c = a\gg (b\gg c). \tag{27}$$

A *codendriform coalgebra* is a coalgebra C whose coproduct Δ splits as $\Delta(c) = \overline{\Delta}(c) + c\otimes 1 + 1\otimes c$ and $\overline{\Delta} = \Delta_{\ll} + \Delta_{\gg}$, such that, for all c in C:

$$(\Delta_{\ll}\otimes Id)\circ\Delta_{\ll}(c) = (Id\otimes\overline{\Delta})\circ\Delta_{\ll}(c), \tag{28}$$

$$(\Delta_{\gg}\otimes Id)\circ\Delta_{\ll}(c) = (Id\otimes\Delta_{\ll})\circ\Delta_{\gg}(c), \tag{29}$$

$$(\overline{\Delta}\otimes Id)\circ\Delta_{\gg}(c) = (Id\otimes\Delta_{\gg})\circ\Delta_{\gg}(c). \tag{30}$$

The Loday-Ronco algebra of planar binary trees introduced in [**14**] arises as the free dendriform dialgebra on one generator. This is moreover a Hopf algebra, which turns out to be self-dual, so that it is also codendriform. There is some compatibility between the dendriform and the codendriform structures, leading to what has been called by Foissy [**6**] a *bidendriform bialgebra*, defined as a bialgebra which is both a dendriform dialgebra and a codendriform coalgebra, satisfying the following four compatibility relations

$$\Delta_{\gg}(a\gg b) = a'b'_{\gg}\otimes a''\gg b''_{\gg} \ + a'\otimes a''\gg b \ + b'_{\gg}\otimes a\gg b''_{\gg} \ + ab'_{\gg}\otimes b''_{\gg} \ + a\otimes b, \tag{31}$$

$$\Delta_{\gg}(a\ll b) = a'b'_{\gg}\otimes a''\ll b''_{\gg} \ + a'\otimes a''\ll b \ + b'_{\gg}\otimes a\ll b''_{\gg}, \tag{32}$$

$$\Delta_{\ll}(a \gg b) = a'b'_{\ll} \otimes a'' \gg b''_{\ll} + ab'_{\ll} \otimes b''_{\ll} + b'_{\ll} \otimes a \gg b''_{\ll}\,, \tag{33}$$

$$\Delta_{\ll}(a \ll b) = a'b'_{\ll} \otimes a'' \ll b''_{\ll} + a'b \otimes a'' + b'_{\ll} \otimes a \ll b''_{\ll} + b \otimes a\,, \tag{34}$$

where the pairs (x', x'') (resp. $(x'_{\ll}, x''_{\ll})$ and $(x'_{\gg}, x''_{\gg})$) correspond to all possible elements occuring in $\overline{\Delta}x$ (resp. $\Delta_{\ll}x$ and $\Delta_{\gg}x$), summation signs being understood (Sweedler's notation).

Foissy has shown [**6**] that a connected bidendriform bialgebra $\mathcal{B}$ is always free as an associative algebra and self-dual as a Hopf algebra. Moreover, its primitive Lie algebra is free, and as a dendriform dialgebra, $\mathcal{B}$ is also free over the space of totally primitive elements (those annihilated by $\Delta_{\ll}$ and $\Delta_{\gg}$). It is also proved in [**6**] that **FQSym** is bidendriform, so that it satisfies all these properties. In [**19**], we have proved that **PQSym**, the Hopf algebra of parking functions, as also bidendriform.

The realization of **PQSym*** given in [**18, 19**] implies that

$$\mathbf{M}_u = \sum_{\mathrm{pack}(\mathbf{a})=u} \mathbf{G}_{\mathbf{a}}\,. \tag{35}$$

Hence, **WQSym** is a subalgebra of **PQSym***. Since in both cases the coproduct correponds to $A \to A' \hat{+} A''$, it is actually a Hopf subalgebra. It also stable by the tridendriform operations, and by the codendriform half-coproducts. Hence,

THEOREM 2.5. **WQSym** *is a sub-bidendriform bialgebra of* **PQSym***. *More precisely, the product rules are*

$$\mathbf{M}_{w'} \ll \mathbf{M}_{w''} = \sum_{w=u.v\in w'*_W w'', |u|=|w'|; \max(v)<\max(u)} \mathbf{M}_w, \tag{36}$$

$$\mathbf{M}_{w'} \gg \mathbf{M}_{w''} = \sum_{w=u.v\in w'*_W w'', |u|=|w'|; \max(v)\geq\max(u)} \mathbf{M}_w, \tag{37}$$

$$\Delta_{\ll}\mathbf{M}_w = \sum_{w\in u \Cup_W v; \mathrm{last}(w)\leq|u|} \mathbf{M}_u \otimes \mathbf{M}_v, \tag{38}$$

$$\Delta_{\gg}\mathbf{M}_w = \sum_{w\in u \Cup_W v; \mathrm{last}(w)>|u|} \mathbf{M}_u \otimes \mathbf{M}_v. \tag{39}$$

where $|u| \geq 1$ *and* $|v| \geq 1$, *and* $\mathrm{last}(w)$ *means the last letter of* w. *As a consequence,* **WQSym** *is free, cofree, self-dual, and its primitive Lie algebra is free.*

2.4. Duality: embedding WQSym* into PQSym. Recall from [**17**] that **PQSym** is the algebra with basis $(\mathbf{F_a})$, the product being given by the shifted shuffle of parking functions, and that $(\mathbf{G_a})$ is the dual basis in **PQSym***.

For a packed word u over the integers, let us define its *maximal unpacking* $\mathrm{mup}(u)$ as the greatest parking function $\mathbf{b}$ for the lexicographic order such that $\mathrm{pack}(\mathbf{b}) = u$. For example, $\mathrm{mup}(321412451) = 641714791$.

Since the basis $(\mathbf{M}_u)$ of **WQSym** can be expressed as the sum of $\mathbf{G_a}$ with a given packed word, the dual basis of $(\mathbf{M}_u)$ in **WQSym*** can be identified with equivalence classes of $(\mathbf{F_a})$ under the relation $\mathbf{F_a} = \mathbf{F_{a'}}$ iff $\mathrm{pack}(\mathbf{a}) = \mathrm{pack}(\mathbf{a}')$. Since the shifted shuffle of two maximally unpacked parking functions contains only maximally unpacked parking functions, the dual algebra **WQSym*** is in fact a subalgebra of **PQSym**. Finally, since, if $\mathbf{a}$ is maximally unpacked then only maximally unpacked parking functions appear in the coproduct $\Delta\mathbf{F_a}$, one has

THEOREM 2.6. **WQSym*** *is a Hopf subalgebra of* **PQSym**. *Its basis element* $\mathbf{M}^*_u$ *can be identified with* $\mathbf{F_b}$ *where* $\mathbf{b} = \mathrm{mup}(u)$.

So we have

$$\mathbf{F_{b'}}\mathbf{F_{b'}} := \sum_{\mathbf{b}\in\mathbf{b'}\Cup\mathbf{b''}} \mathbf{F_b}\,, \qquad \Delta\mathbf{F_b} = \sum_{u\cdot v=\mathbf{b}} \mathbf{F}_{\mathrm{Park}(u)} \otimes \mathbf{F}_{\mathrm{Park}(v)}\,, \tag{40}$$

where Park is the parkization algorithm defined in [**19**]. For example,

$$\mathbf{F}_{113}\mathbf{F}_{11} = \mathbf{F}_{11344} + \mathbf{F}_{11434} + \mathbf{F}_{11443} + \mathbf{F}_{14134} + \mathbf{F}_{14143} + \mathbf{F}_{14413} + \mathbf{F}_{41134} + \mathbf{F}_{41143} + \mathbf{F}_{41413} + \mathbf{F}_{44113}. \tag{41}$$

$$\Delta\mathbf{F}_{531613} = 1\otimes\mathbf{F}_{531613}+\mathbf{F}_1\otimes\mathbf{F}_{31513}+\mathbf{F}_{21}\otimes\mathbf{F}_{1413}+\mathbf{F}_{321}\otimes\mathbf{F}_{312}+\mathbf{F}_{3214}\otimes\mathbf{F}_{12}+\mathbf{F}_{43151}\otimes\mathbf{F}_1\mathbf{F}_{531613}\otimes 1. \tag{42}$$

2.5. The Solomon-Tits algebra. The above realization of $\mathbf{WQSym}^*$ in $\mathbf{PQSym}$ is stable under the internal product of $\mathbf{PQSym}$ defined in [**18**]. Indeed, by definition of the internal product, if $\mathbf{b}'$ and $\mathbf{b}''$ are maximally unpacked, and $\mathbf{F_b} = \mathbf{F_{b'}} * \mathbf{F_{b''}}$, then $\mathbf{b}$ is also maximally unpacked.

Moreover, if one writes $\mathbf{b}' = \{s'_1, \ldots, s'_k\}$ and $\mathbf{b}'' = \{s''_1, \ldots, s''_l\}$ as ordered set partitions, then the parkized word $\mathbf{b} = \mathrm{Park}(\mathbf{b}', \mathbf{b}'')$ corresponds to the ordered set partition obtained from

$$\{s'_1 \cap s''_1, s'_1 \cap s''_2, \ldots, s'_1 \cap s''_l, s'_2 \cap s''_1, \ldots, s'_k \cap s''_l\}. \tag{43}$$

This formula was rediscovered in [**2**] and Bergeron and Zabrocki recognized the Solomon-Tits algebra, in the version given by Bidigare [**3**], in terms of the face semigroup of the braid arrangement of hyperplanes. So,

THEOREM 2.7. *$(\mathbf{WQSym}^*, *)$ is isomorphic to the Solomon-Tits algebra.*

In particular, the product of the Solomon-Tits algebra is dual to the coproduct $\delta\mathbf{G}(A) = \mathbf{G}(A'A'')$.

2.6. The pseudo-permutohedron. We shall now make use of the lattice of pseudo-permutations, a combinatorial structure defined in [**12**] and rediscovered in [**21**]. *Pseudo-permutations* are nothing but ordered set partitions. However, regarding them as generalized permutations helps uncovering their lattice structure. Indeed, let us say that if i is in a block strictly to the right of j with $i < j$ then we have a full inversion (i, j), and that if i is in the same block as j, then we have a *half* inversion $\frac{1}{2}(i, j)$. The total number of inversions is the sum of these numbers. For example, the table of inversions of 45|13|267|8 is

$$\left\{\frac{1}{2}(1,3),\ (1,4),\ (1,5),\ (2,3),\ (2,4),\ (2,5),\ \frac{1}{2}(2,6),\ \frac{1}{2}(2,7),\ (3,4),\ (3,5),\ \frac{1}{2}(4,5),\ \frac{1}{2}(6,7)\right\}, \tag{44}$$

and it has 9.5 inversions.

One can now define a partial order $\preceq$ on pseudo-permutations by setting $p_1 \preceq p_2$ if the value of the inversion (i, j) in the table of inversions of p_1 is smaller than or equal to its value in the table of inversions of p_2, for all (i, j). This partial order is a lattice [**12**]. In terms of packed words, the covering relation reads as follows. The successors of a packed word u are the packed words v such that

- if all the $i-1$ are to the left of all the i in u then u has as successor the element where all letters j greater than or equal to i are replaced by $j-1$.
- if there are k letters i in u, then one can choose an integer j in the interval $[1, k-1]$ and change the j righmost letters i into $i+1$ and the letters l greater than i into $l+1$.

For example, $w = 44253313$ has five successors,

$$33242212,\ 44243313,\ 55264313,\ 55264413,\ 54263313. \tag{45}$$

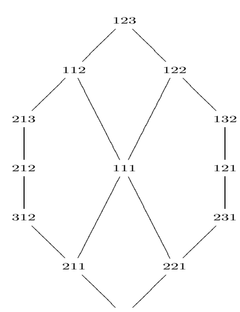


FIGURE 1. The pseudo-permutohedron of degree 3.

THEOREM 2.8 ([**21**]). *Let u and v be two packed words. Then $\mathbf{M}_u\mathbf{M}_v$ is an interval of the pseudo-permutohedron lattice. The minimum of the interval is given by $u{\cdot}v[\max(u)]$ and its maximum by $u[\max(v)]{\cdot}v$.*

For example,

$$\mathbf{M}_{13214}\mathbf{M}_{212} = \sum_{u\in[13214656,35436212]} \mathbf{M}_u. \tag{46}$$

2.7. Other bases of WQSym and WQSym*. Since there is a lattice structure on packed words and since we know that the product $\mathbf{M}_u\mathbf{M}_v$ is an interval of this lattice, we can define several interesting bases, depending on the way we use the lattice.

As in the case of the permutohedron, one can take sums of $\mathbf{M}_u$, over all the elements upper or lower than u in the lattice, or restricted to elements belonging to the same "class" as u (see [**5, 1**] for examples of such bases). In the case of the permutohedron, the classes are the descent classes of permutations. In our case, the classes are the intervals of the pseudo-permutohedron composed of words with the same standardization.

Summing over all elements upper (or lower) than a word u naturally yields multiplicative bases on **WQSym**. Summing over all elements upper (or lower) than u inside its standardization class leads to analogs of the usual bases of *QSym*.

2.7.1. *Multiplicative bases.* Let

$$\mathcal{S}_u := \sum_{v\preceq u}\mathbf{M}_v \quad \text{and} \quad \mathcal{E}_u := \sum_{u\preceq v}\mathbf{M}_v. \tag{47}$$

For example,

$$\mathcal{S}_{212} = \mathbf{M}_{212} + \mathbf{M}_{213} + \mathbf{M}_{112} + \mathbf{M}_{123}. \tag{48}$$

$$\mathcal{E}_{212} = \mathbf{M}_{212} + \mathbf{M}_{312} + \mathbf{M}_{211} + \mathbf{M}_{321}. \tag{49}$$

$$\mathcal{S}_{1122} = \mathbf{M}_{1122} + \mathbf{M}_{1123} + \mathbf{M}_{1233} + \mathbf{M}_{1234}. \tag{50}$$

Since both $\mathcal{S}$ and $\mathcal{E}$ are triangular over the basis $\mathbf{M}_u$ of **WQSym**, we know that these are bases of **WQSym**.

THEOREM 2.9. *The sets $(\mathcal{S}_u)$ and $(\mathcal{E}_u)$ where u runs over packed words are bases of* **WQSym**. *Moreover, their product is given by*

$$\mathcal{S}_{u'}\mathcal{S}_{u''} = \mathcal{S}_{u'[\max(u'')]\cdot u''}. \tag{51}$$

$$\mathcal{E}_{u'}\mathcal{E}_{u''} = \mathcal{E}_{u'\cdot u''[\max(u')]}. \tag{52}$$

For example,

$$\mathcal{S}_{1122}\mathcal{S}_{132} = \mathcal{S}_{4455132}. \tag{53}$$

$$\mathcal{E}_{1122}\mathcal{E}_{132} = \mathcal{E}_{1122354}. \tag{54}$$

2.7.2. *Quasi-ribbon basis of* **WQSym**. Let us first mention that a basis of **WQSym** has been defined in [**2**] by summing over intervals restricted to standardization classes of packed words.

We will now consider similar sums but taken the other way round, in order to build the analogs of **WQSym** of Gessel's fundamental basis F_I of *QSym*. Indeed, as already mentioned, the $\mathbf{M}_u$ are mapped to the M_I of *QSym* under the abelianization $\mathbb{K}\langle A\rangle \to \mathbb{K}[X]$ of **WQSym**. Since the pair of dual bases (F_I, R_I) of (*QSym*, **Sym**) is of fundamental importance, it is natural to ask whether one can find an analogous pair for (**WQSym**, **WQSym***). To avoid confusion in the notations, we will denote the analog of F_I by Φ_u instead of $\mathbf{F}_u$ since this notation is already used in the dual algebra **WQSym**$^* \subset$ **PQSym**, with a different meaning. The analog of R basis in **WQSym*** will still be denoted by R. The representation of packed words by segmented permutations is more suited for the next statements since one easily checks that two words u and v having the same standardized word satisfy $v \preceq u$ iff v is obtained as a segmented permutation from the segmented permutation of u by inserting any number of bars. Let

$$\Phi_\sigma := \sum_{\sigma'}\mathbf{M}_{\sigma'} \tag{55}$$

where σ' runs ver the set of segmented permutations obtained from σ by inserting any number of bars. For example,

$$\Phi_{14|6|23|5} = \mathbf{M}_{14|6|23|5} + \mathbf{M}_{14|6|2|3|5} + \mathbf{M}_{1|4|6|23|5} + \mathbf{M}_{1|4|6|2|3|5}. \tag{56}$$

Since (Φ_u) is triangular over $(\mathbf{M}_u)$, it is a basis of **WQSym**. By construction, it satisfies a product formula similar to that of Gessel's basis F_I of *QSym* (whence the choice of notation). To state it, we need an analogue of the shifted shuffle, defined on the special class of segmented permutations encoding set compositions.

The *shifted shuffle* $\alpha \Cup \beta$ of two such segmented permutations is obtained from the usual shifted shuffle $\sigma \Cup \tau$ of the underlying permutations σ and τ by inserting bars

- between each pairs of letters coming from the same word if they were separated by a bar in this word,
- after each element of β followed by an element of α.

For example,

$$2|1 \Cup 12 = 2|134 + 23|14 + 234|1 + 3|2|14 + 3|24|1 + 34|2|1. \tag{57}$$

THEOREM 2.10. *The product and coproduct in the basis Φ are given by*

$$\Phi_{\sigma'}\Phi_{\sigma''} = \sum_{\sigma\in\sigma'\Cup\sigma''} \Phi_\sigma. \tag{58}$$

$$\Delta\Phi_\sigma = \sum_{\sigma'|\sigma''=\sigma \text{ or } \sigma'\cdot\sigma''=\sigma} \Phi_{\mathrm{Std}(\sigma')} \otimes \Phi_{\mathrm{Std}(\sigma'')}. \tag{59}$$

For example, we have

$$\Phi_1\Phi_{13|2} = \Phi_{124|3} + \Phi_{2|14|3} + \Phi_{24|13} + \Phi_{24|3|1}. \tag{60}$$

$$\Delta\Phi_{35|14|2} = 1 \otimes \Phi_{35|14|2} + \Phi_1 \otimes \Phi_{4|13|2} + \Phi_{12} \otimes \Phi_{13|2} + \Phi_{23|1} \otimes \Phi_{2|1} + \Phi_{24|13} \otimes \Phi_1 + \Phi_{35|14|2} \otimes 1. \tag{61}$$

Note that under abelianization, $\chi(\Phi_u) = F_I$ where I is the evaluation of u.

2.7.3. *Ribbon basis of* **WQSym***. Let us now consider the dual basis of Φ. We have seen that it should be regarded as an analog of the ribbon basis of **Sym**. By duality, one can state:

THEOREM 2.11. *Let R_σ be the dual basis of Φ_σ. Then the product and coproduct in this basis are given by*

$$R_{\sigma'}R_{\sigma''} = \sum_{\sigma=\tau|\nu \text{ or } \sigma=\tau\nu;\mathrm{Std}(\tau)=\sigma',\mathrm{Std}(\nu)=\sigma''} R_\sigma. \tag{62}$$

$$\Delta R_\sigma = \sum_{\sigma'.\sigma''=\sigma} R_{\mathrm{Std}(\sigma')} \otimes R_{\mathrm{Std}(\sigma'')}. \tag{63}$$

Note that there are more elements coming from $\tau|\nu$ than from $\tau\nu$ since the permutation σ has to be *increasing* between two bars.

For example,

$$R_{21}R_1 = R_{212} + R_{221} + R_{213} + R_{231} + R_{321}. \tag{64}$$

3. Hopf algebras based on Schröder sets

In Section 2.2, we recalled that the little Schröder numbers build up the Hilbert series of the free dendriform trialgebra on one generator $\mathfrak{TD}$. We will see that our relization of $\mathfrak{TD}$ endows it with a natural structure of bidendriform bialgebra. In particular, this will prove that there is a natural self-dual Hopf structure on $\mathfrak{TD}$. But there are other ways to arrive at the little Schröder numbers from the other Hopf algebras **WQSym** and **PQSym**. Indeed, the number of classes of packed words of size n under the sylvester congruence is s_n, and the number of classes of parking functions of size n under the hypoplactic congruence is also s_n. The hypoplactic quotient of **PQSym*** has been studied in [**19**]. It is not isomorphic to $\mathfrak{TD}$ nor to the sylvester quotient of **WQSym** since it is a non self-dual Hopf algebra whereas the last two are self-dual, and furthemore isomorphic as bidendriform bialgebras and as dendriform trialgebras.

3.1. The free dendriform trialgebra again. Recall that we realized the free dendriform trialgebra in Section 2.2 as the subtrialgebra of **WQSym** generated by $\mathbf{M}_1$, the sum of all letters. It is immediate that $\mathfrak{TD}$ is stable by the codendriform half-coproducts of **WQSym***. Hence,

THEOREM 3.1. *$\mathfrak{TD}$ is a sub-bidendriform bialgebra, and hence a Hopf subalgebra of* **WQSym****. In particular, $\mathfrak{TD}$ is free, self-dual and its primitive Lie algebra is free.*

3.2. Lattice structure on plane trees. Given a plane tree T, define its *canonical word* as the maximal packed word w in the pseudo-permutohedron such that $\mathcal{T}(w) = T$.

For example, the canonical words up to $n = 3$ are

$$\{1\}, \qquad \{11, 12, 21\}, \qquad \{111,\ 112,\ 211,\ 122,\ 212,\ 221,\ 123,\ 213,\ 231,\ 312,\ 321\} \tag{65}$$

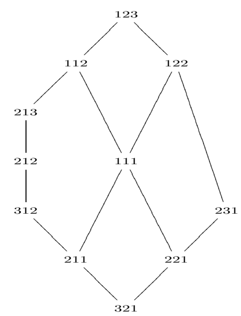


FIGURE 2. The lattice of plane trees represented by their canonical words for $n = 3$.

Define the *second canonical word* of each tree T as the minimal packed word w in the pseudo-permutohedron such that $\mathcal{T}(w) = T$.

A packed word $u = u_1 \cdots u_n$ is said to *avoid the pattern* $w = w_1 \cdots w_k$ if there is no sequence $1 \le i_1 < \cdots < i_k \le n$ such that $u' = u_{i_1} \cdots u_{i_k}$ has same inversions and same half-inversions as w.

For example, 41352312 avoids the patterns 111 and 1122, but not 2311 since 3522 has the same (half)-inversions.

THEOREM 3.2. *The canonical words of trees are the packed words avoiding the patterns 121 and 132. The second canonical words of trees are the packed words avoiding the patterns 121 and 231.*

Set $u \sim_T v$ iff $\mathcal{T}(u) = \mathcal{T}(v)$. We now define two orders $\sim_T$-classes of packed words

1. A class S is smaller than a class S' if the canonical word of S is smaller than the canonical word of S' in the pseudo-permutohedron.
2. A class S is smaller than a class S' if there is a pair (w, w') in $S \times S'$ such that w is smaller than w' in the pseudo-permutohedron.

THEOREM 3.3. *These two orders coincide and are also identical with the one defined in* [**21**]. *Moreover, the restriction of the pseudo-permutohedron to the canonical words of trees is a lattice.*

3.3. Some bases of $\mathfrak{TD}$.

3.3.1. *The basis $\mathcal{M}_T$.* Let us start with the already defined basis $\mathcal{M}_T$. First note that $\mathcal{M}_T$ expressed as a sum of $\mathbf{M}_u$ in **WQSym** is an interval of the pseudo-permutohedron. From the above description of the lattice, we obtain easily:

THEOREM 3.4 ([**21**]). *The product $\mathcal{M}_{T'}\mathcal{M}_{T''}$ is an interval of the lattice of plane trees. On trees, the minimum $T' \wedge T''$ is obtained by gluing the root of T'' at the end of the leftmost branch of T', whereas the maximum $T' \vee T''$ is obtained by gluing the root of T' at the end of the rightmost branch of T''.*

On the canonical words w' and w'', the minimum is the canonical word associated with $w' \cdot w''[\max(w')]$ and the maximum is $w'[\max(w'')] \cdot w''$.

3.3.2. *Complete and elementary bases of $\mathfrak{TD}$.* We can also build two multiplicative bases as in **WQSym**.

THEOREM 3.5. *The set $(\mathcal{S}_w)$ (resp. $(\mathcal{E}_w)$) where w runs over canonical (resp. second canonical) words are multiplicative bases of $\mathfrak{TD}$.*

3.4. Internal product on $\mathfrak{TD}$. If one defines $\mathfrak{TD}$ as the Hopf subalgebra of **WQSym** defined by

$$\mathcal{M}_T = \sum_{\mathcal{T}(u)=T} \mathbf{M}_u \,, \tag{66}$$

then $\mathfrak{TD}^*$ is the quotient of $\mathbf{WQSym}^*$ by the relation $\mathbf{F}_u \equiv \mathbf{F}_v$ iff $\mathcal{T}(u) = \mathcal{T}(v)$. We denote by S_T the dual basis of $\mathcal{M}_T$.

THEOREM 3.6. *The internal product of* $\mathbf{WQSym}_n^*$ *induces an internal product on the homogeneous components* $\mathfrak{TD}_n^*$ *of the dual algebra. More precisely, one has*

$$S_{T'} * S_{T''} = S_T, \tag{67}$$

where T *is the tree obtained by applying* $\mathcal{T}$ *to the biword of the canonical words of the trees* T' *and* T''.

For example, representing trees as their canonical words, one has

$$S_{221} * S_{122} = S_{231}; \qquad S_{221} * S_{321} = S_{321}; \tag{68}$$

$$S_{453223515} * S_{433442214} = S_{674223518}. \tag{69}$$

3.5. Sylvester quotient of WQSym. One can check by direct calculation that the sylvester quotient [**10**] of **WQSym** is also stable by the tridendriform operations, and by the codendriform half-coproducts since the elements of a sylvester class have the same last letter. Hence,

THEOREM 3.7. *The sylvester quotient of* **WQSym** *is a dendriform trialgebra, a bidendriform bialgebra, and hence a Hopf algebra. It is isomorphic to* $\mathfrak{TD}$ *as a dendriform trialgebra, as a bidendriform bialgebra and as a Hopf algebra.*

4. A Hopf algebra of segmented compositions

In [**19**], we have built a Hopf subalgebra $\mathbf{SCQSym}^*$ of the hypoplactic quotient $\mathbf{SQSym}^*$ of $\mathbf{PQSym}^*$, whose Hilbert series is given by

$$1 + \sum_{n\geq 1} 3^{n-1} t^n. \tag{70}$$

This Hopf algebra is not self-dual, but admits lifts of Gessel's fundamental basis F_I of $QSym$ and its dual basis. Since the elements of $\mathbf{SCQSym}^*$ are obtained by summing up hypoplactic classes having the same packed word, thanks to the following diagram, it is obvious that $\mathbf{SCQSym}^*$ is also the quotient of **WQSym** by the hypoplactic congruence.

$$\begin{array}{ccc} \mathbf{PQSym}^* & \xrightarrow{hypo} & \mathbf{SQSym}^* \\ {\scriptstyle(pack)}\uparrow & & \uparrow{\scriptstyle(pack)} \\ \mathbf{WQSym} & \xrightarrow{hypo} & \mathbf{SCQSym}^* \end{array} \tag{71}$$

4.1. Segmented compositions. Define a *segmented composition* as a finite sequence of integers, separated by vertical bars or commas, *e.g.*, $(2,1\,|\,2\,|\,1,2)$.

The number of segmented compositions having the same underlying composition is obviously 2^{l-1} where l is the length of the composition, so that the total number of segmented compositions of sum n is 3^{n-1}. There is a natural bijection between segmented compositions of n and sequences of length $n-1$ over three symbols $<, =, >$: start with a segmented composition $\mathbf{I}$. If the i-th position is not a descent of the underlying ribbon diagram, write $<$; otherwise, if i is followed by a comma, write $=$; if i is followed by a bar, write $>$.

Now, with each word w of length n, associate a segmented composition $S(w) = s_1 \cdots s_{n-1}$ where s_i is the correct comparison sign between w_i and w_{i+1}. For example, given $w = 1615116244543$, one gets the sequence (and the segmented composition):

$$<><>=<><=<>>\Longleftrightarrow (2|2|1,2|2,2|1|1). \tag{72}$$

4.2. A subalgebra of $\mathfrak{TD}$. Given a segmented composition $\mathbf{I}$, define

$$M_{\mathbf{I}} = \sum_{S(T)=\mathbf{I}} \mathcal{M}_T = \sum_{S(u)=\mathbf{I}} \mathbf{M}_u\,. \tag{73}$$

For example,

$$M_{12|1} = \mathcal{M}_{2231} \qquad M_{1|3} = \mathcal{M}_{2123} + \mathcal{M}_{2134} + \mathcal{M}_{3123} + \mathcal{M}_{3124} + \mathcal{M}_{4123}. \tag{74}$$

THEOREM 4.1. *The $M_{\mathbf{I}}$ generate a subalgebra $\mathfrak{TC}$ of $\mathfrak{TD}$. Their product is given by*

$$M_{\mathbf{I}'}M_{\mathbf{I}''} = M_{\mathbf{I}'\triangleright\mathbf{I}''} + M_{\mathbf{I}',\mathbf{I}''} + M_{\mathbf{I}'|\mathbf{I}''}. \tag{75}$$

where $\mathbf{I}' \triangleright \mathbf{I}''$ is obtained by gluing the last part of $\mathbf{I}'$ and the first part of $\mathbf{I}''$, so that $\mathfrak{TC}$ is the free cubical trialgebra on one generator [**15**].

For example,

$$M_{1|21}M_{31} = M_{1|241} + M_{1|2131} + M_{1|21|31}. \tag{76}$$

4.3. A lattice structure on segmented compositions. Given a segmented composition $\mathbf{I}$, define its *canonical word* as the maximal packed word w in the pseudo-permutohedron such that $S(w) = \mathbf{I}$.

For example, the canonical words up to $n = 3$ are

$$\{1\}, \qquad \{11, 12, 21\}, \qquad \{111,\ 112,\ 211,\ 122,\ 221,\ 123,\ 231,\ 312,\ 321\} \tag{77}$$

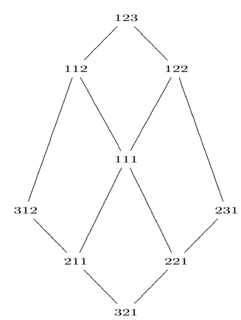


FIGURE 3. The lattice of segmented compositions represented by their canonical words at $n = 3$.

Define the *second canonical word* of a segmented composition $\mathbf{I}$ as the minimal packed word w in the pseudo-permutohedron such that $S(w) = \mathbf{I}$.

THEOREM 4.2. *The canonical words of segmented compositions are the packed words avoiding the patterns 121, 132, 212, and 213. The second canonical words of segmented compositions are the packed words avoiding the patterns 121, 231, 212, and 312.*

Let $u \sim_S v$ iff $S(u) = S(v)$. We define two orders on $\sim_S$-equivalence classes of words.

1. A class S is smaller than a class S' if the canonical word of S is smaller than the canonical word of S' in the pseudo-permutohedron.
2. A class S is smaller than a class S' if there exists two elements (w, w') in $S \times S'$ such that w is smaller than w' in the pseudo-permutohedron.

PROPOSITION 4.3. *The two orders coincide. Moreover, the restriction of the pseudo-permutohedron to the canonical segmented words is a lattice.*

4.4. Multiplicative bases. We can build two multiplicative bases, as in **WQSym**. They are particularly simple:

THEOREM 4.4. *The set $(\mathcal{S}_w)$ where w runs into the set of canonical segmented words is a basis of $\mathfrak{TC}$. The set $(\mathcal{E}_w)$ where w runs into the set of second canonical segmented words is a basis of $\mathfrak{TC}$.*

4.5. Internal product on $\mathfrak{TC}$. If one defines $\mathfrak{TC}$ as the Hopf subalgebra of **WQSym** as in Equation (73), then $\mathfrak{TC}^*$ is the quotient of **WQSym*** by the relation $\mathbf{F}_u \equiv \mathbf{F}_v$ iff $S(u) = S(v)$. We denote by $S_{\mathbf{I}}$ the dual basis of $M_{\mathbf{I}}$.

THEOREM 4.5. *The internal product of* **WQSym*** *induces an internal product on the homogeneous components $\mathfrak{TC}_n^*$ of $\mathfrak{TC}^*$. More precisely, one has*

$$S_{\mathbf{I}'} * S_{\mathbf{I}''} = S_{\mathbf{I}}, \tag{78}$$

where $\mathbf{I}$ is the segmented composition obtained by applying S to the biword of the canonical words of the segmented compositions $\mathbf{I}'$ and $\mathbf{I}''$.

References

[1] M. Aguiar and F. Sottile, *Structure of the Loday-Ronco Hopf algebra of trees*, Adv. Math., **191** (2005), 225–275.

[2] N. Bergeron and M. Zabrocki, *The Hopf algebras of symmetric functions and quasisymmetric functions in non-commutative variables are free and cofree*, preprint math.CO/0509265.

[3] T. P. Bidigare, *Hyperplane arrangements, face algebras and their associated Markov chains*, Ph. D. Thesis, University of Michigan, 1997.

[4] F. Chapoton, *Algèbres de Hopf des permutahèdres, associahèdres et hypercubes*, Advances in Math. **150** (2000), 264–275.

[5] G. Duchamp, F. Hivert, and J.-Y. Thibon, *Noncommutative symmetric functions VI: free quasi-symmetric functions and related algebras*, Internat. J. Alg. Comput. **12** (2002), 671–717.

[6] L. Foissy, *Bidendriform bialgebras, trees, and free quasi-symmetric functions*, math.RA/0505207.

[7] I.M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, and J.-Y. Thibon, *Noncommutative symmetric functions*, Adv. in Math. **112** (1995), 218–348.

[8] F. Hivert, *Combinatoire des fonctions quasi-symétriques*, Thèse de Doctorat, Marne-La-Vallée, 1999.

[9] F. Hivert, *Hecke Algebras, Difference Operators, and Quasi-Symmetric Functions*, Advances in Math. **155** (2000), 181–238.

[10] F. Hivert, J.-C. Novelli, and J.-Y. Thibon, *The algebra of binary search trees*, Theoretical Computer Science **339** (2005), 129–165.

[11] F. Hivert and N. Thiéry, *MuPAD-Combinat, an open-source package for research in algebraic combinatorics*, Sém. Lothar. Combin. **51** (2004), 70p. (electronic).

[12] D. Krob, M. Latapy, J.-C. Novelli, H. D. Phan and S. Schwer, *Pseudo-Permutations I: First Combinatorial and Lattice Properties*, FPSAC'01, H. Barcelo, ed., 2000.

[13] J.-L. Loday, *Dialgebras*, "Dialgebras and related operads", Lecture Notes in Math. **1763**, Springer, Berlin, (2001), 7–66.

[14] J.-L. Loday and M. O. Ronco, *Hopf algebra of the planar binary trees*, Adv. Math. **139** (1998) n. 2, 293–309.

[15] J.-L. Loday and M. O. Ronco, *Trialgebras and families of polytopes*, Contemporary Mathematics **346** (2004).

[16] T. V. Narayana, *Lattice Path Combinatorics with Statistical Applications*, Univ. Toronto Press (1979), 100–101.

[17] J.-C. Novelli and J.-Y. Thibon, *A Hopf algebra of parking functions*, Proc. FPSAC'04, Vancouver.

[18] J.-C. Novelli and J.-Y. Thibon, *Parking Functions and Descent Algebras*, Annals of Combinatorics, to appear.

[19] J.-C. Novelli and J.-Y. Thibon, *Hopf algebras and dendriform structures arising from parking functions*, preprint, math.CO/0511200.

[20] J.-C. Novelli and J.-Y. Thibon, *Construction of dendriform trialgebras*, math.CO/0510218

[21] P. Palacios and M.O. Ronco, *Weak Bruhat order on the set of faces of the permutahedra*, preprint, math.CO/0404352.

[22] N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://www.research.att.com/~njas/sequences/`

Institut Gaspard Monge, Université de Marne-la-Vallée, 5 Boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, FRANCE

E-mail address, Jean-Christophe Novelli: `novelli@univ-mlv.fr`

E-mail address, Jean-Yves Thibon: `jyt@univ-mlv.fr`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The combinatorics of frieze patterns and Markoff numbers

James Propp

ABSTRACT. This article, based on joint work with Gabriel Carroll, Andy Itsara, Ian Le, Gregg Musiker, Gregory Price, Dylan Thurston, and Rui Viana, presents a combinatorial model based on perfect matchings that explains the symmetries of the numerical arrays that Conway and Coxeter dubbed frieze patterns. This matchings model is a combinatorial interpretation of Fomin and Zelevinsky's cluster algebras of type A. One can derive from the matchings model an enumerative meaning for the Markoff numbers, and prove that the associated Laurent polynomials have positive coefficients as was conjectured (much more generally) by Fomin and Zelevinsky. Most of this research was conducted under the auspices of REACH (Research Experiences in Algebraic Combinatorics at Harvard).

RÉSUMÉ. Cet article, basé sur un travail conjoint avec Gabriel Carroll, Andy Itsara, Ian Le, Gregg Musiker, Gregory Price, Dylan Thurston, et Rui Viana presente un modèle combinatoire expliquant les symétries dans les tableaux numérique appelés motifs frieze par Conway et Coxeter. Ce modèle, basé sur les couplages parfaits, donne une interprétation combinatoire des algèbre de cluster de type A de Fomin et Zelevinksy. Ce modèle permet de fournir une interprétation énumérative des nombres Markoff, et on peut démontrer que les polynômes de Laurent associés ont des coefficients positifs, ce qui avait été conjecturé (dans un cadre plus général) par Fomin et Zelevinsky. Cette recherche s'est déroulée dans le cadre du programme REACH (Research Experiences in Algebraic Combinatorics at Harvard).

1. Introduction

A *Laurent polynomial* in the variables $x, y, \ldots$ is a polynomial in the variables $x, x^{-1}, y, y^{-1}, \ldots$. Thus the function $f(x) = (x^2+1)/x = x + x^{-1}$ is a Laurent polynomial, but the composition $f(f(x)) = (x^4+3x^2+1)/x(x^2+1)$ is not. This shows that the set of Laurent polynomials in a single variable is not closed under composition. This failure of closure also holds in the multivariate setting; for instance, if $f(x,y)$, $g(x,y)$ and $h(x,y)$ are Laurent polynomials in x and y, then we would not expect to find that $f(g(x,y), h(x,y))$ is a Laurent polynomial as well. Nonetheless, it has been discovered that, in broad class of instances (embraced as yet by no general rule), "fortuitous" cancellations occur that cause Laurentness to be preserved. This is the "Laurent phenomenon" discussed by Fomin and Zelevinsky [**13**].

Furthermore, in many situations where the Laurent phenomenon holds, there is a certain positivity phenomenon at work as well, and all the coefficients of the Laurent polynomials turn out to be positive. In these cases, the functions being composed are Laurent polynomials with positive coefficients; that is, they are expressions involving only addition, multiplication, and division. It should be noted that subtraction-free expressions do not have all the closure properties one might hope for, as the example $(x^3+y^3)/(x+y)$ illustrates: although the expression is subtraction-free, its reduced form $x^2 - xy + y^2$ is not.

Fomin and Zelevinsky have shown that a large part of the Laurentness phenomenon fits in with their general theory of cluster algebras. In this article I will discuss one important special case of the Laurentness-plus-positivity phenomenon, namely the case associated with cluster algebras of type A, discussed in detail in [**14**]. The purely combinatorial approach taken in sections 2 and 3 of my article obscures the links with deeper issues (notably the representation-theoretic questions that motivated the invention of cluster algebras), but it provides the quickest and most self-contained way to prove the Laurentness-plus-positivity assertion in this case (Theorem 3.1). The frieze patterns of Conway and Coxeter, and their link with triangulations of polygons, will play a fundamental role, as will

---

2000 *Mathematics Subject Classification.* Primary 05C70; Secondary 05E15.

Key words and phrases. Frieze patterns, cluster algebras, Markoff numbers, perfect matchings.

This work was supported by funds from the National Science Foundation, and the National Security Agency.

perfect matchings of graphs derived from these triangulations. (For a different, more algebraic way of thinking about frieze patterns, see [**3**].)

In sections 4 and 5 of this article, two variations on the theme of frieze patterns are considered. One is the tropical analogue, which has bearing on graph-metrics in trees. The other variant is based on a recurrence that looks very similar to the frieze relation; the variant recurrence appears to give rise to tables of positive integers possessing the same glide-reflection symmetry as frieze patterns, but positivity, integrality, and symmetry are currently still unproved.

In section 6, the constructions of sections 2 and 3 are specialized to a case in which the triangulated polygons come from pairs of mutually visible points in a dissection of the plane into equilateral triangles. In this case, counting the matchings of the derived graphs gives us an enumerative interpretation of Markoff numbers (numbers satisfying the ternary cubic $x^2+y^2+z^2=3xyz$). This yields a combinatorial proof of a Laurentness assertion proved by Fomin and Zelevinsky in [**13**] (namely a special case of their Theorem 1.10) that falls outside of the framework of cluster algebras in the strict sense. Fomin and Zelevinsky proved Theorem 1.10 by use of their versatile "Caterpillar Lemma", but this proof did not settle the issue of positivity. The combinatorial approach adopted here shows that all of the Laurent polynomials that occur in the three-variable rational-function analogue of the Markoff numbers — the "Markoff polynomials" — are in fact positive (Theorem 6.2).

Section 7 concludes with some problems suggested by the main result of section 6. One can try to generalize the combinatorial picture by taking other dissections of the plane into triangles, or one can try to generalize by considering other Diophantine equations. There may be a general link here, but its nature is still obscure.

2. Triangulations and frieze patterns

A **frieze pattern** [**7**] is an infinite array such as

$$\begin{array}{cccccccccccccc}
\ldots & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & \ldots \\
\ldots & & 1 & & 5 & & \frac{2}{3} & & 3 & & \frac{5}{3} & & 2 & \ldots \\
\ldots & 1 & & 4 & & \frac{7}{3} & & 1 & & 4 & & \frac{7}{3} & & \ldots \\
\ldots & & 3 & & \frac{5}{3} & & 2 & & 1 & & 5 & & \frac{2}{3} & \ldots \\
\ldots & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & \ldots
\end{array}$$

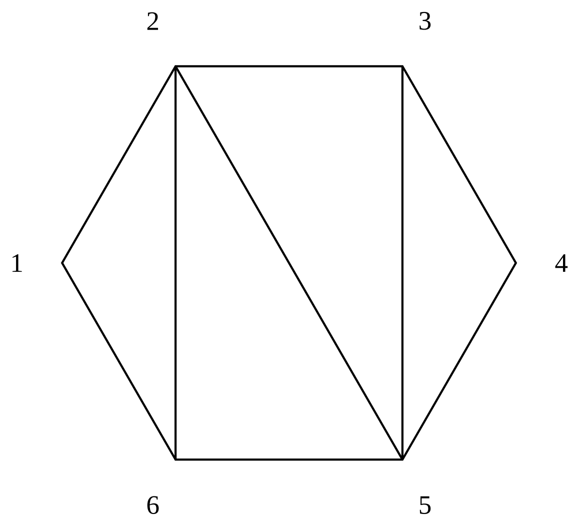
consisting of $n-1$ rows, each periodic with period n, such that all entries in the top and bottom rows are equal to 1 and all entries in the intervening rows satisfy the relation

$$\begin{array}{ccc} & A & \\ B & & C \\ & D & \end{array} \quad : \quad D=(BC-1)/A\ .$$

The rationale for the term "frieze pattern" is that such an array automatically possesses glide-reflection symmetry (as found in some decorative friezes): for $1 \le m \le n-1$, the $n-m$th row is the same as the mth row, shifted. Hence the relation $D=(BC-1)/A$ will be referred to below as the "frieze relation" even though its relation to friezes and their symmetries is not apparent from the algebraic definition.

Frieze patterns arose from Coxeter's study of metric properties of polytopes, and served as useful scaffolding for various sorts of metric data; see e.g. [**9**] (page 160), [**10**], and [**11**]. Typically some of the entries in a frieze pattern are irrational. Conway and Coxeter completely classify the frieze patterns whose entries are positive integers, and show that these frieze patterns constitute a manifestation of the Catalan numbers. Specifically, there is a natural association between positive integer frieze patterns and triangulations of regular polygons with labelled vertices. (In addition to [**7**], see the shorter discussion on pp. 74–76 and 96–97 of [**8**].) Note that for each fixed n, any convex n-gon would serve here just as well as the regular n-gon, since we are only viewing triangulations combinatorially.

From every triangulation T of a regular n-gon with vertices cyclically labelled 1 through n, Conway and Coxeter build an $(n-1)$-rowed frieze pattern determined by the numbers $a_1, a_2, \ldots, a_n$, where a_k is the number of triangles in T incident with vertex k. Specifically: (1) the top row of the array is $\ldots, 1, 1, 1, \ldots$; (2) the second row (offset from the first) is $\ldots, a_1, a_2, \ldots, a_n, a_1, \ldots$ (with period n); and (3) each succeeding row (offset from the one before) is determined by the frieze relation. E.g., the triangulation

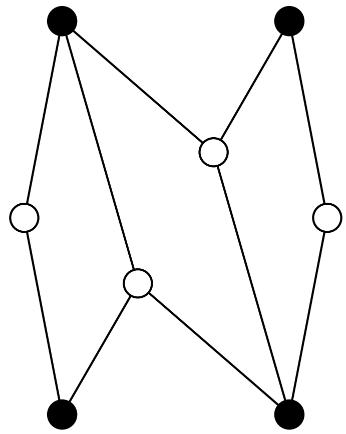


of the 6-gon determines the data $(a_1, \ldots, a_6) = (1, 3, 2, 1, 3, 2)$ and 5-row frieze pattern

$$\begin{array}{cccccccccccccccc}
\ldots & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & \ldots \\
\ldots & & 1 & & 3 & & 2 & & 1 & & 3 & & 2 & & & \ldots \\
\ldots & 1 & & 2 & & 5 & & 1 & & 2 & & 5 & & 1 & & \ldots \\
\ldots & & 1 & & 3 & & 2 & & 1 & & 3 & & 2 & & & \ldots \\
\ldots & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & \ldots
\end{array}$$

Conway and Coxeter show that the frieze relation, applied to the initial rows $\ldots, 1, 1, 1, \ldots$ and $\ldots, a_1, a_2, \ldots, a_n, \ldots$, yields a frieze pattern. Note that implicit in this assertion is the assertion that every entry in rows 1 through $n-3$ is non-zero (so that the recurrence $D = (BC-1)/A$ never involves division by 0). It is not a priori obvious that each of the entries in the array is positive (since the recurrence involves subtraction) or that each of the entries is an integer (since the recurrence involves division). Nor is it immediately clear why for $1 \leq m \leq n-1$, the $n-m$th row of the table given by repeated application of the recurrence should be the same as the mth row, shifted, so that in particular the $n-1$st row, like the first row, consists entirely of 1's.

These and many other properties of frieze patterns are explained by a combinatorial model of frieze patterns discovered by Carroll and Price [**5**] (based on earlier work of Itsara, Le, Musiker, Price, and Viana). Given a triangulation T as above, define a bipartite graph $G = G(T)$ whose n black vertices v correspond to the vertices of T, whose $n-2$ white vertices w correspond to the triangular faces of T, and whose edges correspond to all incidences between vertices and faces in T (that is, v and w are joined by an edge precisely if v is one of the three vertices of the triangle in T associated with w). For $i \neq j$ in the range $1, \ldots, n$, let $G_{i,j}$ be the graph obtained from G by removing black vertices i and j and all edges incident with them, and let $m_{i,j}$ be the number of perfect matchings of $G_{i,j}$ (that is, the number of ways to pair all $n-2$ of the black vertices with the $n-2$ white vertices, so that every vertex is paired to a vertex of the opposite color adjacent to it). For instance, for the triangulation T of the 6-gon defined in the preceding figure, the graph $G_{1,4}$ is



and we put $m_{1,4} = 5$ since this graph has 5 perfect matchings.

THEOREM 2.1 (Gabriel Carroll and Gregory Price [**5**]). *The Conway-Coxeter frieze pattern of a triangulation T is just the array*

$$\begin{array}{cccccccccc}
\cdots & m_{1,2} & & m_{2,3} & & m_{3,4} & & m_{4,5} & \cdots \\
\cdots & & m_{1,3} & & m_{2,4} & & m_{3,5} & & \cdots \\
\cdots & m_{n,3} & & m_{1,4} & & m_{2,5} & & m_{3,6} & \cdots \\
\cdots & & m_{n,4} & & m_{1,5} & & m_{2,6} & & \cdots \\
 & \vdots & & \vdots & & \vdots & & \vdots &
\end{array}$$

where here as hereafter we interpret all subscripts mod n.

Note that this claim makes the glide-reflection symmetry of frieze patterns a trivial consequence of the fact that $G_{i,j} = G_{j,i}$.

PROOF. Here is a sketch of the main steps of the proof:

(1) $m_{i,i+1} = 1$: This holds because there is a tree structure on the set of triangles in T that induces a tree structure on the set of white vertices of G. If we examine the white vertices of G, proceeding from outermost to innermost, we find that we have no freedom in how to match them with black vertices, when we keep in mind that every black vertex must be matched with a white vertex. (In fact, the same reasoning shows that $m_{i,j} = 1$ whenever the triangulation T contains a diagonal connecting vertices i and j.)

(2) $m_{i-1,i+1} = a_i$: The argument is similar, except now we have some freedom in how the ith black vertex is matched: it can be matched with any of the a_i adjacent white vertices.

(3) $m_{i,j}\, m_{i-1,j+1} = m_{i-1,j}\, m_{i,j+1} - 1$: If we move the 1 to the left-hand side, we can use (1) to write the equation in the form

$$m_{i,j}\, m_{i-1,j+1} + m_{i-1,i}\, m_{j,j+1} = m_{i-1,j}\, m_{i,j+1}.$$

This relation is a direct consequence of a lemma due to Eric Kuo (Theorem 2.5 in [**17**]), which I state here for the reader's convenience:

Condensation lemma: If a bipartite planar graph G has 2 more black vertices than white vertices, and the black vertices a, b, c, d lie in cyclic order on some face of G, then

$$m(a,c)m(b,d) = m(a,b)m(c,d) + m(a,d)m(b,c),$$

where $m(x,y)$ denotes the number of perfect matchings of the graph obtained from G by deleting vertices x and y and all incident edges.

(1) and (2) tell us that Carroll and Price's theorem applies to the first two rows of the frieze pattern, and (3) tells us (by induction) that the theorem applies to all subsequent rows. □

It should be mentioned that Conway and Coxeter give an alternative way of describing the entries in frieze patterns, as determinants of tridiagonal matrices. Note that $m_{i-1,i+1} = a_i$ which equals the determinant of the 1-by-1 matrix whose sole entry is a_i, while $m_{i-1,i+2} = a_i a_{i+1} - 1$ which equals the determinant of the 2-by-2 matrix

$$\begin{pmatrix} a_i & 1 \\ 1 & a_{i+1} \end{pmatrix}.$$

One can show by induction using Dodgson's determinant identity (for a statement and a pretty proof of this identity see [**21**]) that $m_{i-1,i+k}$ equals the determinant of the k-by-k matrix with $a_i, \ldots, a_{i+k-1}$ down the diagonal, 1's in the two flanking diagonals, and 0's everywhere else. This is true for any arrays satisfying the frieze relation whose initial row consists of 1's, whether or not it is a frieze pattern. Thus, any numerical array constructed via the frieze relation from initial data consisting of a first row of 1's and a second row of positive integers will be an array of positive integers; entries in subsequent rows will be positive since they are defined by subtraction-free expressions, and they will be integers since they are equal to determinants of integer matrices. (One caveat is in order here: although the table of tridiagonal determinants always satisfies the frieze relation, it may not be possible to compute the table using just the frieze relation, since some of the expressions that arise might be indeterminate fractions of the form $0/0$.) However, for most choices of positive integers $a_1, \ldots, a_n$, the resulting table of positive integers will not be an $(n-1)$-rowed frieze pattern. Indeed, Conway and Coxeter show that every $(n-1)$-rowed frieze pattern whose entries are positive integers arises from a triangulated n-gon in the fashion described above.

3. The sideways construction and its periodicity

Recall that any $(n-1)$-rowed array of real numbers that begins and ends with rows of 1's and satisfies the frieze relation in between qualifies as a frieze pattern.

Note that if the vertices $1, \ldots, n$ of an n-gon lie on a circle and we let $d_{i,j}$ be the distance between points i and j, Ptolemy's theorem on the lengths of the sides and diagonals of an inscriptible quadrilateral gives us the three-term quadratic relation

$$d_{i,j}\, d_{i-1,j+1} + d_{i-1,i}\, d_{j-1,j} = d_{i-1,j}\, d_{i,j+1}$$

(with all subscripts interpreted mod n). Hence the numbers $d_{i,j}$ with $i \neq j$, arranged just as the numbers $m_{i,j}$ were, form an $(n-1)$-rowed array that almost qualifies as a frieze pattern (the array satisfies the frieze relation and has glide-reflection symmetry because $m_{i,j} = m_{j,i}$ for all i, j, but the top and bottom rows do not in general consist of 1's). The nicest case occurs when the n-gon is a regular n-gon of side-length 1; then we get a genuine frieze pattern and each row of the frieze pattern is constant.

Another source of frieze patterns is an old result from spherical geometry: the *pentagramma mirificum* of Napier and Gauss embodies the assertion that the arc-lengths of the sides in a right-angled spherical pentagram can be arranged to form the middle two rows of a four-rowed frieze pattern.

Conway and Coxeter show that frieze patterns are easy to construct if one proceeds not from top to bottom (since one is unlikely to choose numbers $a_1, \ldots, a_n$ in the second row that will yield all 1's in the $(n-1)$st row) but from left to right, starting with a zig-zag of entries connecting the top and bottom rows (where the zig-zag path need not alternate between leftward steps and rightward steps but may consist of any pattern of leftward steps and rightward steps), using the sideways frieze relation

$$\begin{array}{ccc} & A & \\ B & & C \\ & D & \end{array} \quad : \quad C = (AD+1)/B$$

E.g., consider the partial frieze pattern

$$\begin{array}{cccccccccccc}
\ldots & 1 & & 1 & & 1 & & 1 & & 1 & \ldots \\
 & & x & & x' & & & & & & \\
 & y & & y' & & & & & & & \\
 & & z & & z' & & & & & & \\
\ldots & 1 & & 1 & & 1 & & 1 & & 1 & \ldots
\end{array}$$

Given non-zero values of x, y, and z, one can successively compute $y' = (xz+1)/y$, $x' = (y+1)/x$, and $z' = (y+1)/z$, obtaining a new zig-zag of entries x', y', z' connecting the top and bottom rows. For generic choices of non-zero x, y, z, one has x', y', z' non-zero as well, so the procedure can be repeated, yielding further zig-zags of entries. Happily (and perhaps surprisingly), after six iterations of the procedure one will recover the original numbers x, y, z six places to the right of their original position (unless one has unluckily chosen x, y, z so as to cause one to encounter an indeterminate expression of the form $0/0$ from the recurrence).

To dodge the issue of indeterminate expressions, we embrace indeterminacy by regarding x, y, z as formal quantities, not specific numbers, so that x', y', z', etc. become rational functions of x, y, and z. Then our recurrence ceases to be problematic. Indeed, one finds that the rational functions that arise are of a special kind, namely, Laurent polynomials with positive coefficients.

We can see why this is so by incorporating weighted edges into our matchings model. Returning to the triangulated hexagon from section 2, associate the values x, y, and z with the diagonals joining vertices 2 and 6, vertices 2 and 5, and vertices 3 and 5, respectively. Call these the formal weights of the diagonals. Also assign weight 1 to each of the 6 sides of the hexagon. Now construct the graph G as before, only this time assigning weights to all the edges. Specifically, if v is a black vertex of G that corresponds to a vertex of the n-gon and w is a white vertex of G that corresponds to a triangle in the triangulation T that has v as one of its three vertices (and has v' and v'' as the other two vertices), then the edge in G that joins v and w should be assigned the weight of the side or diagonal in T that joins v' and v''. We now define $W_{i,j}$ as the sum of the weights of all the perfect matchings of the graph $G_{i,j}$ obtained by deleting vertices i and j (and all their incident edges) from G, where the weight of a perfect matching is the product of the weights of its constituent edges, and we define $M_{i,j}$ as $W_{i,j}$ divided by the product of the weights of all the diagonals (this product is xyz in our running example). These $M_{i,j}$'s, which are rational functions of x, y, and z, generalize the numbers denoted by $m_{i,j}$ earlier, since we recover the $m_{i,j}$'s from the $M_{i,j}$'s by setting $x = y = z = 1$. It is clear that each $W_{i,j}$ is a polynomial with positive coefficients, so each $M_{i,j}$ is a Laurent polynomial with positive coefficients. And, because of the normalization (division by xyz), we have gotten each $M_{i,i+1}$ to equal 1. So the table of rational functions $M_{i,j}$ is exactly what we get by running our recurrence from left to right. When we pass from x, y, z to x', y', z', we are effectively rotating our triangulation by one-sixth of a full turn; six iterations bring us back to where we started.

We have proved:

THEOREM 3.1. *Given any assignment of formal weights to $n-3$ entries in an $(n-1)$-rowed table that form a zig-zag joining the top row (consisting of all 1's) to the bottom row (consisting of all 1's), there is a unique assignment of rational functions to all the entries in the table so that the frieze relation is satisfied. These rational functions of the original $n-3$ variables have glide-reflection symmetry that gives each row period n. Furthermore, each of the rational functions in the table is a Laurent polynomial with positive coefficients.*

Note that a zig-zag joining the top row to the bottom row corresponds to a triangulation T whose dual tree is just a path. Not every triangulation is of this kind. In general, the entries in a frieze pattern that correspond to the diagonals of a triangulation T do not form a zig-zag path, so it is not clear from the frieze pattern how to extend the known entries to the unknown entries. In such a case, it is best to refer directly to the triangulation itself, and to use a generalization of the frieze relation, namely the (formal) Ptolemy relation [**5**]

$$M_{i,j}M_{k,l} + M_{j,k}M_{i,l} = M_{i,k}M_{j,l}$$

where i, j, k, l are four vertices of the n-gon listed in cyclic order. Since every triangulation of a convex n-gon can be obtained from every other by means of flips that replace one diagonal of a quadrilateral by the other diagonal, we can iterate the Ptolemy relation so as to solve for all of the $M_{i,j}$'s in terms of the ones whose values were given.

Up until now we have allowed the diagonals, but not the sides, of our n-gon to have indeterminate weights; that is, the sides have all had weight 1. We can remedy this seeming lack of generality by noting that if we multiply the weights of the three sides of any triangle in the triangulation T by some constant c, the effect is to multiply by c the weights of three edges of the graph G, namely, the three edges incident with the white vertex w associated with T. This has the effect of multiplying the weight of every perfect matching of every graph $G_{i,j}$ by c, and such a scaling has no effect on the Laurentness phenomenon.

Our combinatorial construction of Laurent polynomials associated with the diagonals of an n-gon is essentially nothing more than the type A case (more precisely, the A_{n-3} case) of the cluster algebra construction of Fomin and Zelevinsky [**14**]. The result that our matchings model yields, stated in a self-contained way, is as follows:

THEOREM 3.2. *Given any assignment of formal weights $x_{i,j}$ to the $2n-3$ edges of a triangulated convex n-gon, there is a unique assignment of rational functions to all $n(n-3)/2$ diagonals of the n-gon such that the rational* *functions assigned to the four sides and two diagonals of any quadrilateral determined by four of the n vertices satisfies the Ptolemy relation. These rational functions of the original* $2n-3$ *variables are Laurent polynomials with positive coefficients.*

The formal weights are precisely the cluster variables in the cluster algebra of type A_{n-3}, and the triangulations are the clusters. The periodicity phenomenon is a special case of a more general periodicity conjectured by Zamolodchikov and proved in the type A case independently by Frenkel and Szenes and by Gliozzi and Tateo; see [**14**] for details.

4. The tropical analogue

Since the sideways frieze relation involves only subtraction-free expressions in the cluster variables, our whole picture admits a tropical analogue (for background on tropical mathematics, see [**19**]) in which multiplication is replaced by addition, division by subtraction, addition by max, and 1 by 0. (One could use min instead of max, but max will be more useful for us.) In this new picture, the Ptolemy relation

$$d_{i,j}\, d_{k,l} + d_{j,k}\, d_{i,l} = d_{i,k}\, d_{j,l}$$

becomes the ultrametric relation

$$\max(d_{i,j} + d_{k,l}, d_{j,k} + d_{i,l}) = d_{i,k} + d_{j,l}.$$

Metrics satisfying this relation arise from finite collections of non-intersecting arcs that join points on the sides of the n-gon (not vertices) in pairs (which we will call finite laminations). For any pair of vertices i, j, we define $d_{i,j}$ as the smallest possible number of intersections between a path in the n-gon from i to j and the arcs in the finite lamination (we choose the path so as to avoid crossing any arc in the finite lamination more than once). Then these quantities $d_{i,j}$ satisfy the ultrametric relation. As in the non-tropical case, we can find all the quantities $d_{i,j}$ once we know the values associated with the sides of the n-gon and the diagonals belonging to some triangulation.

For an alternative picture, one can divide the laminated n-gon into a finite number of sub-regions, each of which is bounded by pieces of the boundary of the n-gon and/or arcs of the finite lamination; the vertices of the n-gon correspond to n special sub-regions (some of which may coincide with one another, if there is no arc in the finite lamination separating the associated vertices of the n-gon). Then the dual of this dissection of the n-gon is a tree with n specified leaf vertices (some of which may coincide), and $d_{i,j}$ is the graph-theoretic distance between leaf i and leaf j (which could be zero). We see that if we know $2n-3$ of these leaf-to-leaf distances, and the $2n-3$ pairs of leaves correspond to the sides and diagonals of a triangulated n-gon, then all of the other leaf-to-leaf distances can be expressed as piecewise-linear functions of the $2n-3$ specified distances. (For more on the graph metric on trees, see [**2**].)

5. A variant

Before leaving the topic of frieze patterns, I mention an open problem concerning a variant of Conway and Coxeter's definition, in which the frieze relation is replaced by the relation

$$\begin{array}{ccc} & A & \\ B & C & D \\ & E & \end{array} \quad : \; E = (BD - C)/A$$

and its sideways version

$$\begin{array}{ccc} & A & \\ B & C & D \\ & E & \end{array} \quad : \; D = (AE + C)/B \; .$$

Here, too, it appears that we can construct arrays that have the same sort of symmetries as frieze patterns by starting with a suitable zig-zag of entries (where successive downwards steps can go left, right, or straight) and proceeding from left to right. E.g., consider the partial table

$$\begin{array}{ccccccc} \dots & 1 & 1 & 1 & 1 & 1 & \dots \\ & & A & D & x & & \\ & B & E & y & & & \\ & & C & F & z & & \\ \dots & 1 & 1 & 1 & 1 & 1 & \dots \end{array}$$

where $A,\dots,F$ are pre-specified, and where we compute $y=(AC+E)/B$, $x=(y+D)/A$, $z=(y+F)/C$, etc. Then one can check that after exactly fourteen iterations of the procedure, one gets back the original numbers (in their original order). Moreover, along the way one sees Laurent polynomials with positive coefficients.

Define a "double zig-zag" to be a subset of the entries of an $(n-2)$-rowed table consisting of a pair of adjacent entries in each of the middle $n-4$ rows, such that the pair in each row is displaced with respect to the pair in the preceding and succeeding rows by at most one position.

CONJECTURE: Given any assignment of formal weights to the $2(n-4)$ entries in a double zig-zag in an $(n-2)$-rowed table, there is a unique assignment of rational functions to all the entries in the table so that the variant frieze relation is satisfied. These rational functions of the original $2(n-4)$ variables have glide-reflection symmetry that gives each row period $2n$. Furthermore, each of the rational functions in the table is a Laurent polynomial with positive coefficients.

There ought to be a way to prove this by constructing the numerators of these Laurent polynomials as sums of weights of perfect matchings of some suitable graph (or perhaps sums of weights of combinatorial objects more general than perfect matchings), and the numerators undoubtedly contain abundant clues as to how this can be done.

For $n=5,6,7,8$, it appears that the number of positive integer arrays satisfying the variant frieze relation is respectively 1, 5, 51, 868. This variant of the Catalan sequence does not appear to have been studied before. However, it should be said that these numbers were not computed in a rigorous fashion. Indeed, it is not clear that there really is a variant of the Catalan sequence operating here; that is to say, it is conceivable that beyond some point, the sequence becomes infinite (i.e., for some n there could be infinitely many $(n-2)$-rowed positive integer arrays satisfying the variant frieze relation).

6. Markoff numbers

A **Markoff triple** is a triple (x,y,z) of positive integers satisfying $x^2+y^2+z^2=3xyz$; e.g., the triple (2,5,29). A **Markoff number** is a positive integer that occurs in at least one such triple.

Writing the Markoff equation as $z^2-(3xy)z+(x^2+y^2)=0$, a quadratic equation in z, we see that if (x,y,z) is a Markoff triple, then so is (x,y,z'), where $z'=3xy-z=(x^2+y^2)/z$, the other root of the quadratic in z. (z' is positive because $z'=(x^2+y^2)/z$, and is an integer because $z'=3xy-z$.) Likewise for x and y.

The following claim is well-known (for an elegant proof, see [**1**]): Every Markoff triple (x,y,z) can be obtained from the Markoff triple $(1,1,1)$ by a sequence of such exchange operations, in fact, by a sequence of exchange operations that leaves two numbers alone and increases the third. E.g., $(1,1,1)\to(2,1,1)\to(2,5,1)\to(2,5,29)$.

Create a graph whose vertices are the Markoff triples and whose edges correspond to the exchange operations $(x,y,z)\to(x',y,z)$, $(x,y,z)\to(x,y',z)$, $(x,y,z)\to(x,y,z')$ where $x'=\frac{y^2+z^2}{x}$, $y'=\frac{x^2+z^2}{y}$, $z'=\frac{x^2+y^2}{z}$. This 3-regular graph is connected (see the claim in the preceding paragraph), and it is not hard to show that it is acyclic. Hence the graph is the 3-regular infinite tree.

This tree can be understood as the dual of the triangulation of the upper half plane by images of the modular domain under the action of the modular group. Concretely, we can describe this picture by using Conway's terminology of "lax vectors", "lax bases", and "lax superbases" ([**6**]).

A **primitive** vector $\mathbf{u}$ in a lattice L is one that cannot be written as $k\mathbf{v}$ for some vector $\mathbf{v}$ in L, with $k>1$. A **lax vector** is a primitive vector defined only up to sign; if $\mathbf{u}$ is a primitive vector, the associated lax vector is written $\pm\mathbf{u}$. A **lax base** for L is a set of two lax vectors $\{\pm\mathbf{u},\pm\mathbf{v}\}$ such that $\mathbf{u}$ and $\mathbf{v}$ form a basis for L. A **lax superbase** for L is a set of three lax vectors $\{\pm\mathbf{u},\pm\mathbf{v},\pm\mathbf{w}\}$ such that $\pm\mathbf{u}\pm\mathbf{v}\pm\mathbf{w}=\mathbf{0}$ (with appropriate choice of signs) and any two of $\mathbf{u},\mathbf{v},\mathbf{w}$ form a basis for L.

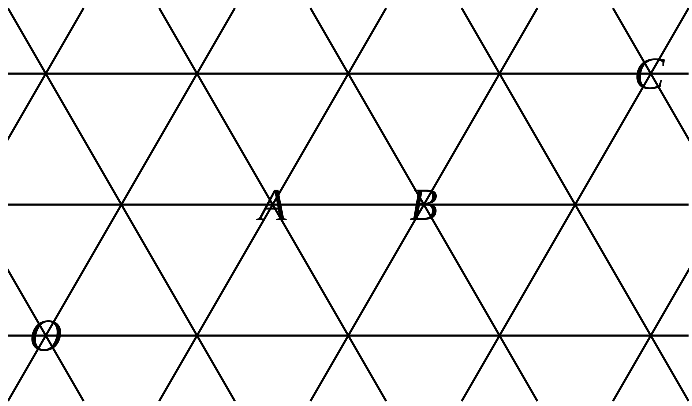
Each superbase $\{\pm\mathbf{u},\pm\mathbf{v},\pm\mathbf{w}\}$ contains the three bases $\{\pm\mathbf{u},\pm\mathbf{v}\}$, $\{\pm\mathbf{u},\pm\mathbf{w}\}$, $\{\pm\mathbf{v},\pm\mathbf{w}\}$ and no others. In the other direction, each base $\{\pm\mathbf{u},\pm\mathbf{v}\}$ is in the two superbases $\{\pm\mathbf{u},\pm\mathbf{v},\pm(\mathbf{u}+\mathbf{v})\}$, $\{\pm\mathbf{u},\pm\mathbf{v},\pm(\mathbf{u}-\mathbf{v})\}$ and no others.

The **topograph** is the graph whose vertices are lax superbases and whose edges are lax bases, where each superbase is incident with the three bases in it. This gives a 3-valent tree whose vertices correspond to the lax superbases of L, whose edges correspond to the lax bases of L, and whose "faces" correspond to the lax vectors in L.

The lattice L that we will want to use is the triangular lattice $L=\{(x,y,z)\in\mathbb{Z}^3\ :\ x+y+z=0\}$ (or $\mathbb{Z}^3/\mathbb{Z}\mathbf{v}$ where $\mathbf{v}=(1,1,1)$, if you prefer).

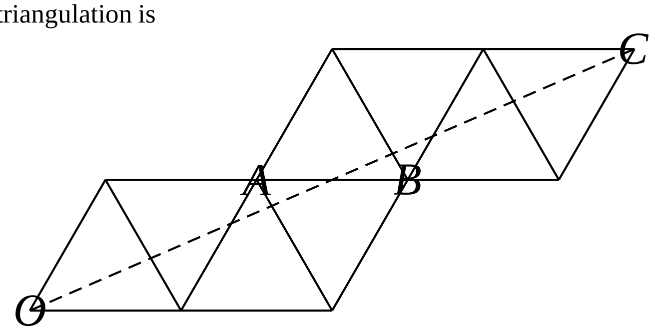
Using this terminology, I can now state the main idea of this section: Unordered Markoff triples are associated with lax superbases of the triangular lattice, and Markoff numbers with lax vectors of the triangular lattice. For example, the unordered Markoff triple $2,5,29$ will correspond to the lax superbase $\{\pm\mathbf{u},\pm\mathbf{v},\pm\mathbf{w}\}$ where $\mathbf{u}=\vec{OA}$,

$\mathbf{v} = \vec{OB}$, and $\mathbf{w} = \vec{OC}$, with O, A, B, and C forming a fundamental parallelogram for the triangular lattice, as shown below.



The Markoff numbers 1, 2, 5, and 29 will correspond to the primitive vectors $\vec{AB}$, $\vec{OA}$, $\vec{OB}$, and $\vec{OC}$.

To find the Markoff number associated with a primitive vector $\vec{OX}$, take the union R of all the triangles that segment OX passes through. The underlying lattice provides a triangulation of R. E.g., for the vector $\mathbf{u} = \vec{OC}$ from the previous figure, the triangulation is



Turn this into a planar bipartite graph as in Part I, let $G(\mathbf{u})$ be the graph that results from deleting vertices O and C, and let $M(\mathbf{u})$ be the number of perfect matchings of $G(\mathbf{u})$. (If $\mathbf{u}$ is a shortest vector in the lattice, put $M(\mathbf{u}) = 1$.)

THEOREM 6.1 (Gabriel Carroll, Andy Itsara, Ian Le, Gregg Musiker, Gregory Price, and Rui Viana). *If* $\{\mathbf{u},\mathbf{v},\mathbf{w}\}$ *is a lax superbase of the triangular lattice, then* $(M(\mathbf{u}),M(\mathbf{v}),M(\mathbf{w}))$ *is a Markoff triple. Every Markoff triple arises in this fashion. In particular, if* $\mathbf{u}$ *is a primitive vector, then* $M(\mathbf{u})$ *is a Markoff number, and every Markoff number arises in this fashion.*

(The association of Markoff numbers with the topograph is not new; what is new is the combinatorial interpretation of the association, by way of perfect matchings.)

PROOF. The base case, with

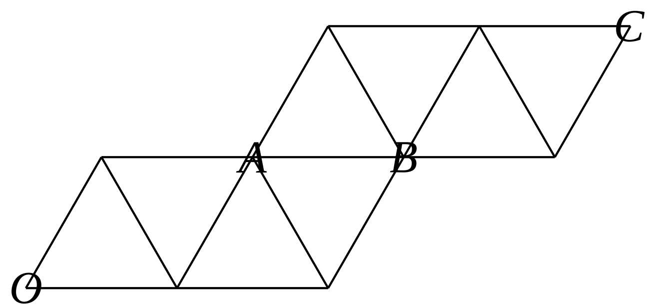
$$(M(\mathbf{e}_1),M(\mathbf{e}_2),M(\mathbf{e}_3)) = (1,1,1),$$

is clear. The only non-trivial part of the proof is the verification that

$$M(\mathbf{u}+\mathbf{v}) = (M(\mathbf{u})^2 + M(\mathbf{v})^2)/M(\mathbf{u}-\mathbf{v}).$$

E.g., in the picture below, we need to verify that

$$M(\vec{OC})M(\vec{AB}) = M(\vec{OA})^2 + M(\vec{OB})^2.$$



But if we rewrite the desired equation as

$$M(\vec{OC})M(\vec{AB}) = M(\vec{OA})M(\vec{BC}) + M(\vec{OB})M(\vec{AC})$$

we see that this is just Kuo's lemma. □

Remark 1: Some of the work done by the REACH students used a square lattice picture; this way of interpreting the Markoff numbers combinatorially was actually discovered first, in 2001–2002 (see [**4**]).

Remark 2: the original combinatorial model for the Conway-Coxeter numbers (found by Price) involved paths, not perfect matchings. Carroll turned this into a perfect matchings model, which made it possible to arrive at the matchings model of Itsara, Le, Musiker, and Viana via a different route.

More generally, one can put $M(\mathbf{e}_1) = x$, $M(\mathbf{e}_2) = y$, and $M(\mathbf{e}_3) = z$ (with $x, y, z > 0$) and recursively define

$$M(\mathbf{u}+\mathbf{v}) = (M(\mathbf{u})^2 + M(\mathbf{v})^2)/M(\mathbf{u}-\mathbf{v}).$$

Then for all primitive vectors $\mathbf{u}$, $M(\mathbf{u})$ is a Laurent polynomial in x, y, z; that is, it can be written in the form $P(x,y,z)/x^a y^b z^c$, where $P(x,y,z)$ is an ordinary polynomial in x, y, z (with non-zero constant term). The numerator $P(x,y,z)$ of each Markoff polynomial is the sum of the weights of all the perfect matchings of the graph $G(\mathbf{u})$, where edges have weight x, y, or z according to orientation. The triples $X = M(\mathbf{u})$, $Y = M(\mathbf{v})$, $Z = M(\mathbf{w})$ of rational functions associated with lax superbases are solutions of the equation

$$X^2 + Y^2 + Z^2 = \frac{x^2 + y^2 + z^2}{xyz} XYZ.$$

We have seen that these numerators $P(x,y,z)$ are polynomials with positive coefficients. This proves the following theorem:

THEOREM 6.2. *Consider the initial triple (x,y,z), along with any triple of rational functions in x, y, and z that can be obtained from the initial triple by a sequence of operations of the form $(X,Y,Z) \mapsto (X',Y,Z)$, $(X,Y,Z) \mapsto (X,Y',Z)$, or $(X,Y,Z) \mapsto (X,Y,Z')$, where $X' = (Y^2+Z^2)/X$, $Y' = (X^2+Z^2)/Y$, and $Z' = (X^2+Y^2)/Z$, Every rational function of x, y, and z that occurs in such a triple is a Laurent polynomial with positive coefficients.*

Fomin and Zelevinsky proved in [**13**] (Theorem 1.10) that the rational functions $X(x,y,z), Y(x,y,z), Z(x,y,z)$ are Laurent polynomials, but their methods did not prove positivity. An alternative proof of positivity, based on topological ideas, was given by Dylan Thurston [**20**].

It can be shown that if $\mathbf{u}$ inside the cone generated by $+\mathbf{e}_1$ and $-\mathbf{e}_3$, then $a < b > c$ and $(c+1)\mathbf{e}_1 - (a+1)\mathbf{e}_3 = \mathbf{u}$. (Likewise for the other sectors of L.) This implies that all the "Markoff polynomials" $M(\mathbf{u})$ are distinct (aside from the fact that $M(\mathbf{u}) = M(-\mathbf{u})$), and thus $M(\mathbf{u})(x,y,z) \neq M(\mathbf{v})(x,y,z)$ for all primitive vectors $\mathbf{u} \neq \pm\mathbf{v}$ as long as (x,y,z) lies in a dense G_δ set of real triples. This fact can be used to show [**20**] that, for a generic choice of hyperbolic structure on the once-punctured torus, no two simple geodesics have the same length.

7. Other directions for exploration

7.1. Other ternary cubics. Neil Herriot (another member of REACH) showed [**15**] that if we replace the triangular lattice used above by the tiling of the plane by isosceles right triangles (generated from one such triangle by repeated reflection in the sides), superbases of the square lattice correspond to triples (x,y,z) of positive integers satisfying either

$$x^2 + y^2 + 2z^2 = 4xyz$$

or

$$x^2 + 2y^2 + 2z^2 = 4xyz.$$

(Note that these two Diophantine equations are essentially equivalent, as the map $(x,y,z) \mapsto (2z,y,x)$ gives a bijection between solutions to the former and solutions to the latter.) This result, considered in conjunction with the result on Markoff numbers, raises the question of whether there might be some more general combinatorial approach to ternary cubic equations of similar shape.

Rosenberger [**18**] showed that there are exactly three ternary cubic equations of the shape $ax^2+by^2+cz^2 = (a+b+c)xyz$ for which all the positive integer solutions can be derived from the solution $(x,y,z) = (1,1,1)$ by means of the exchange operations $(x,y,z) \to (x',y,z)$, $(x,y,z) \to (x,y',z)$, and $(x,y,z) \to (x,y,z')$, with $x' = (by^2+cz^2)/ax$, $y' = (ax^2+cz^2)/by$, and $z' = (ax^2+by^2)/cz$. These three ternary cubic equations are

$$x^2+y^2+z^2 = 3xyz,$$

$$x^2+y^2+2z^2 = 4xyz,$$

and

$$x^2+2y^2+3z^2 = 6xyz.$$

Note that the triples of coefficients that occur here — (1,1,1), (1,1,2), and (1,2,3) — are precisely the triples that occur in the classification of finite reflection groups in the plane. Specifically, the ratios 1:1:1, 1:1:2, and 1:2:3 describe the angles of the three triangles — the 60-60-60 triangle, the 45-45-90 triangle, and the 30-60-90 triangle — that arise as the fundamental domains of the three irreducible two-dimensional reflection groups.

Since the solutions to the ternary cubic $x^2+y^2+z^2 = 3xyz$ describe properties of the tiling of the plane by 60-60-60 triangles, and solutions to the ternary cubic $x^2+y^2+2z^2 = 4xyz$ describe properties of the tiling of the plane by 45-45-90 triangles, the solutions to the ternary cubic $x^2+2y^2+3z^2 = 6xyz$ "ought" to be associated with some combinatorial model involving the reflection-tiling of the plane by 30-60-90 triangles. Unfortunately, the most obvious approach (based on analogy with the 60-60-60 and 45-45-90 cases) does not work. So we are left with two problems that may or may not be related: first, to find a combinatorial interpretation for the integers (or, more generally, the Laurent polynomials) that arise from solving the ternary cubic $x^2+2y^2+3z^2 = 6xyz$; and second, to find algebraic recurrences that govern the integers (or, more generally, the Laurent polynomials) that arise from counting (or summing the weights of) perfect matchings of graphs derived from the reflection-tiling of the plane by 30-60-90 triangles.

If there is a way to make the analogy work, one might seek to extend the analysis to other ternary cubics. It is clear how this might generalize on the algebraic side. On the geometric side, one might drop the requirement that the triangle tile the plane by reflection, and insist only that each angle be a rational multiple of 360 degrees. There is a relatively well-developed theory of "billiards flow" in such a triangle (see e.g. [**16**]) where a particle inside the triangle bounces off the sides following the law of reflection (angle of incidence equals angle of reflection) and travels along a straight line in between bounces. The path of such a particle can be unfolded by repeatedly reflecting the triangular domain in the side that the particle is bouncing off of, so that the unfolded path of the particle is just a straight line in the plane. Of special interest in the theory of billiards are trajectories joining a corner to a corner (possibly the same corner or possibly a different one); these are called saddle connections. The reflected images of the triangular domain form a triangulated polygon, and the saddle connection itself is a combinatorial diagonal of this polygon. It is unclear whether the combinatorics of such triangulations might contain dynamical information about the billiards flow, but if this prospect were to be explored, enumeration of matchings on the derived bipartite graphs would be one thing to try.

7.2. More variables. Another natural variant of the Markoff equation is $w^2+x^2+y^2+z^2 = 4wxyz$ (one special representative of a broader class called Markoff-Hurwitz equations; see [**1**]). The Laurent phenomenon applies here too: The four natural exchange operations convert an initial formal solution (w,x,y,z) into a quadruple of Laurent polynomials. (This is a special case of Theorem 1.10 in [**13**].)

Furthermore, the coefficients of these Laurent polynomials appear to be positive, although this has not been proved.

The numerators of these Laurent polynomials ought to be weight-enumerators for some combinatorial model, but I have no idea what this model looks like.

References

[1] A. Baragar, Integral solutions of the Markoff-Hurwitz equations, *J. of Number Theory* **49** (1994), 27–44.

[2] P. Buneman, A note on metric properties of trees, *J. Comb. Theory, Ser. B* **17** (1974), 48–50.

[3] P. Caldero and F. Chapoton, Cluster algebras as Hall algebras of quiver representations, arXiv:`math.RT/0410187`.

[4] G. Carroll, A. Itsara, I. Le, J. Propp, Markov numbers, Farey sequences, and the Ptolemy recurrence, unpublished memo (2003).

[5] G. Carroll and G. Price, Two new combinatorial models for the Ptolemy recurrence, unpublished memo (2003).

[6] J. Conway, The Sensual (Quadratic) Form, Mathematical Association of America, Washington, D.C. (1997).

[7] J.H. Conway and H.S.M. Coxeter, Triangulated Polygons and Frieze Patterns, *Math. Gaz.* **57** (1973), 87–94.
[8] J.H. Conway and R.K. Guy, The Book of Numbers, Springer-Verlag, New York (1996).
[9] H.S.M. Coxeter, Regular Polytopes, Macmillan, New York (1963); reprinted by Dover.
[10] H.S.M. Coxeter, Frieze patterns, *Acta Arithmetica* **18** (1971), 297–310.
[11] H.S.M. Coxeter, Cyclic sequences and frieze patterns, *Vinculum* **8** (1971), 4–7.
[12] H.S.M. Coxeter, Regular Complex, Polytopes, Cambridge University Press, London (1974).
[13] S. Fomin and A. Zelevinsky, The Laurent phenomenon, arXiv:`math.CO/0104241`.
[14] S. Fomin and A. Zelevinsky, Y-systems and generalized associahedra, arXiv:`hep-th/0111053`.
[15] N. Herriot, personal communication.
[16] R. Kenyon and J. Smillie, Billiards on rational-angled triangles, *Comment. Math. Helv.* **75** (2000), 65–108.
[17] E. Kuo, Applications of graphical condensation for enumerating matchings and tilings, *Theoret. Comput. Sci.* **319** (2004), 29–57; arXiv:`math.CO/0304090`.
[18] G. Rosenberger, Uber die diophantische Gleichung $ax^2 + by^2 + cz^2 = dxyz$, *J. Reine Angew. Math.* **305** (1979), 122–125.
[19] D. Speyer and B. Sturmfels, Tropical mathematics, arXiv:`math.CO/0408099`.
[20] D. Thurston, personal communication.
[21] D. Zeilberger, Dodgson's determinant-evaluation rule proved by two-timing men and women, *Elec. J. Comb.* **4**:2 (1997), R22; arXiv:`math.CO/9808079`.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN, MADISON, WISCONSIN, USA, 53706
E-mail address: `propp@math.wisc.edu`
URL: `http://www.math.wisc.edu/~propp`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The Horn recursion for Schur P- and Q- functions
Extended Abstract

Kevin Purbhoo and Frank Sottile

Abstract. A consequence of work of Klyachko and of Knutson-Tao is the Horn recursion to determine when a Littlewood-Richardson coefficient is non-zero. Briefly, a Littlewood-Richardson coefficient is non-zero if and only if it satisfies a collection of Horn inequalities which are indexed by smaller non-zero Littlewood-Richardson coefficients. There are similar Littlewood-Richardson numbers for Schur P- and Q- functions. Using a mixture of combinatorics of root systems, combinatorial linear algebra in Lie algebras, and the geometry of certain cominuscule flag varieties, we give Horn recursions to determine when these other Littlewood-Richardson numbers are non-zero. Our inequalities come from the usual Littlewood-Richardson numbers, and while we give two very different Horn recursions, they have the same sets of solutions. Another combinatorial by-product of this work is a new Horn-type recursion for the usual Littlewood-Richardson coefficients.

Résumé. Une des conséquences du travail de Klyachko et de Knutson-Tao est un système de récurrences de Horn pour déterminer quand un coefficient de Littlewood-Richardson est non nul. En bref, un tel coefficient est non nul si et seulement si il satisfait une collection d'inégalités de type Horn, dont les indices sont des coefficients de Littlewood-Richardson plus petits et non nuls. Il existe des nombres de Littlewood-Richardson comparables pour les P- et Q- fonctions de Schur. En utilisant des outils provenant combinatoire des systèmes de racines, d'algèbre linéaire dans le contexte des algébre de Lie, et de la géométrie des variétés de drapeaux cominiscules, nous obtenons un système de récurrences de type Horn pour déterminer quand cette famille de nombres de Littlewood-Richardson sont non nuls. Ces inégalités sont basées sur les nombres de Littlewood-Richardson habituels, et même si les deux systèmes sont très différents, ils ont la même solution. Une autre conséquence de ce travail est une nouvelle récurrence de type Horn pour les coefficients Littlewood-Richardson habituels.

Introduction

The Littlewood-Richardson numbers $a^{\lambda}_{\mu,\nu}$ for partitions λ, μ, ν are important in many areas of mathematics. For example, they are the structure constants of several related rings with distinguished bases: the ring of symmetric functions with its basis of Schur functions, the representation ring of $\mathfrak{sl}_n$ with its basis of irreducible highest weight modules, the external representation ring of the tower of symmetric groups with its basis of irreducible modules, and the cohomology ring of the Grassmannian with its basis of Schubert classes [**Ful97, Mac95, Sta99**]. The combinatorics of Littlewood-Richardson numbers are extremely interesting and now we have many formulas for them, including the original Littlewood-Richardson formula [**LR34**]. Despite this deep and prolonged interest in Littlewood-Richardson numbers, one of the most fundamental questions about them was not asked until about a decade ago:

$$\text{When is } a^{\lambda}_{\mu,\nu} \text{ non-zero?}$$

This question came from (of all places) a problem in linear algebra concerning the possible eigenvalues of a sum of hermitian matrices. The answer to this problem is given by the Horn inequalities: the eigenvalues

---

2000 *Mathematics Subject Classification.* Primary 05E15; Secondary 14M15.

Key words and phrases. Littlewood-Richardson numbers, Schur functions, flag manifolds.

Work of Sottile supported by the Clay Mathematical Institute and NSF CAREER grant DMS-0134860.
Work of Purbhoo supported by NSERC.

which can and do occur are the solutions to a set of linear inequalities, and the inequalities themselves come from non-negative integral eigenvalues solving this problem for smaller matrices.

The same inequalities answer our question about Littlewood-Richardson numbers. A Littlewood-Richardson number $a^{\lambda}_{\mu,\nu}$ is non-zero if and only if the triple of partitions (λ, μ, ν) satisfy certain linear inequalities, and the inequalities themselves come from triples indexing smaller non-zero Littlewood-Richardson coefficients. This description is a consequence of work of Klyachko [**Kly98**] which linked eigenvalues of sums of hermitian matrices, highest weight modules of $\mathfrak{sl}_n$, and the Schubert calculus for Grassmannians, and then Knutson and Tao's proof [**KT99**] of Zelevinsky's Saturation Conjecture [**Zel99**]. This work implies Horn's Conjecture [**Hor62**] about the eigenvalues of sums of Hermitian matrices. These results have wide implications in mathematics (see the surveys [**Ful98, Ful00**]) and have raised many new and evocative questions. For example, the Horn inequalities give the answer to questions in several different realms of mathematics (representation theory, combinatorics, Schubert calculus, eigenvalues), but it was initially mysterious why any one of these questions should have a recursive answer, as the proofs travelled through so many other parts of mathematics.

Another question, which was the *point de départ* for the results we describe here, is the following: are there related numbers whose non-vanishing has a similar recursive answer? Our main result is a recursive answer for the non-vanishing of the analogs of Littlewood-Richardson coefficients for Schur P-functions, and the same for Schur Q-functions. We give one set of inequalities which determine non-vanishing for the P-functions and a different set of inequalities for the Q-functions. Because each Schur P-function is a non-zero multiple of a corresponding Schur Q-function, a Littlewood-Richardson number for P-functions is non-zero if and only if the corresponding number for Q-functions is non-zero, and thus our two sets of inequalities have the same sets of solutions.

Another consequence of our work is a new set of recursive Horn-type inequalities for the ordinary Littlewood-Richardson numbers $a^{\lambda}_{\mu,\nu}$. While these new inequalities are clearly related to the ordinary Horn inequalities, they are definitely quite different. (We explain this below.)

Before we define some of these objects and give the different recursions, we remark that our results were proved using a mixture of the combinatorics of root systems, combinatorial linear algebra in Lie algebras, and the geometry of certain *cominuscule* flag varieties G/P. Cominuscule flag varieties are (almost all of) the flag varieties whose geometrically defined Bruhat order (which is the Bruhat order on the cosets W/W_P of the Weyl groups) is a distributive lattice.

The alert reader will notice that these inequalities for Schur P- and Q-functions are not strictly recursive because they are indexed by ordinary Littlewood-Richardson numbers which are non-zero. The reason for the term recursive is that the inequalities stem from a geometric recursion among all cominuscule flag varieties which is not evident from viewing only the subclass corresponding to, say the Schur Q-functions.

This abstract does not dwell on the geometry, but rather on the fascinating combinatorial consequences of these recursions. The last section of this extended abstract does however give a broad view of some of the key geometric ideas which underly our recursion. The results here are proved in the forthcoming preprint by the authors, "The recursive nature of the cominuscule Schubert calculus".

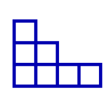
1. The classical Horn recursion

For more details and definitions concerning the various flavors of Schur functions that arise here, we recommend the book of Macdonald [**Mac95**]. *Schur functions S_λ* are symmetric functions indexed by *partitions λ*, which are weakly decreasing sequences of nonnegative integers, $\lambda\colon \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. The Schur function S_λ is homogeneous of degree $|\lambda| := \lambda_1 + \cdots + \lambda_n$. Schur functions form a basis for the $\mathbb{Z}$-algebra of symmetric functions. Thus there are integral *Littlewood-Richardson* numbers $a^{\lambda}_{\mu,\nu}$ defined by the identity

$$S_\mu \cdot S_\nu \;=\; \sum_{\lambda} a^{\lambda}_{\mu,\nu}\, S_\lambda\,.$$

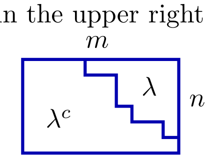
Homogeneity gives the necessary relation $|\lambda| = |\mu| + |\nu|$ for $a^{\lambda}_{\mu,\nu} \neq 0$.

A partition λ is represented by its *diagram*, which is a left-justified array of boxes in the positive quadrant with λ_i boxes in row i. Thus

$(4, 2, 1) \;\leftrightarrow$ 

Partitions are partially ordered by the inclusion of their diagrams. Let $n \times m$ be the rectangular partition with n parts, each of size m.

For $\lambda \subset n \times m$ ($\lambda_1 \leq m$ and $\lambda_{n+1} = 0$), define λ^c to be the partition obtained from the set-theoretic difference $n \times m - \lambda$ of diagrams (placing λ in the upper right corner of $n \times m$). Thus we have



.

The Horn-type inequalities we give are naturally stated in terms of *symmetric Littlewood-Richardson numbers*. For $\lambda, \mu, \nu \subset n \times m$, define

$$\begin{aligned} a_{\lambda,\mu,\nu} &:= \text{Coefficient of } S_{n\times m} \text{ in } S_\lambda S_\mu S_\nu \\ &= \text{Coefficient of } S_{\lambda^c} \text{ in } S_\mu S_\nu \;=\; a^{\lambda^c}_{\mu,\nu}\,. \end{aligned}$$

We say that a triple of partitions $\lambda, \mu, \nu \subset n \times m$ is *feasible* if $a_{\lambda,\mu,\nu} \neq 0$. This convenient terminology comes from geometry.

Definition 1.1. Suppose that $\lambda \subset n \times m$ and $\alpha \subset r \times (n-r)$, where $0 < r < n$. Define

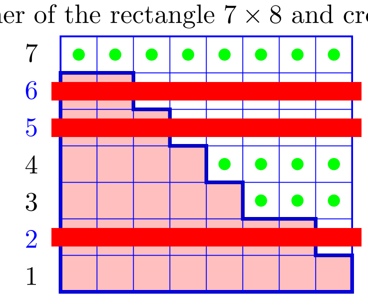
$$I_n(\alpha) \;:=\; \{n-r+1-\alpha_1,\; n-r+2-\alpha_2,\; \ldots,\; n-\alpha_r\}\,.$$

Draw λ in the upper right corner of the $n \times m$ rectangle, and number the rows Cartesian-style. Define $|\lambda|_\alpha$ to be the number of boxes that remain in λ after crossing out the rows indexed by $I_n(\alpha)$.

Example 1.2. Suppose that $n = 7$, $m = 8$, and $r = 3$, and we have $\lambda = 8654310$ and $\alpha = 311$. Then the set $I_7(\alpha)$ is

$$\{7-3+1-3,\; 7-3+2-1,\; 7-3+3-1\} \;=\; \{2,5,6\}\,.$$

If we place λ in the upper-right corner of the rectangle 7×8 and cross out the rows indexed by $I_7(\alpha)$,



we count the dots $\bullet$ to see that $|\lambda|_\alpha = 15$.

Theorem 1.3 (Classical Horn Recursion: Klyachko [**Kly98**], Knutson-Tao [**KT99**]).
A triple $\lambda, \mu, \nu \subset n \times m$ is feasible if and only if $|\lambda| + |\mu| + |\nu| = nm$, and

$$|\lambda|_\alpha + |\mu|_\beta + |\nu|_\gamma \;\leq\; (n-r)m$$

for all feasible triples $\alpha, \beta, \gamma \subset r \times (n-r)$ and for all $0 < r < n$.

The first condition, $|\lambda| + |\mu| + |\nu| = nm$, is due to homogeneity.

2. Symmetric Horn recursion

Since replacing a partition λ by its conjugate λ^t (interchanging rows with columns) induces an involution on the algebra of symmetric functions, there is a version of the Horn recursion where one crosses out columns instead of rows. It turns out that there are necessary inequalities obtained by crossing out both rows and columns, including possibly a different number of each. The cominuscule recursion reveals a sufficient subset of these.

Definition 2.1. Let $0 < r < \min\{n, m\}$ and suppose that $\lambda \subset n \times m$, $\alpha \subset r \times (n-r)$, and we have another partition $\alpha' \subset r \times (m-r)$. Define $I_n(\alpha)$ as before, and set

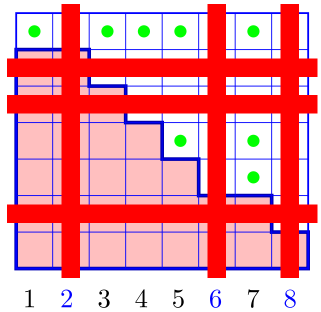
$$I_m(\alpha') \;:=\; \{m-r+1-\alpha'_1,\; m-r+2-\alpha'_2,\; \ldots,\; m-\alpha'_r\}\,.$$

Draw λ in the upper right corner of the $n\times m$ rectangle and cross out the rows indexed by $I_n(\alpha)$ and the columns indexed by $I_m(\alpha')$. Define $|\lambda|_{\alpha,\alpha'}$ to be the number of boxes that remain in λ.

Example 2.2. We use the same data as in Example 1.2, and set $\alpha'=410$. Then

$$I_8(\alpha') \;=\; \{8-3+1-4,\ 8-3+2-1,\ 8-3+3-0\} \;=\; \{2,6,8\}\,.$$

If we now cross out the rows indexed by $I_7(\alpha)$ and the columns indexed by $I_8(\alpha')$,



we count the dots • to see that $|\lambda|_{\alpha,\alpha'}=8$.

Theorem 2.3 (Symmetric Horn Recursion).
A triple $\lambda,\mu,\nu\subset n\times m$ is feasible if and only if $|\lambda|+|\mu|+|\nu|=nm$, and

$$|\lambda|_{\alpha,\alpha'}+|\mu|_{\beta,\beta'}+|\nu|_{\gamma,\gamma'} \;\le\; (m-r)(n-r)$$

for all pairs of feasible triples $\alpha,\beta,\gamma\subset r\times(n-r)$ and $\alpha',\beta',\gamma'\subset r\times(m-r)$ and for all $0<r<\min\{m,n\}$.

3. Schur P- and Q- functions

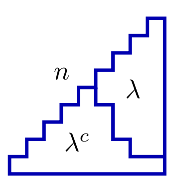
The algebra of symmetric functions has a natural odd subalgebra which comes from its structure as a combinatorial Hopf algebra [**ABS06**]. This algebra was first studied by Schur in the context of the theory of projective representations of the symmetric group. This odd subalgebra has a pair of distinguished bases, the *Schur P-functions* and the *Schur Q-functions*. They are indexed by strict partitions, which are strictly decreasing sequences of positive integers λ: $\lambda_1>\lambda_2>\cdots>\lambda_k>0$. They are proportional: $Q_\lambda=2^kP_\lambda$, where λ has k parts.

We have Littlewood-Richardson coefficients $c^\lambda_{\mu,\nu}$ and $d^\lambda_{\mu,\nu}$ indexed by triples of strict partitions and defined by the identities

$$Q_\mu\cdot Q_\nu \;=\; \sum_\lambda c^\lambda_{\mu,\nu}\,Q_\lambda \qquad\text{and}\qquad P_\mu\cdot P_\nu \;=\; \sum_\lambda d^\lambda_{\mu,\nu}\,P_\lambda\ .$$

Combinatorial formulas for these numbers were given in work of Worley [**Wor84**], Sagan [**Sag87**], and Stembridge [**Ste89**].

Let $\triangle_n\colon n>n{-}1>\cdots>2>1$ be the strict partition of staircase shape. Then $\lambda\subset\triangle_n$ if $\lambda_1\le n$. If $\lambda\subset\triangle_n$, define λ^c to be the partition obtained from the set-theoretic difference $\triangle_n-\lambda$ of diagrams (placing λ in the upper right corner of $\triangle_n$). Thus we have



As before, the Horn-type inequalities are naturally stated in terms of symmetric Littlewood-Richardson numbers. For $\lambda,\mu,\nu\subset\triangle_n$, define

$$\begin{aligned} c_{\lambda,\mu,\nu} &:= \text{Coefficient of } Q_{\triangle_n} \text{ in } Q_\lambda Q_\mu Q_\nu\\ &= \text{Coefficient of } Q_{\lambda^c} \text{ in } Q_\mu Q_\nu \;=\; c^{\lambda^c}_{\mu,\nu}\,. \end{aligned}$$

A triple of strict partitions $\lambda,\mu,\nu\subset n\times m$ is *feasible* if $c_{\lambda,\mu,\nu}\neq 0$.

We similarly define symmetric Littlewood-Richardson numbers $d_{\lambda,\mu,\nu}$ for the Schur P-functions. Since the two bases are proportional, the corresponding coefficients are as well. In particular the sets of triples λ, μ, ν for which the corresponding coefficients are feasible are the same. Nevertheless, we give two very different sets of inequalities which determine the feasibility of these numbers, arising from the different geometric origins of Schur Q-functions and Schur P-functions.

Definition 3.1. Let $0 < r < n$ and suppose that $\lambda \subset \triangle_n$ is a strict partition and $\alpha \subset r \times (n-r)$ is an ordinary partition. Draw λ in the upper right corner of the staircase $\triangle_n$. Number the inner corners $1, 2, \ldots, n$ and, for each number in $I_n(\alpha)$, cross out the row and column emanating from that inner corner. Then let $[\lambda]_\alpha$ be the number of boxes that remain in λ.

Definition 3.2. Let $0 < r < n+1$ and suppose that $\lambda \subset \triangle_n$ is a strict partition and $\alpha \subset r \times (n+1-r)$ is an ordinary partition. Draw λ in the upper right corner of the staircase $\triangle_n$. Number the outer corners $1, 2, \ldots, n, n+1$ and for each number in $I_{n+1}(\alpha)$, cross out the row and column emanating from the corresponding outer corner. Then let $\{\lambda\}_\alpha$ be the number of boxes that remain in λ.

Example 3.3. For example, suppose that $n = 8$ and $r = 4$, we have $\lambda = 8643$ and $\alpha = 4220$. Then

$$\begin{aligned} I_8(\alpha) &= \{8-4+1-4,\ 8-4+2-2,\ 8-4+3-2,\ 8-4+4-0\} = \{1,4,5,8\}, \\ I_9(\alpha) &= \{9-4+1-4,\ 9-4+2-2,\ 9-4+3-2,\ 9-4+4-0\} = \{2,5,6,9\}. \end{aligned}$$

and if we place λ in the upper-right corner of the rectangle $\triangle_8$ and cross out the rows and columns emanating from the inner corners indexed by $I_8(\alpha)$, we see that $[\lambda]_\alpha = 6$. If we instead cross out the rows and columns emanating from the outer corners indexed by $I_9(\alpha)$, we see that $\{\lambda\}_\alpha = 5$. The two diagrams are shown in Figure 1, on the left and right, respectively.

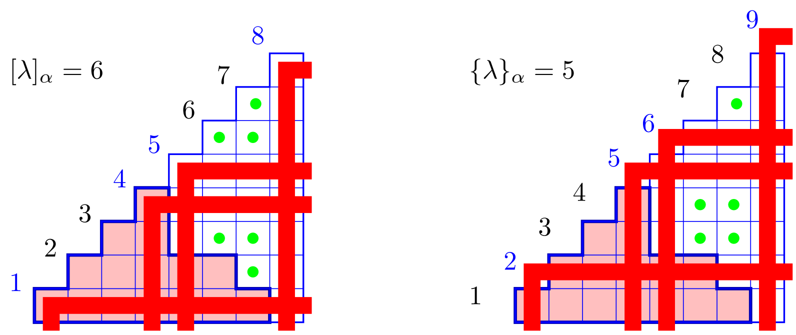


Figure 1. Computation of $[\lambda]_\alpha = 6$ and of $\{\lambda\}_\alpha = 5$

Note that the homogeneity of the multiplication of Schur P-functions and Schur Q-functions implies that

$$|\lambda| + |\mu| + |\nu| = |\triangle_n| = \binom{n+1}{2}, \tag{3.1}$$

is necessary for a triple $\lambda, \mu, \nu \subset \triangle_n$ to be feasible.

Theorem 3.4 (Horn recursion for Schur P- and Q-functions).
A triple $\lambda, \mu, \nu \subset \triangle_n$ of strict partitions is feasible if and only if one of the following two equivalent conditions hold:
(1) *The homogeneity condition* (3.1) *holds, and for all feasible $\alpha, \beta, \gamma \subset r \times (n-r)$ and all $0 < r < n$, we have*

$$[\lambda]_\alpha + [\mu]_\beta + [\nu]_\gamma \leq \binom{n+1-r}{2},$$

or else

(2) *The homogeneity condition* (3.1) *holds, and for all feasible* $\alpha, \beta, \gamma \subset r \times (n+1-r)$ *and all* $0 < r < n+1$ *with* r *even, we have*

$$\{\lambda\}_\alpha + \{\mu\}_\beta + \{\nu\}_\gamma \ \leq \ \binom{n+1-r}{2}.$$

4. Remarks on the geometry of the proof

We first give some general idea of the ingredients in our proof, and then explain a little bit of the relation of this geometry to the combinatorics given here.

A flag manifold G/P (G is a reductive algebraic group and P is a parabolic subgroup) has a Bruhat decomposition into Schubert cells indexed by cosets W/W_P, where W is the Weyl group of G and W_P that of P. The closures of the Schubert cells are the Schubert varieties, and cohomology classes associated to them (Schubert classes) form bases for the cohomology ring of G/P. Standard results in geometry show that the structure constants (generalized Littlewood-Richardson numbers) are the number of points in triple intersections of general translates of Schubert varieties (and hence are non-negative).

If a structure constant is non-zero, then any triple intersection of corresponding Schubert varieties (not just a general intersection) is non-empty, and general intersections are transverse. Conversely, if a structure constant is zero, then any corresponding general intersection is empty, and a non-empty intersection is never transverse. The key idea is to replace the difficult question on non-emptiness of a general intersection of Schubert varieties by the easier question of the transversality of a (not completely general) intersection. This was used by one of us to show transversality of intersections in the Grassmannian of lines [**Sot97**], but its use to study the Horn problem is due to Belkale [**Bel02**], who first gave a geometric proof of the Horn recursion for the Grassmannian.

This idea transfers the analysis from the flag manifold G/P to its tangent space T_pG/P at a given point p. In fact, all of our diagrams are just pictures of the root-space decompositions of T_pG/P for the corresponding varieties. In our proof, we consider three Schubert varieties which contain the point p, and then move them independently by the stabilizer P of p so that they are otherwise general. If it is possible to move the three tangent spaces inside of T_pG/P so that they meet transversally, then the triple is feasible, and if not, then it is not.

This explains where cominuscule flag manifolds come in. The maximal reductive, or Levi, subgroup L of the parabolic group P acts on the tangent space T_pG/P to G/P at that point. Our arguments (moving the tangent spaces to Schubert varieties around by elements of L) require that L act on T_pG/P with finitely many orbits, and this is one characterization of cominuscule flag manifolds G/P.

It also explains why there is a recursion. The argument requires us to consider the stabilizer Q in L of a certain linear subspace of T_pG/P—the tangent space to an orbit of L through a general point of the intersection of general translates of the tangents to the three Schubert varieties. Then the Schubert calculus inside of the smaller flag manifold L/Q is used to analyze the transversality of that triple intersection. Fortunately, the flag manifold L/Q is itself cominuscule, which is the source of our recursion.

We briefly illustrate these comments on the three flag manifolds that arose in this extended abstract.

4.1. The classical Grassmannian. Let $Gr(n, m{+}n)$ be the Grassmannian of n-planes in $m{+}n$ space. The general linear group $GL(m+n, \mathbb{C})$ acts on $Gr(n, m{+}n)$. If $H \in Gr(n, m{+}n)$ then $T_H Gr(n, m{+}n)$ may be identified with the set of n by m matrices (more precisely with $\mathrm{Hom}(H, \mathbb{C}^{m+n}/H)$). The Levi subgroup is the group $GL(n, \mathbb{C}) \times GL(m, \mathbb{C})$ which acts linearly on the rows and columns of n by m matrices. The orbits of this group are simply matrices of a fixed rank, r, and the subgroup Q is the stabilizer of a pair (K, K'), where $K \subset H$ and $K' \subset \mathbb{C}^{m+n}/H$ both have dimension r. This explains why in Definition 2.1, the number of rows crossed out equals the number of columns crossed out. The smaller cominuscule flag variety L/Q is the product of two Grassmannians, $Gr(r, n) \times Gr(r, m)$.

The Schubert varieties of $Gr(n, m{+}n)$ are indexed by partitions λ which fit in the $n \times m$ rectangle, and its cohomology ring is the algebra of Schur functions with these restricted indices.

4.2. The Lagrangian Grassmannian. Fix a non-degenerate alternating bilinear (symplectic) form on $\mathbb{C}^{2n}$. Let $LG(n)$ be the set of maximal isotropic (Lagrangian) subspaces in $\mathbb{C}^{2n}$, each of which has dimension n. This is the quotient of the symplectic group by the parabolic subgroup corresponding to the long root, $LG(n) = Sp(2n, \mathbb{C})/P_0$.

Since $H \in LG(n)$ is isotropic the symplectic form identifies $\mathbb{C}^{2n}/H$ with the dual of H, and $T_H LG(n)$ is identified with the space of quadratic forms on H. The Levi subgroup is the general linear group $GL(H)$. Identifying H with $\mathbb{C}^n$, the Levi becomes $GL(n,\mathbb{C})$ and $T_H LG(n)$ is the set of $n \times n$ symmetric matrices. (Symmetric matrices are parametrized by their weakly lower triangular parts, which correspond to the staircase shape $\triangle_n$ where the order of the columns has been reversed.) The general linear group acts simultaneously on the rows and columns of symmetric matrices. The orbits are simply symmetric matrices of a fixed rank, r, and the subgroup Q is the stabilizer of the null space of such a matrix. The smaller cominuscule flag variety L/Q is the Grassmannian $G(r,n)$.

The Schubert varieties of $LG(n)$ are indexed by strict partitions λ which fit inside the staircase $\triangle_n$, and its cohomology ring is the algebra of Schur Q-functions with these restricted indices.

4.3. The Orthogonal Grassmannian. Fix a non-degenerate symmetric bilinear form on $\mathbb{C}^{2n+2}$. The set of maximal isotropic subspaces (each of which has dimension $n+1$) of $\mathbb{C}^{2n+2}$ has two isomorphic components. Let $OG(n+1)$ be one of these components. This is the quotient of the even orthogonal group by a parabolic subgroup P corresponding to one of the roots at the fork in the Dynkin diagram, $OG(n+1) = SO(2n+2,\mathbb{C})/P$.

If $H \in OG(n+1)$ is isotropic, then $T_H OG(n+1)$ is identified with the space of alternating forms on H. The Levi subgroup is the general linear group $GL(H)$. Identifying H with $\mathbb{C}^{n+1}$, then the Levi becomes $GL(n+1,\mathbb{C})$ and $T_H OG(n+1)$ is the set of $(n+1) \times (n+1)$ anti-symmetric matrices. (Anti-symmetric matrices are parametrized by their lower triangular parts, and these strictly lower triangular matrices correspond to the staircase shape where the order of the columns has been reversed.) The general linear group acts simultaneously on the rows and columns of anti-symmetric matrices. The orbits are simply anti-symmetric matrices of a fixed rank. However, and this comes from the details of the proof and the roots of $SO(2n+2,\mathbb{C})$, the subgroup Q is the stabilizer of an even-dimensional subspace of H. The smaller cominuscule flag variety L/Q is the Grassmannian $G(r, n+1)$, where r is even.

The Schubert varieties of $OG(n+1)$ are indexed by strict partitions λ which fit inside the staircase $\triangle_n$, and its cohomology ring is the algebra of Schur P-functions with these restricted indices.

References

[ABS06] Marcelo Aguiar, Nantel Bergeron, and Frank Sottile, *Combinatorial Hopf Algebras*, Compositio Mathematica **142** (2006), 1–30.

[Bel02] Prakash Belkale, *Geometric proofs of Horn and saturation conjectures*, 2002, `math.AG/0208107`.

[Ful97] William Fulton, *Young tableaux*, Cambridge University Press, Cambridge, 1997, With applications to representation theory and geometry.

[Ful98] ———, *Eigenvalues of sums of Hermitian matrices (after A. Klyachko)*, Astérisque (1998), no. 252, Exp. No. 845, 5, 255–269, Séminaire Bourbaki. Vol. 1997/98.

[Ful00] ———, *Eigenvalues, invariant factors, highest weights, and Schubert calculus*, Bull. Amer. Math. Soc. (N.S.) **37** (2000), no. 3, 209–249 (electronic).

[Hor62] Alfred Horn, *Eigenvalues of sums of Hermitian matrices*, Pacific J. Math. **12** (1962), 225–241.

[Kly98] A.A. Klyachko, *Stable bundles, representation theory, and Hermitian operators*, Selecta Math. **4** (1998), 419–445.

[KT99] Allen Knutson and Terence Tao, *The honeycomb model of* $\mathrm{GL}_n(\mathbf{C})$ *tensor products. I. Proof of the saturation conjecture*, J. Amer. Math. Soc. **12** (1999), no. 4, 1055–1090.

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

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

[Sag87] Bruce Sagan, *Shifted tableaux, schur q-functions and a conjecture of stanley*, J. Combin. Th., Ser. A **45** (1987), 62–103.

[Sot97] F. Sottile, *Enumerative geometry for the real Grassmannian of lines in projective space*, Duke Math. J. **87** (1997), no. 1, 59–85.

[Sta99] Richard P. Stanley, *Enumerative combinatorics. Vol. 2*, Cambridge University Press, Cambridge, 1999, With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Ste89] John Stembridge, *Shifted tableaux and the projective representations of the symmetric group*, Adv. Math. **74** (1989), 87–134.

[Wor84] Dale Worley, *A theory of shifted young tableaux*, Ph.D. thesis, MIT, 1984.

[Zel99] Andrei Zelevinsky, *Littlewood-Richardson semigroups*, New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97), Math. Sci. Res. Inst. Publ., vol. 38, Cambridge Univ. Press, Cambridge, 1999, pp. 337–345.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF BRITISH COLUMBIA, 1984 MATHEMATICS ROAD, VANCOUVER, B.C., V6T 1Z2, CANADA
E-mail address: kevinp@math.ubc.ca
URL: http://www.math.ubc.ca/~kevinp/

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY, COLLEGE STATION, TX 77843, USA
E-mail address: sottile@math.tamu.edu
URL: http://www.math.tamu.edu/~sottile

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Clusters, Coxeter-sortable elements and noncrossing partitions

Extended abstract. Please cite only the full version, available on the arXiv.

Nathan Reading

Abstract. We introduce Coxeter-sortable elements of a Coxeter group W. For finite W, we give bijective proofs that Coxeter-sortable elements are equinumerous with clusters and with noncrossing partitions. We characterize Coxeter-sortable elements in terms of their inversion sets and, in the classical cases, in terms of permutations.

Résumé. Nous introduisons dans ce travail la notion d'éléments sortables pour un groupe de Coxeter W. Dans le cas où W est fini, nous montrons que les éléments sortables sont en bijection avec les clusters ainsi qu'avec les partitions non croisées. Nous donnons une caractérisation des éléments sortables au moyen de leurs ensembles d'inversion et, dans les cas classiques, en terme de permutations.

Introduction

The famous Catalan numbers can be viewed as a special case of the W-*Catalan number*, which counts various objects related to a finite Coxeter group W. In many cases, the common numerology is the only known connection between different objects counted by the W-Catalan number. One collection of objects counted by the W-Catalan number is the set of noncrossing partitions associated to W, which play a role in low-dimensional topology, geometric group theory and non-commutative probability [**17**]. Another is the set of maximal cones of the cluster fan. The cluster fan is dual to the generalized associahedron [**9, 11**], a polytope whose combinatorial structure underlies cluster algebras of finite type [**12**].

This paper connects noncrossing partitions to associahedra via certain elements of W which we call *Coxeter-sortable* elements or simply *sortable* elements. For each Coxeter element c of W, there is a set of c-sortable elements, defined in the context of the combinatorics of reduced words. We prove bijectively that sortable elements are equinumerous with clusters and with noncrossing partitions. Sortable elements and the bijections are defined without reference to the classification of finite Coxeter groups, but the proof that these maps are indeed bijections refers to the classification. The bijections are well-behaved with respect to the refined enumerations associated to the Narayana numbers and to positive clusters.

In the course of establishing the bijections, we characterize[1] the sortable elements in terms of their inversion sets. Loosely speaking, we "orient" each rank-two parabolic subgroup of W and require that the inversion set of the element be "aligned" with these orientations. In particular, we obtain a characterization of the sortable elements in types A_n, B_n and D_n as permutations.

Because sortable elements are defined in terms of reduced words, it is natural to partially order them as an induced subposet of the weak order. Indeed, the definition of sortable elements arose from the study of certain lattice quotients of the weak order called Cambrian lattices. In the sequel [**22**] to this paper, we show that sortable elements are indeed a combinatorial model for the Cambrian lattices.

2000 *Mathematics Subject Classification.* 20F55 (Primary) 05E15, 05A15 (Secondary).

Key words and phrases. clusters, Coxeter-sortable elements, noncrossing partitions, W-Catalan combinatorics.

The author was partially supported by NSF grant DMS-0202430.

[1]This characterization is described in the full version.

Recently, Brady and Watt [**8**] observed that the cluster fan arises naturally in the context of noncrossing partitions. Their work and the present work constitute two seemingly different approaches to connecting noncrossing partitions to clusters. The relationship between these approaches is not yet understood.

The term "sortable" has reference to the special case where W is the symmetric group. For one particular choice of c, the definition of c-sortable elements of the symmetric group is as follows: Perform a bubble sort on a permutation π by repeatedly passing from left to right in the permutation and, whenever two adjacent elements are out of order, transposing them. For each pass, record the set of positions at which transpositions were performed. Then π is c-sortable if this sequence of sets is weakly decreasing in the containment order. There is also a choice of c (see Example 1.8) such that the c-sortable elements are exactly the 231-avoiding or stack-sortable permutations [**15**, Exercise 2.2.1.4–5].

1. Coxeter-sortable elements

Throughout this paper, W denotes a finite Coxeter group with simple generators S and reflections T. Some definitions apply also to the case of infinite W, but we confine the treatment of the infinite case to a series of remarks (Remarks 1.5, 2.4 and 3.5.).

The term "word" always means "word in the alphabet S." Later, we consider words in the alphabet T which, to avoid confusion, are called "T-words." A *cover reflection* of $w \in W$ is a reflection t such that $tw = ws$ with $l(ws) < l(w)$. The term "cover reflection" refers to the (right) weak order. This is the partial order on W whose cover relations are the relations of the form $w \gtrdot ws$ for $l(ws) < l(w)$, or equivalently, $w \gtrdot tw$ for a cover reflection t of w. For each $J \subseteq S$, let W_J be the subgroup of W generated by J. The notation $\langle s \rangle$ stands for the set $S - \{s\}$.

For the rest of the paper, c denotes a *Coxeter element*, that is, an element of W with a reduced word which is a permutation of S. An *orientation* of the Coxeter diagram for W is obtained by replacing each edge of the diagram by a single directed edge, connecting the same pair of vertices in either direction. Orientations of the Coxeter diagram correspond to Coxeter elements (cf. [**25**]) as follows: Given a Coxeter element c, any two reduced words for c are related by commutations of simple generators. An edge s—t in the diagram represents a pair of noncommuting simple generators, and the edge is oriented $s \longrightarrow t$ if and only if s precedes t in every reduced word for c.

We now define Coxeter-sortable elements. Fix a reduced word $a_1 a_2 \cdots a_n$ for a Coxeter element c. Write a half-infinite word

$$c^\infty = a_1 a_2 \cdots a_n a_1 a_2 \cdots a_n a_1 a_2 \cdots a_n \ldots$$

The *c-sorting word* for $w \in W$ is the lexicographically first (as a sequence of positions in c^∞) subword of c^∞ which is a reduced word for w. The c-sorting word can be interpreted as a sequence of subsets of S by rewriting

$$c^\infty = a_1 a_2 \cdots a_n | a_1 a_2 \cdots a_n | a_1 a_2 \cdots a_n | \ldots$$

where the symbol "|" is called a *divider*. The subsets in the sequence are the sets of letters of the c-sorting word which occur between adjacent dividers. This sequence contains a finite number of non-empty subsets, and furthermore if any subset in the sequence is empty, then every later subset is also empty. For clarity in examples, we often retain the dividers when writing c-sorting words for c-sortable elements.

An element $w \in W$ is *c-sortable* if its c-sorting word defines a sequence of subsets which is decreasing under inclusion. This definition of c-sortable elements requires a choice of reduced word for c. However, for a given w, the c-sorting words for w arising from different reduced words for c are related by commutations of letters, with no commutations across dividers. In particular, the set of c-sortable elements does not depend on the choice of reduced word for c.

Remark 1.1. The c-sortable elements have a natural search-tree structure rooted at the identity element. The edges are pairs v, w of c-sortable elements such that the c-sorting word for v is obtained from the c-sorting word for w by deleting the rightmost letter. This makes possible an efficient traversal of the set of c-sortable elements of W which, although it does not explicitly appear in what follows, allows various properties of c-sortable elements to be checked computationally in low ranks. Also, in light of the bijections of Theorems 2.1 and 3.2, an efficient traversal of the c-sortable elements leads to an efficient traversal of noncrossing partitions or of clusters.

The next two lemmas are immediate from the definition of c-sortable elements. Together with the fact that 1 is c-sortable for any c, they completely characterize c-sortability. A simple generator $s \in S$ is *initial* in (or is an *initial letter* of) a Coxeter element c if it is the first letter of some reduced word for c.

LEMMA 1.2. *Let s be an initial letter of c and let $w \in W$ have $l(sw) > l(w)$. Then w is c-sortable if and only if it is an sc-sortable element of $W_{\langle s \rangle}$.*

LEMMA 1.3. *Let s be an initial letter of c and let $w \in W$ have $l(sw) < l(w)$. Then w is c-sortable if and only if sw is scs-sortable.*

REMARK 1.4. In the dictionary between orientations of Coxeter diagrams (i.e. quivers) and Coxeter elements, the operation of replacing c by scs corresponds to the operation on quivers which changes a source into a sink by reversing all arrows from the source. This operation was used in [**16**] in generalizing the clusters of [**11**] to Γ-clusters, where Γ is a quiver of finite type. We thank Andrei Zelevinsky for pointing out the usefulness of this operation, which plays a key role throughout the paper.

REMARK 1.5. The definition of c-sortable elements is equally valid for infinite W. Lemmas 1.2 and 1.3 are valid and characterize c-sortability in the infinite case as well. However, we remind the reader that for all stated results in this paper, W is assumed to be finite.

EXAMPLE 1.6. Consider $W = B_2$ with $c = s_0 s_1$. The c-sortable elements are 1, s_0, $s_0 s_1$, $s_0 s_1 | s_0$, $s_0 s_1 | s_0 s_1$ and s_1. The elements $s_1 | s_0$ and $s_1 | s_0 s_1$ are not c-sortable.

We close the section with a discussion of the sortable elements of the Coxeter group $W = A_n$, realized combinatorially as the symmetric group S_{n+1}. Permutations $\pi \in S_{n+1}$ are written in *one-line notation* as $\pi_1 \pi_2 \cdots \pi_{n+1}$ with $\pi_i = \pi(i)$. The simple generators of S_{n+1} are $s_i = (i \;\; i{+}1)$ for $i \in [n]$.

A *barring* of a set U of integers is a partition of that set into two sets $\overline{U}$ and $\underline{U}$. Elements of $\overline{U}$ are *upper-barred* integers denoted $\overline{i}$ and *lower-barred* integers are elements of $\underline{U}$, denoted $\underline{i}$.

Recall that orientations of the Coxeter diagram correspond to Coxeter elements. The Coxeter diagram for S_{n+1} has unlabeled edges connecting s_i to s_{i+1} for $i \in [n-1]$. We encode orientations of the Coxeter diagram for S_{n+1} as barrings of $[2, n]$ by directing $s_i \to s_{i-1}$ for every $\overline{i} \in [2, n]$ and $s_{i-1} \to s_i$ for every $\underline{i} \in [2, n]$, as illustrated in Figure 1 for $c = s_8 s_7 s_4 s_1 s_2 s_3 s_5 s_6$ in S_9. Given a choice of Coxeter element, the corresponding barring is assumed.

barA.ps

FIGURE 1. Orientation and barring in S_9

We now define condition (A), which characterizes c-sortability of permutations. Condition (A) depends on the choice of c as follows: a fixed choice of c defines a barring as described above, and condition (A) depends on that fixed barring. A permutation $\pi \in S_{n+1}$ satisfies condition (A) if both of the following conditions hold:

(A1) π contains no subsequence $\overline{j}\,k\,i$ with $i < j < k$, and
(A2) π contains no subsequence $k\,i\,\underline{j}$ with $i < j < k$.

Notice that i and k appear in (A1) and (A2) without explicit barrings. This is because the barrings of i and k are irrelevant to the conditions. For example, to satisfy (A1), π may not contain any sequence of the form $\overline{j}\,k\,i$, regardless of the barrings of i and k.

THEOREM 1.7. *A permutation $\pi \in S_{n+1}$ is c-sortable if and only if it satisfies condition (A) with respect to the barring corresponding to c.*

EXAMPLE 1.8. For $W = S_{n+1}$ and $c = (n \;\; n{+}1) \cdots (2\;3)(1\;2)$, the c-sortable elements are exactly the 231-avoiding or stack-sortable permutations defined in [**15**, Exercise 2.2.1.4–5].

2. Sortable elements and noncrossing partitions

In this section, we define a bijection between sortable elements and noncrossing partitions. Recall that T is the set of reflections of W. Any element $w \in W$ can be written as a word in the alphabet T. To avoid confusion we always refer to a word in the alphabet T as a *T-word*. Any other use of the term "word" is assumed to refer to a word in the alphabet S. A *reduced* T-word for w is a T-word for w which has minimal length among all T-words for w. The *absolute length* of an element w of W is the length of a reduced T-word for w. This is not the usual length $l(w)$ of w, which is the length of a reduced word for w in the alphabet S.

The notion of reduced T-words leads to a prefix partial order on W, analogous to the weak order. Say $x \leq_T y$ if x possesses a reduced T-word which is a prefix of some reduced T-word for y. Equivalently, $x \leq_T y$ if every reduced T-word for x is a prefix of some reduced T-word for y. Since the alphabet T is closed under conjugation by arbitrary elements of W, the partial order $\leq_T$ is invariant under conjugation. The partial order $\leq_T$ can also be defined as a subword order: $x \leq_T y$ if and only if there is a reduced T-word for y having as a subword some reduced T-word for x. In particular, $x \leq_T y$ if and only if $x^{-1}y \leq y$.

The *noncrossing partition lattice* in W (with respect to the Coxeter element c) is the interval $[1, c]_T$, and the elements of this interval are called *noncrossing partitions*. The poset $[1, c]_T$ is graded and the rank of a noncrossing partition is its absolute length.

Let w be a c-sortable element and let $a = a_1 a_2 \cdots a_k$ be a c-sorting word for w. Totally order the inversions of w such that the ith reflection in the order is $a_1 a_2 \cdots a_{i-1} a_i a_{i-1} \cdots a_2 a_1$. Equivalently, t is the ith reflection in the order if and only if $tw = a_1 a_2 \cdots \hat{a}_i \cdots a_k$, where $\hat{a}_i$ indicates that a_i is deleted from the word. Write the set of cover reflections of w as a subsequence $t_1, t_2, \ldots, t_l$ of this order on inversions. Let nc_c be the map which sends w to the product $t_1 t_2 \cdots t_l$. Recall that the construction of a c-sorting word begins with an arbitrary choice of a reduced word for c. However, since any two c-sorting words for w are related by commutation of simple generators, $\mathrm{nc}_c(w)$ does not depend of the choice of reduced word for c.

Theorem 2.1. *For any Coxeter element c, the map $w \mapsto \mathrm{nc}_c(w)$ is a bijection from the set of c-sortable elements to the set of noncrossing partitions with respect to c. Furthermore nc_c maps c-sortable elements with k descents to noncrossing partitions of rank k.*

Recall that the descents of w are the simple generators $s \in S$ such that $l(ws) < l(w)$. Recall also that these are in bijection with the cover reflections of w. The basic tool for proving Theorem 2.1 is induction on rank and length, using Lemmas 1.2 and 1.3. A more complicated induction is used to prove the existence of the inverse map.

Example 2.2. We again consider the case $W = B_2$ and $c = s_0 s_1$. As a special case of the combinatorial realization of noncrossing partitions of type B given in [**24**], the noncrossing partitions in B_2 with respect to c are the centrally symmetric noncrossing partitions of the cycle shown below.

B2cycle.ps

Figure 2 illustrates the map nc_c for this choice of W and c.

w	1	$\hat{s}_0$	$s_0\hat{s}_1$	$s_0s_1\vert\hat{s}_0$	$\hat{s}_0s_1\vert s_0\hat{s}_1$	$\hat{s}_1$
$nc_c(w)$	1	s_0	$s_0s_1s_0$	$s_1s_0s_1$	$s_0 \cdot s_1$	s_1
	B2.1.ps	*B2.a.ps*	*B2.ab.ps*	*B2.aba.ps*	*B2.ababi.ps* ↓ *B2.ababii.ps*	*B2.b.ps*

Figure 2. The map nc_c

Example 2.3. Covering reflections of a permutation $\pi \in S_{n+1}$ are the transpositions corresponding to descents (pairs (π_i, π_{i+1}) with $\pi_i > \pi_{i+1}$). The map nc_c sends π to the product of these transpositions. The relations $\pi_i \equiv \pi_{i+1}$ for descents (π_i, π_{i+1}) generate an equivalence relation on $[n+1]$ which can be interpreted as a noncrossing partition (in the classical sense) of the cycle c. For $c = (n\ \ n{+}1) \cdots (2\ 3)(1\ 2)$ as in Example 1.8, this map between 231-avoiding permutations and classical (i.e. type A) noncrossing partitions is presumably known.

Remark 2.4. The definition of nc_c is valid for infinite W. However, Theorem 2.1 address the finite case only. In particular, for infinite W it is not even known whether nc_c maps c-sortable elements into the interval $[1, c]_T$.

Remark 2.5. As a byproduct of Theorem 2.1, we obtain a canonical reduced T-word for every x in $[1, c]_T$. The letters are the canonical generators of the associated parabolic subgroup, or equivalently the cover reflections of $(\mathrm{nc}_c)^{-1}(x)$, occurring in the order induced by the c-sorting word for $(\mathrm{nc}_c)^{-1}(x)$. This is canonical, up to the choice of reduced word for c. Changing the reduced word for c alters the choice of canonical reduced T-word for x by commutations of letters.

In [**1**], it is shown that for c bipartite, the natural labeling of $[1, c]_T$ is an EL-shelling with respect to the reflection order obtained from what we here call the c-sorting word for w_0. In particular, the labels on the unique maximal chain in $[1, x]_T$ constitute a canonical reduced T-word for x. It is apparent from the proof of [**1**, Theorem 3.5(ii)] that these two choices of canonical reduced T-words are identical in the bipartite case, for W crystallographic. Presumably the same is true for non-crystallographic W.

3. Sortable elements and clusters

In this section we define c-clusters, a slight generalization (from crystallographic Coxeter groups to all finite Coxeter groups) of the Γ-clusters of [**16**]. These in turn generalize the clusters of [**11**]. The main result of this section is a bijection between c-sortable elements and c-clusters.

We build the theory of clusters within the framework of Coxeter groups, rather than in the framework of root systems. This is done in order to avoid countless explicit references to the map between positive roots and reflections in what follows. Readers familiar with root systems will easily make the translation to the language of *almost positive roots* of [**11**] and [**16**].

Let $-S$ denote the set $\{-s : s \in S\}$ of formal negatives of the simple generators of W, and let $T_{\geq -1}$ be $T \cup (-S)$. (Recall that T is the set of all reflections of W.) The notation T_J stands for $T \cap W_J$ and $(T_J)_{\geq -1}$ denotes $T_J \cup (-J)$.

For each $s \in S$, define an involution $\sigma_s : T_{\geq -1} \to T_{\geq -1}$ by

$$\sigma_s(t) := \begin{cases} -t & \text{if } t = \pm s, \\ t & \text{if } t \in (-S) \text{ and } t \neq -s, \text{ or} \\ sts & \text{if } t \in T - \{s\}. \end{cases}$$

We now define a symmetric binary relation $\|_c$ called the *c-compatibility* relation.

Proposition 3.1. *There exists a unique family of symmetric binary relations $\|_c$ on $T_{\geq -1}$, indexed by Coxeter elements c, with the following properties:*

(i) *For any $s \in S$, $t \in T_{\geq -1}$ and Coxeter element c,*

$$-s \,\|_c\, t \text{ if and only if } t \in (T_{\langle s \rangle})_{\geq -1}.$$

(ii) *For any $t_1, t_2 \in T_{\geq -1}$ and any initial letter s of c,*

$$t_1 \,\|_c\, t_2 \text{ if and only if } \sigma_s(t_1) \,\|_{scs}\, \sigma_s(t_2).$$

A *c-compatible subset* of $T_{\geq -1}$ is a set of pairwise c-compatible elements of $T_{\geq -1}$. A *c-cluster* is a maximal c-compatible subset. A c-cluster is called *positive* if it contains no element of $-S$.

Let w be a c-sortable element with c-sorting word $a_1 a_2 \cdots a_k$. If $s \in S$ occurs in a then the *last reflection* for s in w is $a_1 a_2 \cdots a_j a_{j-1} \cdots a_2 a_1$, where a_j is the rightmost occurrence of s in a. If s does not occur in a then the last reflection for s in w is the formal negative $-s$. Let $\mathrm{cl}_c(w)$ be the set of last reflections of w. This is an n-element subset of $T_{\geq -1}$. This map does not depend on the choice of reduced word for c, because any two c-sorting words for w are related by commutations of simple generators.

Theorem 3.2. *The map $w \mapsto \mathrm{cl}_c(w)$ is a bijection from the set of c-sortable elements to the set of c-clusters. Furthermore, cl_c restricts to a bijection between c-sortable elements with full support and positive c-clusters.*

The strategy for proving Theorem 3.2 is the same as for Theorem 2.1, but with fewer complications. We argue by induction on rank and length.

Example 3.3. We continue the example of $W = B_2$ and $c = s_0 s_1$. Clusters in B_2 correspond to collections of diagonals which define centrally symmetric triangulations of the hexagon shown below.

B2seed.ps

Each element of $T_{\geq -1}$ is represented by a diameter or a centrally symmetric pair of diagonals. For details, see [**11**, Section 3.5]. Figure 3 illustrates the map cl_c on c-sortable elements for this choice of W and c.

w	1	s_0	s_0s_1	$s_0s_1\vert s_0$	$s_0s_1\vert s_0s_1$	s_1
$\mathrm{cl}_c(w)$	$-s_0, -s_1$	$s_0, -s_1$	$s_0, s_0s_1s_0$	$s_1s_0s_1, s_0s_1s_0$	$s_1s_0s_1, s_1$	$-s_0, s_1$
	B2cl.1.ps	*B2cl.a.ps*	*B2cl.ab.ps*	*B2cl.aba.ps*	*B2cl.abab.ps*	*B2cl.b.ps*

FIGURE 3. The map cl_c

EXAMPLE 3.4. By way of contrast with Example 2.3, we offer no characterization of the map cl_c on permutations satisfying (A), even in the 231-avoiding case. Such a characterization is not immediately apparent, due to the dependence of $\mathrm{cl}_c(w)$ on a specific choice of reduced word for w.

REMARK 3.5. Even for infinite W, the map cl_c associates to each c-sortable element an n-element subset of $T_{\geq -1}$. However, for infinite W, it is not even clear how c-compatibility should be defined, and in particular the proofs in this section apply to the finite case only. As mentioned in the proof of Proposition 3.1, Theorem 3.2 implies the following characterization of c-compatibility: Distinct elements t_1 and t_2 of $T_{\geq -1}$ are c-compatible if and only if there exists a c-sortable element w such that $t_1, t_2 \in \mathrm{cl}_c(w)$. Thus the map cl_c itself might conceivably provide some insight into compatibility in the infinite case.

4. Enumeration

In this section we briefly discuss the enumeration of sortable elements. The *W-Catalan number* is given by the following formula, in which h is the Coxeter number of W and the e_i are the exponents of W.

$$\mathrm{Cat}(W) = \prod_{i=1}^{n} \frac{e_i + h + 1}{e_i + 1}.$$

The values of the W-Catalan number for finite irreducible W are tabulated below.

A_n	B_n	D_n	E_6	E_7	E_8	F_4	G_2	H_3	H_4	$I_2(m)$
$\frac{1}{n+2}\binom{2n+2}{n+1}$	$\binom{2n}{n}$	$\frac{3n-2}{n}\binom{2n-2}{n-1}$	833	4160	25080	105	8	32	280	$m+2$

The noncrossing partitions (with respect to any Coxeter element) in an irreducible finite Coxeter group W are counted by the W-Catalan number [**3, 18, 24**]. The c-clusters (for any Coxeter element c) of an irreducible finite Coxeter group W are also counted by $\mathrm{Cat}(W)$. This follows from [**16**, Corollary 4.11] and [**11**, Proposition 3.8] for the crystallographic case, or is proved in any finite case by combining Theorems 2.1 and 3.2. We refer the reader to [**13**, Section 5.1] for a brief account of other objects counted by the W-Catalan number. By Theorem 2.1 or Theorem 3.2 we have the following.

THEOREM 4.1. *For any Coxeter element c of W, the c-sortable elements of W are counted by* $\mathrm{Cat}(W)$.

The *positive W-Catalan number* is the number of positive c-clusters (c-clusters containing no element of $-S$). The following is an immediate corollary of Theorem 3.2.

COROLLARY 4.2. *For any Coxeter element c, the c-sortable elements not contained in any proper standard parabolic subgroup are counted by the positive W-Catalan number.*

The map nc_c also respects this positive W-Catalan enumeration: The map nc_c maps the sortable elements not contained in any proper standard parabolic subgroup to the noncrossing partitions not contained in any proper standard parabolic subgroup.

The W-Narayana numbers count noncrossing partitions by their rank. That is, the kth W-Narayana number is the number of elements of $[1, c]_T$ whose absolute length is k. The following is an immediate corollary of Theorem 2.1.

COROLLARY 4.3. *For any Coxeter element c, the c-sortable elements of W which have exactly k descents are counted by the kth W-Narayana number.*

Remark 4.4. The kth W-Narayana number is also the kth component in the h-vector of the simplicial W-associahedron. Using results from [**20**] and [**22**], one associates a complete fan to c-sortable elements. This fan has the property that any linear extension of the weak order on c-sortable elements is a shelling. In [**23**], David Speyer and the author show that the map cl_c induces a combinatorial isomorphism. Thus as a special case of a general fact explained in the discussion following [**20**, Proposition 3.5], the h-vector of Δ_c has entry h_k equal to the number of c-sortable elements with exactly k descents. This gives an alternate proof of Corollary 4.3 and, by composing bijections, a bijective explanation of why counting noncrossing partitions by rank recovers the h-vector of the W-associahedron.

Acknowledgments

I am grateful to Sara Billey, Sergey Fomin, Jon McCammond, Vic Reiner, John Stembridge, Hugh Thomas and Andrei Zelevinsky for helpful conversations. Thanks also to Stembridge for advice on efficient computations and to an anonymous francophone for providing a much better translation of the abstract than I could have made.

References

[1] C. Athanasiadis, T. Brady and C. Watt, Shellability of noncrossing partition lattices, preprint `math.CO/0503007`, 2005.
[2] I. Bernšteĭn, I. Gel'fand, and V. Ponomarev, Coxeter functors, and Gabriel's theorem. *Uspehi Mat. Nauk* **28** (1973) no. 2(170), 19–33. English translation in *Russian Math. Surveys* **28** (1973), no. 2, 17–32.
[3] D. Bessis, The dual braid monoid, *Ann. Sci. Ecole Norm. Sup.* **36** (2003) 647-683.
[4] S. Billey and T. Braden, Lower bounds for Kazhdan-Lusztig polynomials from patterns, *Transform. Groups* **8** (2003), no. 4, 321–332.
[5] A. Björner and M. Wachs, Shellable nonpure complexes and posets. II, *Trans. Amer. Math. Soc.* **349** (1997) no. 10, 3945–3975.
[6] N. Bourbaki, *Lie groups and Lie algebras, Chapters 4–6*, Springer-Verlag, Berlin, 2002.
[7] T. Brady and C. Watt, A partial order on the orthogonal group, *Comm. Algebra* **30** (2002) no. 8, 3749–3754.
[8] T. Brady and C. Watt, Lattices in finite real reflection groups, *Trans. Amer. Math. Soc.*, to appear.
[9] F. Chapoton, S. Fomin, and A. Zelevinsky, Polytopal realizations of generalized associahedra, *Canad. Math. Bull.* **45** (2002), 537–566.
[10] M. Dyer, Hecke algebras and shellings of Bruhat intervals. *Compositio Math.* **89** (1993), no. 1, 91–115.
[11] S. Fomin and A. Zelevinsky, Y-systems and generalized associahedra, *Ann. of Math.* **158** (2003), 977–1018.
[12] S. Fomin and A. Zelevinsky, Cluster algebras. II. Finite type classification. *Invent. Math.* **154** (2003), no. 1, 63–121.
[13] S. Fomin and N. Reading, Root systems and generalized associahedra, lecture notes for the *IAS/Park City Graduate Summer School in Geometric Combinatorics*, 2004.
[14] J. Humphreys, Reflection Groups and Coxeter Groups, Cambridge Studies in Advanced Mathematics **29**, Cambridge Univ. Press, 1990.
[15] D. Knuth, *The art of computer programming, Volume 1: Fundamental algorithms*, Addison-Wesley, Reading, Mass., second edition, 1973.
[16] R. Marsh, M. Reineke and A. Zelevinsky, Generalized associahedra via quiver representations, *Trans. Amer. Math. Soc.* **355** (2003) no. 10, 4171–4186.
[17] J. McCammond, Noncrossing partitions in surprising locations, *Amer. Math. Monthly*, to appear.
[18] M. Picantin, Explicit presentations for the dual braid monoids, *C. R. Math. Acad. Sci. Paris* **334** (2002), 843–848.
[19] N. Reading, Lattice congruences of the weak order, *Order*, **21** (2004) no. 4, 315–344.
[20] N. Reading, Lattice congruences, fans and Hopf algebras, *J. Combin. Theory Ser. A* **110** (2005) no. 2, 237–273.
[21] N. Reading, Cambrian Lattices, *Adv. Math.*, to appear.
[22] N. Reading, Sortable elements and Cambrian lattices, preprint, 2005 (`math.CO/0512339`).
[23] N. Reading and D. Speyer, Cambrian Fans, in preparation.
[24] V. Reiner, Non-crossing partitions for classical reflection groups, *Discrete Math.* **177** (1997), 195–222.
[25] J. Shi, The enumeration of Coxeter elements, *J. Algebraic Combin.* **6** (1997), no. 2, 161–171.
[26] C. Ingalls and H. Thomas, Generalized Catalan phenomena via representation theory of quivers, in preparation.

Mathematics Department, University of Michigan, Ann Arbor, MI 48109-1043, USA
E-mail address: `nreading@umich.edu`
URL: `http://www.math.lsa.umich.edu/~nreading/`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Coincidences among skew Schur functions

Victor Reiner, Kristin M. Shaw, and Stephanie van Willigenburg

Abstract. We define an equivalence relation on skew diagrams such that two skew diagrams are equivalent if and only if they give rise to equal skew Schur functions. Then we derive some necessary and sufficient conditions for equivalence.

Résumé. Nous étudions quand deux fonctions de skew Schur sont égales. Avec plus précision, nous derivons quelques règles nécessaires et quelques règles suffisantes pour l'égalité.

1. Introduction

Schur functions are ubiquitous in algebraic combinatorics. They have recently been connected to branching rules for classical Lie groups [**8, 11**], and eigenvalues and singular values of sums of Hermitian and of complex matrices [**1, 5, 8**] via the study of *inequalities* among products of skew Schur functions.

With this in mind, a natural problem is to describe all *equalities* among products of skew Schur functions, or equivalently, to describe all *binomial syzygies* among skew Schur functions. As we shall see in Section 2 this is equivalent to describing all equalities among individual skew Schur functions indexed by *connected* skew diagrams. This is a more tractable instance of a problem that currently seems intractable: describe *all* syzygies among skew Schur functions. Famous non-binomial syzygies include various formulations of the Littlewood-Richardson rule, which give some indication of the complexity that any such description would involve.

The study of equalities among skew Schur functions can also be regarded as part of the calculus of shapes. For an arbitrary subset D of $\mathbb{Z}^2$, there are polynomial representations $\mathcal{S}^D$ and $\mathcal{W}^D$ of $GL_N(\mathbb{C})$ known as a *Schur* or *Weyl modules* respectively. These $GL_N(\mathbb{C})$-representations are obtained by row-symmetrizing and column-antisymmetrizing tensors whose tensor positions are indexed by the cells of D. In general, these representations have $GL_N(\mathbb{C})$-character equal to a symmetric function $s_D(x_1, \ldots, x_N)$; when D is a skew diagram, this is a skew Schur function. Therefore, the question of when two skew Schur or Weyl modules are equivalent in characteristic zero is precisely the question of equalities among skew Schur functions.

Thus we wish to study the following equivalence relation.

Definition 1.1. Given two skew diagrams D_1 and D_2 we say they are *skew-equivalent* denoted $D_1 \sim D_2$ if and only if $s_{D_1} = s_{D_2}$.

For the sake of brevity, in this abstract we assume that the reader is familiar with the basic tenets of algebraic combinatorics such as skew diagrams and Schur functions. If this is not the case, then we refer them to the excellent texts [**9, 12, 13**]*, whose lead we follow by using english notation throughout. One further indispensable tool for us will be the more recent Hamel-Goulden determinant, expressing a skew Schur function s_D, for a skew diagram D, in terms of a determinant based on an outside decomposition of D, and the cutting strip associated to the decomposition; see* [**4, 7**] *for further details.*

2000 *Mathematics Subject Classification.* Primary 05E05, 20C30.

Key words and phrases. Symmetric function, skew Schur function, ribbon Schur function, Weyl module.

The first author was supported by NSF grant DMS-0245379. The second and third authors were supported in part by the National Sciences and Engineering Research Council of Canada.

2. Reduction to connected diagrams

We begin by explaining two easy reductions:

A. Understanding all binomial syzygies among the skew Schur functions is equivalent to understanding the equivalence relation $\sim$ on all skew diagrams, and

B. the latter is equivalent to understanding $\sim$ among *connected* skew diagrams.

These reductions follow from some observations about the matrix

$$JT(\lambda/\mu) := (h_{\lambda_i-\mu_j-i+j})_{i,j=1}^{\ell(\lambda)},$$

which appears in the Jacobi-Trudi formula for a skew diagram λ/μ. We collect these observations in the following proposition, whose straightforward proof is omitted in this abstract.

PROPOSITION 2.1. *Let λ/μ be a skew diagram with $\ell := \ell(\lambda)$.*

(i) The largest subscript k occurring on any nonzero entry h_k in the Jacobi-Trudi matrix $JT(\lambda/\mu)$ is

$$L := \lambda_1 + \ell - 1$$

and this subscript occurs exactly once, on the $(1, \ell)$-entry h_L.

(ii) The subscripts on the diagonal entries in $JT(\lambda/\mu)$ are exactly the row lengths

$$(r_1, \ldots, r_\ell) := (\lambda_1 - \mu_1, \ldots, \lambda_\ell - \mu_\ell)$$

and the monomial $h_{r_1} \cdots h_{r_\ell}$ occurs in the determinant s_D

(a) with coefficient $+1$, and

(b) as the monomial whose subscripts rearranged into weakly decreasing order give the smallest partition of $|\lambda/\mu|$ in dominance order among all nonzero monomials.

(iii) The subscripts on the nonzero subdiagonal entries in $JT(\lambda/\mu)$ are exactly one less than the adjacent row overlap lengths:

$$(\lambda_2 - \mu_1, \lambda_3 - \mu_2, \ldots, \lambda_\ell - \mu_{\ell-1}).$$

COROLLARY 2.1. *For a disconnected skew diagram $D = D_1 \oplus D_2$, one has the factorization $s_D = s_{D_1} s_{D_2}$. For a connected skew diagram D, the polynomial s_D is irreducible in $\mathbb{Z}[h_1, h_2, \ldots]$.*

PROOF. (sketch) The first assertion of the proposition is well-known, and follows, for example, immediately from the definition of skew Schur functions using tableaux.

For the second assertion, let $D = \lambda/\mu$ with $\ell := \ell(\lambda)$ and $L := \lambda_1 + \ell - 1$. Then the Jacobi-Trudi formula and Proposition 2.1(i) imply that the expansion of s_D as a polynomial in the h_r is of the form

$$s \cdot h_L + r \tag{2.1}$$

where s, r are polynomials containing no occurrences of h_L. Proposition 2.1(ii) implies that r is not the zero polynomial, and hence if one can show that s is also nonzero, Equation (2.1) would exhibit s_D as a linear polynomial in h_L with nonzero constant term, and hence clearly irreducible in $\mathbb{Z}[h_1, h_2, \ldots]$. The latter is argued using the fact that D is connected, so that its adjacent row overlaps are all positive, along with Proposition 2.1(iii). $\square$

We can now infer reductions A and B from the beginning of the section. Given a binomial syzygy

$$c\ s_{D_1} s_{D_2} \cdots s_{D_m} - c'\ s_{D'_1} s_{D'_2} \cdots s_{D'_m} = 0$$

among the skew Schur functions, one can rewrite this as $c\, s_D = c' s_{D'}$, where

$$D := D_1 \oplus D_2 \oplus \cdots \oplus D_m$$
$$D' := D'_1 \oplus D'_2 \oplus \cdots \oplus D'_m.$$

Then Proposition 2.1 (ii) implies the unitriangular expansion $s_D = h_\rho + \sum_{\mu:\mu >_{dom} \rho} c_\mu h_\mu$ in which ρ is the weakly decreasing rearrangement of the row lengths in D. This forces $c = c'$ and hence $s_D = s_{D'}$, achieving reduction A.

For reduction B, use the fact that $\Lambda = \mathbb{Z}[h_1, h_2, \ldots]$ is a unique factorization domain, along with Corollary 2.1.

3. Sufficient conditions

The most basic skew-equivalence is the following well-known fact.

Proposition 3.1. [**13**, Exercise 7.56(a)] *If D is a skew diagram then $D \sim D^*$, where D^* is the antipodal rotation of D.*

It transpires that there are several other constructions and operations on skew diagrams that give rise to more skew-equivalences.

3.1. Composition with ribbons. Recall the subset of skew diagrams that contain no 2×2 subdiagram, often known as *ribbons*. In this subsection we generalize in two different ways the composition operation $\alpha \circ \beta$ on ribbons α, β that was introduced in [**2**].

Given two skew diagrams D_1, D_2, aside from their disjoint sum $D_1 \oplus D_2$, there are two closely related important operations called their *concatentation* $D_1 \cdot D_2$ and their *near-concatenation* $D_1 \odot D_2$. The concatentation $D_1 \cdot D_2$ (resp. near concatentation $D_1 \odot D_2$) is obtained from the disjoint sum $D_1 \oplus D_2$ by moving all cells of D_2 one column west (resp. one row south), so that the same column (resp. row) is occupied by the rightmost column (resp. topmost row) of D_1 and the leftmost column (resp. bottommost row) of D_2. For example, if $D_1 = (2,2), D_2 = (3,2)/(1)$ then

$$D_1 \oplus D_2 = \begin{array}{ccccc} & & & 2 & 2 \\ & & 2 & 2 & \\ 1 & 1 & & & \\ 1 & 1 & & & \end{array} \qquad D_1 \cdot D_2 = \begin{array}{cccc} & & 2 & 2 \\ & 2 & 2 & \\ 1 & 1 & & \\ 1 & 1 & & \end{array} \qquad D_1 \odot D_2 = \begin{array}{ccccc} & & & 2 & 2 \\ 1 & 1 & 2 & 2 & \\ 1 & 1 & & & \end{array}.$$

Observe we have used the numbers 1 and 2 to distinguish between those cells in D_1 and those cells in D_2. The reason for the names "concatentation" and "near-concatentation" becomes clearer when we restrict to ribbons. Observe that in this case there exists a natural correspondence that identifies a composition $\alpha = \alpha_1 \ldots \alpha_k$ with the ribbon whose row lengths are $\alpha_1, \ldots, \alpha_k$ read from the bottom. Hence, if we identify ribbons with compositions via this natural correspondence, to get $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ and $\beta = (\beta_1, \ldots, \beta_m)$, we have

$$\begin{aligned} \alpha \cdot \beta &= (\alpha_1, \ldots, \alpha_\ell, \beta_1, \ldots, \beta_m) \\ \alpha \odot \beta &= (\alpha_1, \ldots, \alpha_{\ell-1}, \alpha_\ell + \beta_1, \beta_2, \ldots, \beta_m), \end{aligned}$$

which are the definitions for concatenation and near concatenation given in [**6**].

Note that the operations $\cdot$ and $\odot$ are each associative, and associate with each other:

$$\begin{aligned} (D_1 \cdot D_2) \cdot D_3 &= D_1 \cdot (D_2 \cdot D_3) \\ (D_1 \odot D_2) \odot D_3 &= D_1 \odot (D_2 \odot D_3) \\ (D_1 \odot D_2) \cdot D_3 &= D_1 \odot (D_2 \cdot D_3) \\ (D_1 \cdot D_2) \odot D_3 &= D_1 \cdot (D_2 \odot D_3) \end{aligned}. \tag{3.1}$$

Consequently a string of operations $D_1 \star_1 D_2 \star_2 \cdots \star_{k-1} D_k$ in which each $\star_i$ is either $\cdot$ or $\odot$ is well-defined without any parenethesization. Also note that ribbons are exactly the skew diagrams that can be written uniquely as a string of the form

$$\alpha = \square \star_1 \square \star_2 \cdots \star_{k-1} \square \tag{3.2}$$

where $\square$ is the Ferrers diagram with exactly one cell.

Given a ribbon α and a skew diagram D, define $\alpha \circ D$ to be the result of replacing each cell $\square$ in the expression (3.2) for α with D:

$$\alpha \circ D := D \star_1 D \star_2 \cdots \star_{k-1} D.$$

For example, if

$$\alpha = \begin{array}{cccc} & & & \times \\ & \times & \times & \times \\ \times & \times & & \end{array} \quad \text{and} \quad D = \begin{array}{cc} \times & \times \\ \times & \times \end{array}$$

then

$$\alpha = \square \odot \square \cdot \square \odot \square \odot \square \cdot \square$$
$$\alpha \circ D = D \odot D \cdot D \odot D \odot D \cdot D$$

$$= \quad \begin{array}{cccccccccc}
 & & & & & & & & 6 & 6 \\
 & & & & & & & & 6 & 6 \\
 & & & & & & & 5 & 5 & \\
 & & & & & 4 & 4 & 5 & 5 & \\
 & & & 3 & 3 & 4 & 4 & & & \\
 & & & 3 & 3 & & & & & \\
 & & 2 & 2 & & & & & & \\
1 & 1 & 2 & 2 & & & & & & \\
1 & 1 & & & & & & & &
\end{array}$$

where we have used numbers to distinguish between copies of D.

It is easily seen that when $D = \beta$ is a ribbon, then $\alpha \circ \beta$ is also a ribbon, and agrees with the definition in [**2**].

Similarly, given a skew diagram D and a ribbon β, we can also define $D \circ \beta$ as follows. Create a copy $\beta^{(i)}$ of the ribbon β for each of the cells of D, numbered $i = 1, 2, \ldots, n$ arbitrarily. Then assemble the ribbons $\beta^{(i)}$ into a disjoint decomposition of $D \circ \beta$ by translating them in the plane in such a way that $\beta^{(i)} \sqcup \beta^{(j)}$ forms a copy of

$$\begin{cases} \beta^{(i)} \odot \beta^{(j)} & \text{if } i \text{ is just left of } j \text{ in some row of } D, \\ \beta^{(i)} \cdot \beta^{(j)} & \text{if } i \text{ is just below } j \text{ in some column of } D. \end{cases}$$

For example, if

$$D = \begin{array}{ccc} & 1 & 2 \\ 3 & 4 & 5 \end{array}, \qquad \beta = \begin{array}{cccc} & \times & \times & \times \\ \times & \times & & \end{array}$$

then $D \circ \beta$ is the skew diagram

$$\begin{array}{ccccccccccccccc}
 & & & & & & & & & & & & 2 & 2 & 2 \\
 & & & & & & & & 1 & 1 & 1 & 2 & 2 & & \\
 & & & & & & & 1 & 1 & 5 & 5 & 5 & & & \\
 & & & & & 4 & 4 & 4 & 5 & 5 & & & & & \\
 & 3 & 3 & 3 & 4 & 4 & & & & & & & & & \\
3 & 3 & & & & & & & & & & & & &
\end{array}$$

where we have used numbers to distinguish between copies of β.

Again it is clear that when $D = \alpha$ is a ribbon, then $\alpha \circ \beta$ is another ribbon agreeing with that in [**2**]. The following distributivity properties should also be clear.

Proposition 3.2. *For a skew diagram D and ribbons α and β the operation $\circ$ distributes over $\cdot$ and $\odot$, that is*

$$(\alpha \cdot \beta) \circ D = (\alpha \circ D) \cdot (\beta \circ D)$$
$$(\alpha \odot \beta) \circ D = (\alpha \circ D) \odot (\beta \circ D)$$

and

$$(D_1 \cdot D_2) \circ \beta = (D_1 \circ \beta) \cdot (D_2 \circ \beta)$$
$$(D_1 \odot D_2) \circ \beta = (D_1 \circ \beta) \odot (D_2 \circ \beta).$$

Remark 3.1. Observe that $D_1 \circ D_2$ has not been defined for both D_1 and D_2 being non-ribbons, and we invite the reader to investigate this situation in order to appreciate the complexities that can arise.

In the meantime, we show that the notation for the operations $\alpha \circ D$ and $D \circ \beta$ is consistent with the notation for skew Schur functions. These operations also lead to nontrivial skew-equivalences, generalizing the constructions of [**2**].

We begin by reviewing the presentation of the ring Λ of symmetric functions by the generating set of ribbon Schur functions s_α, that is, those skew Schur functions indexed by ribbons. Let $\mathcal{Q}[z_\alpha]$ denote a polynomial algebra in infinitely many variables z_α indexed by all compositions α.

Proposition 3.3. [**2**, Proposition 2.2]. *The algebra homomorphism*

$$\begin{array}{ccc} \mathcal{Q}[z_\alpha] & \to & \Lambda \\ z_\alpha & \mapsto & s_\alpha \end{array}$$

is a surjection, whose kernel is the ideal generated by the relations

$$z_\alpha z_\beta - (z_{\alpha\cdot\beta} + z_{\alpha\odot\beta}). \tag{3.3}$$

In fact, this same syzygy is well-known to be satisfied [**9**, Chapter 1.5, Example 21 part (a)] by *all* skew diagrams D_1, D_2:

$$s_{D_1} s_{D_2} = s_{D_1\cdot D_2} + s_{D_1\odot D_2}. \tag{3.4}$$

As a consequence, one deduces the following.

Corollary 3.2. *For a fixed skew diagram D the map*

$$\begin{array}{ccc} \mathcal{Q}[z_\alpha] & \xrightarrow{(-)\circ s_D} & \Lambda \\ z_\alpha & \longmapsto & s_{\alpha\circ D} \end{array}$$

descends to a well-defined map $\Lambda \longrightarrow \Lambda$. In other words, for any symmetric function f, one can arbitrarily write f as a polynomial in ribbon Schur functions $f = p(s_\alpha)$ and then set $f \circ s_D := p(s_{\alpha\circ D})$.

We are abusing notation here by using $\circ$ both for the map $(-)\circ s_D$ on symmetric functions, as well as the two diagrammatic operations $\alpha \circ D$ and $D \circ \beta$. The previous corollary says that it is well-defined to set

$$s_\alpha \circ s_D = s_{\alpha\circ D} \tag{3.5}$$

so that we are at least consistent with one of the diagrammatic operations. The next result says that we are also consistent with the other.

Proposition 3.4. *For any skew diagram D and ribbon β*

$$s_{D\circ\beta} = s_D \circ s_\beta.$$

Proof. (sketch) One uses the Hamel-Goulden determinant for s_D, which starts with an outside decomposition of D into ribbons $(\theta_1, \dots, \theta_m)$. The induced outside decomposition $(\theta_1 \circ \beta, \dots, \theta_m \circ \beta)$ for $D \circ \beta$ leads to a Hamel-Goulden determinant for $s_{D\circ\beta}$. The proposition then follows because various operations commute with each other. ☐

Theorem 3.3. *Assume one has ribbons α, α' and skew diagrams D, D' satisfying $\alpha \sim \alpha'$ and $D \sim D'$. Then*

- *(i)* $\alpha \circ D \sim \alpha' \circ D$,
- *(ii)* $D \circ \alpha \sim D' \circ \alpha$,
- *(iii)* $D \circ \alpha \sim D \circ \alpha'$, *and*
- *(iv)* $\alpha \circ D \sim \alpha \circ D^*$.

Proof. Assertions (i) and (ii) both follow from the fact that if E is any skew diagram, then $D \sim D'$ means $s_D = s_{D'}$, and hence

$$s_D \circ s_E = s_{D'} \circ s_E. \tag{3.6}$$

The third follows by Proposition 3.4 if one can show it when $D = \alpha$ is a ribbon. This special case $\alpha\circ\beta_1 \sim \alpha\circ\beta_2$ was shown in [**2**]. Assertion (iv) follows from assertion (i) and Proposition 3.1:

$$\alpha \circ D \sim (\alpha \circ D)^* = \alpha^* \circ D^* \sim \alpha \circ D^*.$$

☐

Remark 3.4. The last skew-equivalence begs the question of whether $D_1 \sim D_2$ for skew diagrams D_1, D_2 implies $\alpha \circ D_1 \sim \alpha \circ D_2$ for any ribbon α. This turns out to be false. For example, one can check that

$$D_1 = \begin{array}{cccc} & \times & \times & \times \\ & \times & \times & \\ \times & \times & & \\ \times & & & \end{array} \quad \sim \quad \begin{array}{cccc} & & \times & \times \\ \times & \times & \times & \\ \times & \times & & \\ \times & & & \end{array} = D_2,$$

e.g. by Corollary 3.20. However, if one takes $\alpha = (2)$, that is, the ribbon having one row with two cells, then we find

$$\alpha \circ D_1 = \begin{array}{cccccccc} & & & & & \times & \times & \times \\ & & & & & \times & \times & \\ & & & & \times & \times & & \\ & \times & \times & \times & \times & & & \\ & \times & \times & & & & & \\ \times & \times & & & & & & \\ \times & & & & & & & \end{array}$$

$$\alpha \circ D_2 = \begin{array}{cccccccc} & & & & & & \times & \times \\ & & & & \times & \times & \times & \\ & & & & \times & \times & & \\ & & \times & \times & \times & & & \\ \times & \times & \times & & & & & \\ \times & \times & & & & & & \\ \times & & & & & & & \end{array}$$

and $\alpha \circ D_1 \not\sim \alpha \circ D_2$, e.g. by Theorem 4.4.

3.2. Amalgamation and amalgamated composition of ribbons. Now in a third way we generalize the operation $\alpha \circ \beta$ to an operation $\alpha \circ_\omega \beta$, which we will call the *amalgamated composition* of α and β with respect to ω.

Definition 3.5.
Given a skew diagram D and a nonempty ribbon ω, say that ω *lies in the top (resp. bottom) of* D if the restriction of D to its $|\omega|$ northeasternmost (resp. southwesternmost) diagonals is (a translated copy of) the ribbon ω.

Given two skew diagrams D_1, D_2 and a nonempty ribbon ω lying in the top of D_1 and the bottom of D_2, the *amalgamation of* D_1 *and* D_2 *along* ω, denoted $D_1 \amalg_\omega D_2$, is the new ribbon obtained from the disjoint union $D_1 \oplus D_2$ by identifying the copy of ω in the northeast of D_1 with the copy of ω in the southwest of D_2.

Say that ω *protrudes from the top (resp. bottom) of* D if there is another ribbon ω^+ having $|\omega^+| = |\omega|+1$ such that both ω, ω^+ lie at the top (resp. bottom) of D. Equivlalently, ω protrudes from the top (resp. bottom) of D if it lies at the top (resp. bottom) of D and the restriction of D to its $|\omega|+1$ northeasternmost (resp. southwesternmost) diagonals is also a ribbon, namely ω^+.

Example 3.6.
Consider the skew diagram

$$D = \begin{array}{cccc} & \times & \times & \times \\ \times & \times & \times & \end{array}.$$

Then D has $\omega_1 = \times$ protruding from the top and bottom. It has $\omega_2 = \times\times$ lying in its top and bottom, but protruding from neither top nor bottom. Furthermore,

$$D \amalg_{\omega_1} D = \begin{array}{ccccccc} & & & & \times & \times & \times \\ & \times & \times & o & \times & \times & \\ \times & \times & \times & & & & \end{array}, \qquad D \amalg_{\omega_2} D = \begin{array}{cccccc} & & & \times & \times & \times \\ & \times & o & o & \times & \\ \times & \times & \times & & & \end{array}$$

in which the copies of ω_1 and ω_2 that have been amalgamated are indicated with the letter o.

Definition 3.7.
When ω lies in the top of D_1 and bottom of D_2, one can form the *outer (resp. inner) projection of* D_1 *onto* D_2 *with respect to* ω. This is a new diagram in the plane, not necessarily skew, obtained from the disjoint union $D_1 \oplus D_2$ by translating D_2 until it is underneath and to the right (resp. above and to the left) of D_1, in such a way that the two copies of ω in D_1, D_2 are adjacent and occupy the same set of diagonals.

One can see that if ω not only lies in the top of D_1 and bottom of D_2, but actually *protrudes* from the top of D_1 and from the bottom of D_2, then at most one of these two projections can be a skew diagram (and possibly neither one is). When one of them is a skew diagram, call it $D_1 \cdot_\omega D_2$, and say that $D_1 \cdot_\omega D_2$ *is defined* in this case.

Example 3.8.
Let D, ω_1, ω_2 be as in the previous example. Then the outer and inner projections of D onto D with respect to ω_2 are

$$\begin{array}{cccccc} & \times & o & o & \times & \times & \times \\ \times & \times & \times & o & o & \times & \end{array} \qquad \begin{array}{ccccc} & & \times & \times & \times \\ & o & o & \times & \\ & \times & o & o & \\ \times & \times & \times & & \end{array} ,$$

which are both skew diagrams. On the other hand, the outer and inner projections of D onto D with respect to ω_1 are

$$\begin{array}{cccccccc} & \times & \times & o & & \times & \times & \times \\ \times & \times & \times & & o & \times & \times & \end{array} \qquad \begin{array}{cccccc} & & & \times & \times & \times \\ & & o & \times & \times & \\ & \times & \times & o & & \\ \times & \times & \times & & & \end{array} = D \cdot_{\omega_1} D$$

in which only the latter is a skew diagram.

Definition 3.9.
Given a skew diagram D, and ω a ribbon lying in both the top and bottom of D, one can define

$$D^{\amalg_\omega n} = \underbrace{D \amalg_\omega D \amalg_\omega \cdots \amalg_\omega D}_{n \text{ factors}} := ((D \amalg_\omega D) \amalg_\omega D) \amalg_\omega \cdots \amalg_\omega D.$$

If one assumes that $D \cdot_\omega D$ is also defined so that, in particular, ω protrudes from the top and bottom of D, then one can check that this will imply that for any positive integers m, n, we have $(D^{\amalg_\omega m}) \cdot_\omega (D^{\amalg_\omega n})$ is also defined. Under this assumption, for any ribbon $\alpha = (\alpha_1, \ldots, \alpha_k)$, define the *amalgamated composition* of α and D with respect to ω to be the diagram

$$\alpha \circ_\omega D := (D^{\amalg_\omega \alpha_1}) \cdot_\omega \ldots \cdot_\omega (D^{\amalg_\omega \alpha_k}). \tag{3.7}$$

The following theorems are obtained by consideration of an appropriate Hamel-Goulden determinant.

Theorem 3.10. *Let D be a connected skew diagram, and ω a ribbon which protrudes from the top and bottom of D, with $D \cdot_\omega D$ defined. Assume further that the two copies of ω in the top and bottom of D are separated by at least one diagonal,that is, there is a nonempty diagonal in D intersecting neither copy of ω.*
Then for any ribbon α one has

$$s_{\alpha \circ_\omega D} = s_\alpha \circ_\omega s_D.$$

Theorem 3.11. *Let α, α' be ribbons with $\alpha \sim \alpha'$, and assume that D, ω satisfy the hypotheses of Theorem 3.10. Then one has the following skew-equivalences:*

$$\alpha' \circ_\omega D \ \sim \ \alpha \circ_\omega D \ \sim \ \alpha \circ_{\omega^*} D^*.$$

Theorem 3.12. *Let $\{\beta_i\}_{i=1}^k, \{\gamma_i\}_{i=1}^k$ be ribbons, and for each i either $\gamma_i = \beta_i$ or $\gamma_i = \beta_i^*$. If the skew diagrams D, ω satisfy the hypotheses of Theorem 3.10, then*

$$\begin{aligned} &\gamma_1 \circ_\omega \gamma_2 \circ_\omega \ldots \circ_\omega \gamma_k \circ_\omega D \\ &\quad \sim \beta_1 \circ_\omega \beta_2 \circ_\omega \ldots \circ_\omega \beta_k \circ_\omega D \\ &\quad\quad \sim \beta_1 \circ_{\omega^*} \beta_2 \circ_{\omega^*} \ldots \circ_{\omega^*} \beta_k \circ_{\omega^*} D^* \end{aligned}$$

where all the operations $\circ_\omega$ or $\circ_{\omega^*}$ are performed from right to left.

Remark 3.13. Theorem 3.11 is analogous to [**2**, Theorem 4.4 parts 1 and 2], whereas Theorem 3.12 is analogous to the reverse direction of [**2**, Theorem 4.1].

3.3. Conjugation and ribbon staircases. The goal here is to construct skew diagrams D that are skew-equivalent to their conjugates D^t. We first define two decompositions of a connected skew diagram D into ribbons; when one of these decompositions takes on a very special form, we will show that implies $D \sim D^t$.

Definition 3.14.
Given a connected skew diagram D define the *southeast decomposition* to be the following unique decomposition into ribbons. The first ribbon θ is the unique ribbon that starts at the cell on the lower left, traverses the *southeast* border of D, and ends at the cell on the upper right. Now consider D with θ removed, which

may decompose into several connected component skew shapes, and iterate the above procedure on each of these shapes. The *northwest decomposition* is similarly defined, starting with a ribbon θ that traverses the *northwest* border of D.

Note that both of these are outside decompositions of D, and hence give rise to Hamel-Goulden determinants for s_D. In both cases, the associated cutting strip for the decomposition coincides with its first and largest ribbon θ. We will be interested in the case where all of the ribbons in the southeast or northwest decomposition of D arise in a very special way from the amalgamation construction of Section 3.2.

Definition 3.15.
Let $\alpha = (\alpha_1, \ldots, \alpha_k)$ and $\beta = (\beta_1, \ldots, \beta_\ell)$ be ribbons. For an integer $m \geq 1$, say that the *m-intersection $\alpha \cap_m \beta$ exists* if there is a ribbon $\omega = (\omega_1, \ldots, \omega_m)$ with m rows lying in the top of α and the bottom of β for which $\omega_1 = \beta_1$ and $\omega_m = \alpha_k$; when $m = 1$, we set $\omega_1 := \min\{\alpha_k, \beta_1\}$. In this case, define the *m-intersection* $\alpha \cap_m \beta$ and the *m-union* $\alpha \cup_m \beta$ to be

$$\begin{aligned} \alpha \cap_m \beta &:= \omega \\ \alpha \cup_m \beta &:= \alpha \amalg_\omega \beta \end{aligned}$$

If $\alpha \cup_m \beta = \alpha$ or β (resp. or $\alpha \cap_m \beta = \alpha$ or β) then we say the m-union (resp. m-intersection) is *trivial.* If α is a ribbon such that $\alpha \cap_m \alpha$ exists and is non-trivial then

$$\varepsilon_m^k(\alpha) := \underbrace{\alpha \cup_m \alpha \cup_m \ldots \cup_m \alpha}_{k \text{ factors}}$$

is the *ribbon staircase* of height k and depth m generated by α.

Example 3.16.
Let α be the ribbon $(2, 3)$. Then

$$\varepsilon_1^3(\alpha) = \varepsilon_1^3 \left(\begin{array}{cccc} & \times & \times & \times \\ \times & \times & & \end{array} \right) = \begin{array}{cccccccc} & & & & & \times & \times & \times \\ & & & \times & \times & \times & & \\ & \times & \times & \times & & & & \\ \times & \times & & & & & & \end{array} .$$

Definition 3.17.
Say that a skew diagram D has a *southeast ribbon staircase decomposition* if there exists an $m < \ell(\alpha)$ and a ribbon α such that all ribbons in the southeast decomposition of D are of the forms $\alpha \cap_m \alpha$ or $\varepsilon_m^p(\alpha)$ for various integers $p \geq 1$.

In this situation, let k be the maximum value of p occurring among the $\varepsilon_m^p(\alpha)$ above, so that the largest ribbon θ equals $\varepsilon_m^k(\alpha)$. We will think of θ as containing k copies of α, numbered $1, 2, \ldots, k$ from southwest to northeast. We now wish to define the *nesting* $\mathcal{N}$ associated to this decomposition. The nesting $\mathcal{N}$ is a word of length $k-1$ using as letters the four symbols, *dot* ".", *left parenthesis* "(", *right parenthesis* ")" and *vertical slash* "|". Considering the ribbons in the southeast decomposition of D,

- a ribbon of the form $\varepsilon_m^p(\alpha)$ creates a pair of left and right parentheses in positions i and j if the ribbon occupies the same diagonals as the $i+1, i+2, \ldots, j-1, j$ copies of α in θ, while
- a ribbon of the form $\alpha \cap_m \alpha$ creates a vertice slash in position i if it occupies the same diagonals as the intersection of the $i, i+1$ copies of α in θ, and
- all other letters in $\mathcal{N}$ are dots.

With this notation, say that $D = (\varepsilon_m^k(\alpha), \mathcal{N})_{se}$. Analogously define the notation $D = (\varepsilon_m^k(\alpha), \mathcal{N})_{nw}$ using the northwest decomposition.

Lastly, given a nesting $\mathcal{N}$, denote the *reverse nesting*, which is the reverse of the word $\mathcal{N}$, by $\mathcal{N}^*$.

Example 3.18. Consider the following skew diagram D, with its southeast decomposition into ribbons $\theta_1, \theta_2, \theta_3, \theta_4$ distinguished by the numbers $1, 2, 3, 4$ respectively:

$$D = \begin{array}{ccccccc} & & 4 & 3 & 3 & 1 & 1 \\ & & 3 & 3 & 1 & 1 & \\ & & 3 & 1 & 1 & & \\ & 2 & 1 & 1 & & & \\ & 1 & 1 & & & & \\ 1 & 1 & & & & & \\ 1 & & & & & & \end{array} .$$

This happens to be a southeast ribbon staircase decomposition, in which

$$\alpha = \begin{array}{cc} \times & \times \\ \times & \end{array}, \quad m = 1, \quad k = 6, \quad \mathcal{N} = .\,|\,(\,|\,),$$

that is, $D = (\varepsilon_1^6(\alpha), \mathcal{N})_{se}$. Here $\mathcal{N}^* = (\,|\,)\,|\,.$, and additionally note that the skew diagram $D' = (\varepsilon_1^6(\alpha), \mathcal{N}^*)$ is the same as the conjugate skew diagram D^t.

The following is a consequence of the Hamel-Goulden determinant associated to the southeast (or northwest) decomposition of D, when its decomposition is a ribbon staircase decomposition.

Theorem 3.19. *Let α be a ribbon, and let*

$$D_1 = (\varepsilon_m^k(\alpha), \mathcal{N})_x$$
$$D_2 = (\varepsilon_m^k(\alpha), \mathcal{N}^*)_x$$

where $m < \ell(\alpha)$ and $x = se$ or nw. Then $D_1 \sim D_2$.

This leads to the following interesting corollary.

Corollary 3.20. *Let $D = (\varepsilon_m^k(\alpha), \mathcal{N})_x$ where α is a self-conjugate ribbon, $m < l(\alpha)$ and $x = se$ or nw. Then $D \sim D^t$. Furthermore, for any Ferrers diagram μ contained in the staircase partition $\delta_n := (n-1, n-2, \ldots, 1) \vdash \binom{n}{2}$, one has*

$$\delta_n/\mu \sim (\delta_n/\mu)^t .$$

We conjecture the following converse, which has been verified for all skew diagrams D with $|D| \leq 18$.

Conjecture 3.21. *If a skew diagram D satisfies $D \sim D^t$, then $D = (\varepsilon_m^k(\alpha), \mathcal{N})_x$ for some self-conjugate ribbon α, some $m < \ell(\alpha)$ and $x = se$ or nw.*

3.4. Adding a full column/row, and complementation within a rectangle. Let D be thought of as any finite subset of the plane $\mathbb{Z}^2$. We wish to consider two operations on D, which turn out to be closely related.

- *Adding a full column (resp. row)*: Add to the shape a new column (resp. row) which has a cell in every previously nonempty row (resp. column), and possibly in some new rows (resp. columns).
- *Complementation within a rectangle*: If R is a rectangular Ferrers diagram containing D, consider the complementary shape $R \backslash D$.

When D is a Ferrers diagram λ, it is not hard to see that the result of the latter is at least a skew diagram. However, when D is only assumed to be a *skew* diagram, after performing either of these operations, it is generally *not* true that the result is another skew diagram. Nevertheless, in some cases, after performing these operations, one may be able to reorder the columns (resp. rows) so as to obtain a skew diagram again. This combined with the following definition allows us to derive another sufficiency.

Definition 3.22.
A skew diagram D has *spinal columns* if it contains either a single column or a union of two adjacent columns whose union intersects every nonempty row of D. One can similarly define when D has *spinal rows.*

The following can be proved using results from [**10**].

Theorem 3.23. *Let D_i for $i = 1, 2$ be skew diagrams, both having spinal columns, and ℓ nonempty rows. Let R be a rectangle with ℓ rows that contains D_1, D_2. Let D_i^+ be obtained from D_i by adding a full column of length ℓ to form a skew diagram. Then*

$$D_1 \sim D_2 \text{ if and only if } D_1^+ \sim D_2^+ \text{ if and only if } R \backslash D_1 \sim R \backslash D_2.$$

4. Necessary conditions

We now present some combinatorial invariants for the skew-equivalence relation $D_1 \sim D_2$ on connected skew diagrams.

4.1. Frobenius rank. Recall that the *Durfee* or *Frobenius rank* of a skew diagram D is defined to be the minimum number of ribbons in any decomposition of D into ribbons. It was recently conjectured by Stanley [**14**], and proven by Chen and Yang [**3**], that the rank coincides with the highest power of t dividing the polynomial $s_D(1, 1, \ldots, 1, 0, 0, \ldots)$, where t of the variables have been set to 1, and the rest to zero. This implies the following.

Corollary 4.1. *Frobenius rank is an invariant of skew-equivalence, that is, two skew-equivalent diagrams must have the same Frobenius rank.*

In particular, skew-equivalence restricts to the subset of ribbons as they are the skew diagrams of Frobenius rank 1.

4.2. Overlaps. Data about the amount of overlap between sets of rows or columns in the skew diagram D can be recovered from its skew Schur function s_D.

Definition 4.2.
Let D be a skew diagram occupying r rows. For each k in $\{1, 2, \ldots, r\}$, define the *k-row overlap composition* $r^{(k)} = (r_1^{(k)}, \ldots, r_{r-k+1}^{(k)})$ to be the sequence where $r_i^{(k)}$ is the number of columns occupied in common by the rows $i, i+1, \cdots, i+k-1$. Let $\rho^{(k)}$ be the *k-row overlap partition* that is the weakly decreasing rearrangement of $r^{(k)}$. Similarly define column overlap compositions $c^{(k)}$ and column overlap partitions $\gamma^{(k)}$.

Example 4.3. If $D = \begin{matrix} & \times & \times \\ \times & \times & \times \\ \times & & \end{matrix}$, then the 1-row, 2-row and 3-row overlap compositions are

$$\begin{aligned} r^{(1)} &= (2, 3, 1) \\ r^{(2)} &= (2, 1) \\ r^{(3)} &= (0). \end{aligned}$$

With this is mind we are able to prove

Theorem 4.4. *If $D_1 \sim D_2$ then D_1, D_2 have the same k-row overlap partitions and the same k-column overlap partitions for all k.*

It transpires that the row overlap partitions $(\rho^{(k)})_{k \geq 1}$ and the column overlap partitions $(\gamma^{(k)})_{k \geq 1}$ determine each other uniquely. To see this, we define a third form of data on a skew diagram D, which mediates between the two, and which is more symmetric under conjugation.

Proposition 4.1. *Given a skew diagram D, consider the doubly-indexed array $(a_{k,\ell})_{k,\ell \geq 1}$ where $a_{k,\ell}$ is defined to be the number of $k \times \ell$ rectangular subdiagrams contained inside D. Then we have*

$$\begin{aligned} a_{k,\ell} &= \sum_{\ell' \geq \ell} \left(\rho^{(k)}\right)^t_{\ell'} \\ &= \sum_{k' \geq k} \left(\gamma^{(\ell)}\right)^t_{k'}. \end{aligned}$$

Consequently, any one of the three forms of data

$$(\rho^{(k)})_{k \geq 1}, \quad (\gamma^{(k)})_{k \geq 1}, \quad (a_{k,\ell})_{k,\ell \geq 1}$$

on D determines the other two uniquely.

Remark 4.5. Unfortunately, having the same row and column overlap partitions $\rho^{(k)}, \gamma^{(k)}$ is not sufficient for the skew-equivalence of two skew diagrams as

$$\begin{array}{cccc} & & \times & \times \\ \times & \times & \times & \\ \times & & & \end{array} \quad \not\sim \quad \begin{array}{cccc} & \times & \times & \times \\ \times & \times & & \\ \times & & & \end{array}$$

even though they have the same row and column overlap partitions $\rho^{(k)}, \gamma^{(k)}$ for every k.

5. Complete classification

The sufficient conditions discussed in this abstract explain all but six of the skew-equivalences among skew diagrams with up to 18 cells. For example, the following skew-equivalence cannot yet be explained:

$$\begin{array}{cccccc} & & & & \times & \times \\ & & \times & \times & \times & \\ & \times & \times & \times & \times & \\ & \times & \times & & & \\ \times & \times & & & & \\ \times & \times & & & & \end{array} \quad \sim \quad \begin{array}{cccccc} & & & & \times & \times \\ & & & \times & \times & \\ & & & \times & \times & \\ & \times & \times & \times & & \\ \times & \times & \times & \times & & \\ \times & \times & & & & \end{array} .$$

We end with the following conjectures.

Conjecture 5.1. *The skew-equivalence relation $\sim$, when restricted to skew diagrams of Frobenius rank at most 3, is explained by all of the constructions in this paper. In other words, it is the equivalence relation generated by the equivalences listed in*

- *Proposition 3.1,*
- *Theorem 3.3,*
- *Theorem 3.11,*
- *Theorem 3.19, and*
- *Theorem 3.23.*

Conjecture 5.2. *Every skew-equivalence class of skew diagrams has cardinality a power of 2.*

References

[1] F. Bergeron, R. Biagioli, and M.H. Rosas, Inequalities between Littlewood-Richardson coefficients, `math.CO/0403541`.

[2] L.J. Billera, H. Thomas, and S. van Willigenburg, Decomposable compositions, symmetric quasisymmetric functions and equality of ribbon Schur functions, `math.CO/0405434`.

[3] W.Y.C. Chen and A.L.B. Yang, Stanley's Zrank problem on skew partitions, `math.CO/0509181`.

[4] W. Chen, G.-G. Yan, and A. Yang, Transformations of border strips and Schur function determinants, *J. Algebraic Combin.* **21** (2005), 379–394.

[5] S. Fomin, W. Fulton, C.-K. Li, and Y.-T. Poon, Eigenvalues, singular values, and Littlewood-Richardson coefficients, `math.AG/0301307`.

[6] I. M. Gel'fand, D. Krob, A. Lascoux, B. Leclerc, V. Retakh, and J.-Y. Thibon, Noncommutative symmetric functions, *Adv. Math.* **112** (1995), 218–348.

[7] A. Hamel and I. Goulden, Planar decompositions of tableaux and Schur function determinants, *European J. Combin.* **16** (1995), 461–477.

[8] T. Lam, A. Postnikov, and P. Pylyavskyy, Schur positivity and Schur log-concavity, `math.CO/0502446`.

[9] I. Macdonald, *Symmetric Functions and Hall Polynomials, 2nd Edition*, Oxford University Press, New York, USA, 1995.

[10] P. Magyar, Borel-Weil theorem for configuration varieties and Schur modules, *Adv. Math.* **134** (1998), 328–366.

[11] A. Okounkov, Log-concavity of multiplicities with application to characters of $U(\infty)$, *Adv. Math.* **127** (1997), 258-282.

[12] B.E. Sagan, *The symmetric group: Representations, combinatorial algorithms, and symmetric functions (2nd edition), Graduate Texts in Mathematics,* **203**, Springer-Verlag, New York, USA, 2001.

[13] R.P. Stanley, *Enumerative Combinatorics, vol. 2*, Cambridge University Press, Cambridge, UK, 1999.

[14] R.P. Stanley, The rank and minimal border strip decompositions of a skew partition, *J. Combinatorial Theory (A)* **100** (2002), 349–375.

School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
E-mail address: reiner@math.umn.edu

Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
E-mail address: krishaw@math.ubc.ca

Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada
E-mail address: steph@math.ubc.ca

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Virtual crystal structure on rigged configurations

Anne Schilling

ABSTRACT. Rigged configurations are combinatorial objects originating from the Bethe Ansatz, that label highest weight crystal elements. In this note a new *unrestricted* set of rigged configurations is introduced by constructing a crystal structure on the set of rigged configurations.

RÉSUMÉ. Les configurations gréées sont des objets combinatoires inspirés par l'ansatz de Bethe, et qui sont en correspondence avec les éléments cristallins de plus haut poids. Dans cette note, nous introduisons le concept de "configurations gréées généralisées", en construisant une structure cristalline dans l'espace des configurations gréées.

1. Introduction

This note is based on preprint [**33**] which gives a crystal structure on rigged configurations for all simply-laced types. Here we use the virtual crystal method [**29, 30**] to extend these results to nonsimply-laced types.

There are (at least) two main approaches to solvable lattice models and their associated quantum spin chains: the Bethe Ansatz [**6**] and the corner transfer matrix method [**5**].

In his 1931 paper [**6**], Bethe solved the Heisenberg spin chain based on the string hypothesis which asserts that the eigenvalues of the Hamiltonian form certain strings in the complex plane as the size of the system tends to infinity. The Bethe Ansatz has been applied to many models to prove completeness of the Bethe vectors. The eigenvalues and eigenvectors of the Hamiltonian are indexed by rigged configurations. However, numerical studies indicate that the string hypothesis is not always true [**2**].

The corner transfer matrix (CTM) method, introduced by Baxter [**5**], labels the eigenvectors by one-dimensional lattice paths. These lattice paths have a natural interpretation in terms of Kashiwara's crystal base theory [**16, 17**], namely as highest weight crystal elements in a tensor product of finite-dimensional crystals.

Even though neither the Bethe Ansatz nor the corner transfer matrix method are mathematically rigorous, they suggest the existence of a bijection between the two index sets, namely rigged configurations on the one hand and highest weight crystal paths on the other (see Figure 1). For the special case when the spin chain is defined on $V_{(\mu_1)} \otimes V_{(\mu_2)} \otimes \cdots \otimes V_{(\mu_k)}$, where $V_{(\mu_i)}$ is the irreducible $\mathrm{GL}(n)$ representation indexed by the partition (μ_i) for $\mu_i \in \mathbb{N}$, a bijection between rigged configurations and semi-standard Young tableaux was given by Kerov, Kirillov and Reshetikhin [**21, 22**]. This bijection was proven and extended to the case when the (μ_i) are any sequence of rectangles in [**25**]. The bijection has many amazing properties. For example it takes the cocharge statistics cc defined on rigged configurations to the coenergy statistics D defined on crystals.

Rigged configurations and crystal paths also exist for other types. In [**14, 15**] the existence of Kirillov–Reshetikhin crystals $B^{r,s}$ was conjectured, which can be naturally associated with the dominant weight $s\Lambda_r$ where s is a positive integer and Λ_r is the r-th fundamental weight of the underlying algebra of finite type. For a tensor product of Kirillov–Reshetikhin crystals $B = B^{r_k,s_k} \otimes \cdots \otimes B^{r_1,s_1}$ and a dominant weight Λ let $\overline{\mathcal{P}}(B,\Lambda)$ be the set of all highest weight elements of weight Λ in B. In the same papers [**14, 15**], fermionic formulas $\overline{M}(L,\Lambda)$ for the one-dimensional configuration sums $\overline{X}(B,\Lambda) := \sum_{b\in\overline{\mathcal{P}}(B,\Lambda)} q^{D(b)}$ were conjectured. The fermionic formulas admit a combinatorial interpretation in terms of the set of rigged configurations $\overline{\mathrm{RC}}(L,\Lambda)$, where L is the multiplicity array of B. A statistic

2000 *Mathematics Subject Classification.* Primary 17B37, 82B23, 05A15; Secondary 05E99, 81R50.
Key words and phrases. crystal bases, rigged configurations, Kashiwara operators.
Date: November 2005.
Partially supported by NSF grant DMS-0200774 and DMS-0501101.

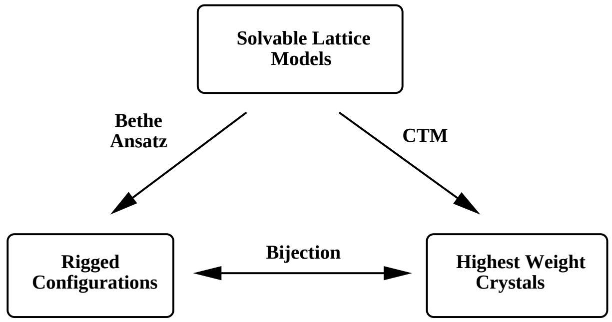


FIGURE 1. Schematic origin of rigged configurations and crystal paths

preserving bijection $\Phi : \overline{\mathcal{P}}(B,\Lambda) \to \overline{\mathrm{RC}}(L,\Lambda)$ has been proven in various cases [**25, 28, 32, 35**] which implies the following identity

$$\overline{X}(B,\Lambda) := \sum_{b\in\overline{\mathcal{P}}(B,\Lambda)} q^{D(b)} = \sum_{(\nu,J)\in\overline{\mathrm{RC}}(L,\Lambda)} q^{\mathrm{cc}(\nu,J)} =: \overline{M}(L,\Lambda). \tag{1.1}$$

Since the sets in (1.1) are finite, these are polynomials in q. When $B = B^{1,s_k}\otimes\cdots\otimes B^{1,s_1}$ of type A, they are none other than the Kostka–Foulkes polynomials.

Rigged configurations corresponding to highest weight crystal paths are only the tip of an iceberg. In this note we extend the definition of rigged configurations to all crystal elements by the explicit construction of a crystal structure on the set of *unrestricted* rigged configurations (see Definition 4.1). For simply-laced types, the proof is given in [**32**] and uses Stembridge's local characterization of simply-laced crystals [**37**]. For nonsimply-laced algebras, we show here how to apply the method of virtual crystals [**29, 30**] to construct the crystal operators on rigged configurations.

The equivalence of the crystal structures on rigged configurations and crystal paths together with the correspondence for highest weight vectors yields the equality of generating functions in analogy to (1.1) (see Theorem 4.10 and Corollary 4.11). Denote the unrestricted set of paths and rigged configurations by $\mathcal{P}(B,\Lambda)$ and $\mathrm{RC}(L,\Lambda)$, respectively. The corresponding generating functions $X(B,\Lambda) = M(L,\Lambda)$ are unrestricted generalized Kostka polynomials or q-supernomial coefficients. A direct bijection $\Phi : \mathcal{P}(B,\Lambda) \to \mathrm{RC}(L,\Lambda)$ for type A along the lines of [**25**] is constructed in [**7, 8**].

Rigged configurations are closely tied to fermionic formulas. Fermionic formulas are explicit expressions for the partition function of the underlying physical model which reflect their particle structure. For more details regarding the background of fermionic formulas see [**14, 19, 20**]. For type A we obtain an explicit characterization of the unrestricted rigged configurations in terms of lower bounds on quantum numbers which yields a new fermionic formula for unrestricted Kostka polynomials of type A. Surprisingly, this formula is different from the fermionic formulas in [**13, 18**] obtained in the special cases of $B = B^{1,s_k}\otimes\cdots\otimes B^{1,s_1}$ and $B = B^{r_k,1}\otimes\cdots\otimes B^{r_1,1}$. The rigged configurations corresponding to the fermionic formulas of [**13, 18**] were related to ribbon tableaux and the cospin generating functions of Lascoux, Leclerc, Thibon [**26, 27**] in reference [**31**]. To distinguish these rigged configurations from the ones introduced in this paper, let us call them ribbon rigged configurations.

The Lascoux–Leclerc–Thibon (LLT) polynomials [**26, 27**] have recently made their debut in the theory of Macdonald polynomials in the seminal paper by Haiman, Haglund, Loehr [**9**]. The main obstacle in obtaining a combinatorial formula for the Macdonald–Kostka polynomials is the Schur positivity of certain LLT polynomials. A related problem is the conjecture of Kirillov and Shimozono [**24**] that the cospin generating function of ribbon tableaux equals the generalized Kostka polynomial. A possible avenue to prove this conjecture would be a direct bijection between the unrestricted rigged configurations of this paper and ribbon rigged configurations.

One of the motivations for considering unrestricted rigged configurations was Takagi's work [**38**] on the inverse scattering transform, which provides a bijection between states in the $\mathfrak{sl}_2$ box ball system and rigged configurations. In this setting rigged configurations play the role of action-angle variables. Box ball systems can be produced from crystals of solvable lattice models for algebras other than $\mathfrak{sl}_2$ [**10, 11, 12**]. The inverse scattering transform can be generalized to the $\mathfrak{sl}_n$ case [**23**], which should give a box-ball interpretation of the unrestricted rigged configurations presented here.

Another motivation for the study of unrestricted configuration sums, fermionic formulas and associated rigged configurations is their appearance in generalizations of the Bailey lemma [**3, 39**]. The Andrews–Bailey construction [**1,**

4] relies on an iterative transformation property of the q-binomial coefficient, which is one of the simplest unrestricted configuration sums, and can be used to prove infinite families of Rogers–Ramanujan type identities. The explicit formulas provided in this paper might trigger further progress towards generalizations to higher-rank or other types of the Andrews–Bailey construction.

The paper is organized as follows. In Section 2 we review basics about crystal bases and virtual crystals. In Section 3 we define rigged configurations. The new crystal structure on rigged configurations is presented in section 4. Section 5 is devoted to type A, where we give an explicit characterization of the unrestricted rigged configurations, a fermionic formula for unrestricted Kostka polynomials, and the affine crystal structure.

2. Crystals

2.1. Axiomatic definition. Kashiwara [**16, 17**] introduced a **crystal** as an edge-colored directed graph satisfying a simple set of axioms. Let $\mathfrak{g}$ be a symmetrizable Kac–Moody algebra with associated root, coroot and weight lattices $Q, Q^\vee, P$. Let I be the index set of the Dynkin diagram and denote the simple roots, simple coroots and fundamental weights by α_i, h_i and Λ_i ($i \in I$), respectively. There is a natural pairing $\langle \cdot\,,\,\cdot \rangle : Q^\vee \otimes P \to \mathbb{Z}$ defined by $\langle h_i\,,\,\Lambda_j\rangle = \delta_{ij}$.

The vertices of the crystal graph are elements of a set B. The edges of the crystal graph are colored by the index set I. A P-weighted I-crystal satisfies the following properties:

(1) Fix an $i \in I$. If all edges are removed except those colored i, the connected components are finite directed linear paths called the i-**strings** of B. Given $b \in B$, define $f_i(b)$ (resp. $e_i(b)$) to be the vertex following (resp. preceding) b in its i-string; if there is no such vertex, declare $f_i(b)$ (resp. $e_i(b)$) to be undefined. Define $\varphi_i(b)$ (resp. $\varepsilon_i(b)$) to be the number of arrows from b to the end (resp. beginning) of its i-string.
(2) There is a function $\mathrm{wt} : B \to P$ such that $\mathrm{wt}(f_i(b)) = \mathrm{wt}(b) - \alpha_i$ and $\varphi_i(b) - \varepsilon_i(b) = \langle h_i\,,\,\mathrm{wt}(b)\rangle$.

2.2. Virtual crystals. There exist natural inclusions of affine Lie algebras as indicated in Figures 2 and 3. Even though these embeddings do not carry over to the corresponding quantum algebras, it is expected that such embeddings exist for crystals. Note that every affine algebra can be embedded into one of type $A^{(1)}, D^{(1)}$ and $E^{(1)}$ which are the untwisted affine algebras whose canonical simple Lie subalgebra is simply-laced. Crystal embeddings corresponding to $C_n^{(1)}, A_{2n}^{(2)}, D_{n+1}^{(2)} \hookrightarrow A_{2n-1}^{(1)}$ have been studied in [**29**], whereas the crystal embeddings $B_n^{(1)}, A_{2n-1}^{(2)} \hookrightarrow D_{n+1}^{(1)}$ have been established in [**30**].

Consider an embedding of the affine algebra with Dynkin diagram X into one with diagram Y. We consider a graph automorphism σ of Y that fixes the 0 node. For type $A_{2n-1}^{(1)}$, $\sigma(i) = 2n - i \pmod{2n}$. For type $D_{n+1}^{(1)}$ the automorphism interchanges the nodes n and $n+1$ and fixes all other nodes. There is an additional automorphism for type $D_4^{(1)}$, namely, the cyclic permutation of the nodes 1,2 and 3. For type $E_6^{(1)}$ the automorphism exchanges nodes 1 and 5 and nodes 2 and 4. In Figures 2 and 3 the automorphism σ is illustrated pictorially by arrows.

Let I^X and I^Y be the vertex sets of the diagrams X and Y respectively, I^Y/σ the set of orbits of the action of σ on I^Y, and $\iota : I^X \to I^Y/\sigma$ a bijection which preserves edges and sends 0 to 0.

EXAMPLE 2.1.
If X is one of $C_n^{(1)}, A_{2n}^{(2)}, D_{n+1}^{(2)}$ and $Y = A_{2n-1}^{(1)}$, then $\iota(0) = 0$, $\iota(i) = \{i, 2n-i\}$ for $1 \le i < n$ and $\iota(n) = n$.
If $X = B_n^{(1)}$ or $A_{2n-1}^{(2)}$ and $Y = D_{n+1}^{(1)}$, then $\iota(i) = i$ for $i < n$ and $\iota(n) = \{n, n+1\}$.
If X is $D_4^{(3)}$ or $G_2^{(1)}$ and $Y = D_4^{(1)}$, then $\iota(0) = 0$, $\iota(1) = 2$ and $\iota(2) = \{1,3,4\}$.
If X is $E_6^{(2)}$ or $F_4^{(1)}$ and $Y = E_6^{(1)}$, then $\iota(0) = 0$, $\iota(1) = 1$, $\iota(2) = 3$, $\iota(3) = \{2,4\}$ and $\iota(4) = \{1,5\}$.

To describe the embedding we endow the bijection ι with additional data. For each $i \in I^X$ we shall define a multiplication factor γ_i that depends on the location of i with respect to a distinguished arrow (multiple bond) in X. Removing the arrow leaves two connected components. The factor γ_i is defined as follows:

(1) Suppose X has a unique arrow.
   (a) Suppose the arrow points towards the component of 0. Then $\gamma_i = 1$ for all $i \in I^X$.
   (b) Suppose the arrow points away from the component of 0. Then γ_i is the order of σ for i in the component of 0 and is 1 otherwise.
(2) Suppose X has two arrows. Then $\gamma_i = 1$ for $1 \le i \le n-1$. For $i \in \{0, n\}$, $\gamma_i = 2$ (which is the order of σ) if the arrow incident to i points away from it and is 1 otherwise.

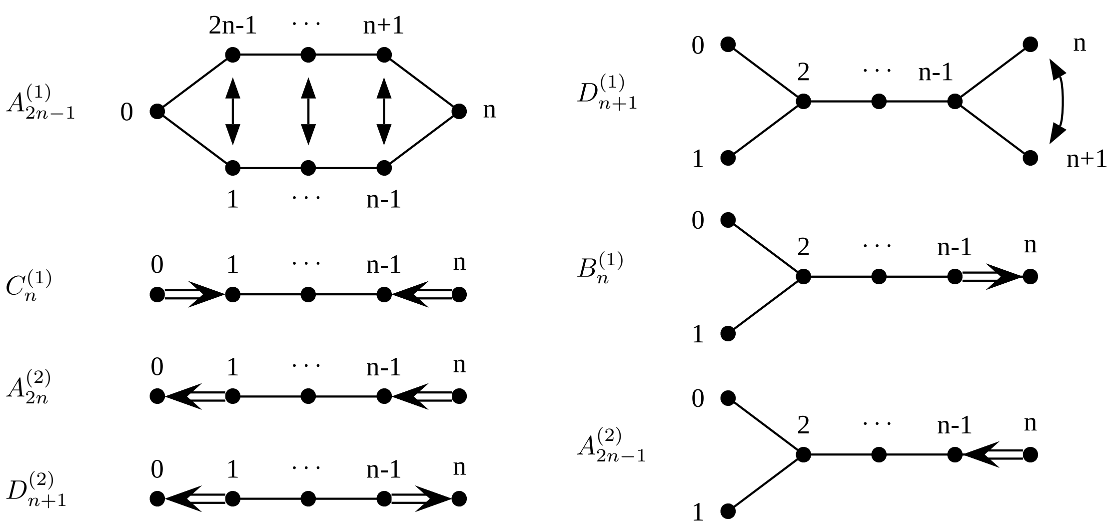


FIGURE 2. Embeddings $C_n^{(1)}, A_{2n}^{(2)}, D_{n+1}^{(2)} \hookrightarrow A_{2n-1}^{(1)}$ and $B_n^{(1)}, A_{2n-1}^{(2)} \hookrightarrow D_{n+1}^{(1)}$

EXAMPLE 2.2. The values of γ_i are summarized in the following table:

X		
$A_{2n-1}^{(2)}$ $D_4^{(3)}$ $E_6^{(2)}$	$\gamma_i = 1$	for all i
$B_n^{(1)}$	$\gamma_i = 2$ $\gamma_n = 1$	for $0 \le i \le n-1$
$G_2^{(1)}$	$\gamma_i = 3$ $\gamma_2 = 1$	for $i = 0, 1$
$F_4^{(1)}$	$\gamma_i = 2$ $\gamma_i = 1$	for $i = 0, 1, 2$ for $i = 3, 4$
$C_n^{(1)}$	$\gamma_i = 1$ $\gamma_0 = \gamma_n = 2$	for $1 \le i < n$
$A_{2n}^{(2)}$	$\gamma_i = 1$ $\gamma_n = 2$	for $0 \le i < n$
$D_{n+1}^{(2)}$	$\gamma_i = 1$	for all i

The embedding $\Psi : P^X \to P^Y$ of weight lattices is defined by

$$\Psi(\Lambda_i^X) = \gamma_i \sum_{j \in \iota(i)} \Lambda_j^Y .$$

Let $\widehat{V}$ be a Y-crystal. We define the virtual crystal operators $\widehat{e}_i, \widehat{f}_i$ for $i \in I^X$ as the composites of Y-crystal operators f_j, e_j given by

$$\widehat{f}_i = \prod_{j \in \iota(i)} f_j^{\gamma_i} \qquad \text{and} \qquad \widehat{e}_i = \prod_{j \in \iota(i)} e_j^{\gamma_i}. \tag{2.1}$$

These are designed to simulate X-crystal operators f_i, e_i for $i \in I^X$. The type Y operators on the right hand side, may be performed in any order, since distinct nodes $j, j' \in \iota(i)$ are not adjacent in Y and thus their corresponding raising and lowering operators commute.

A **virtual crystal** is a pair $(V, \widehat{V})$ such that:

(1) $\widehat{V}$ is a Y-crystal.
(2) $V \subset \widehat{V}$ is closed under $\widehat{e}_i, \widehat{f}_i$ for $i \in I^X$.

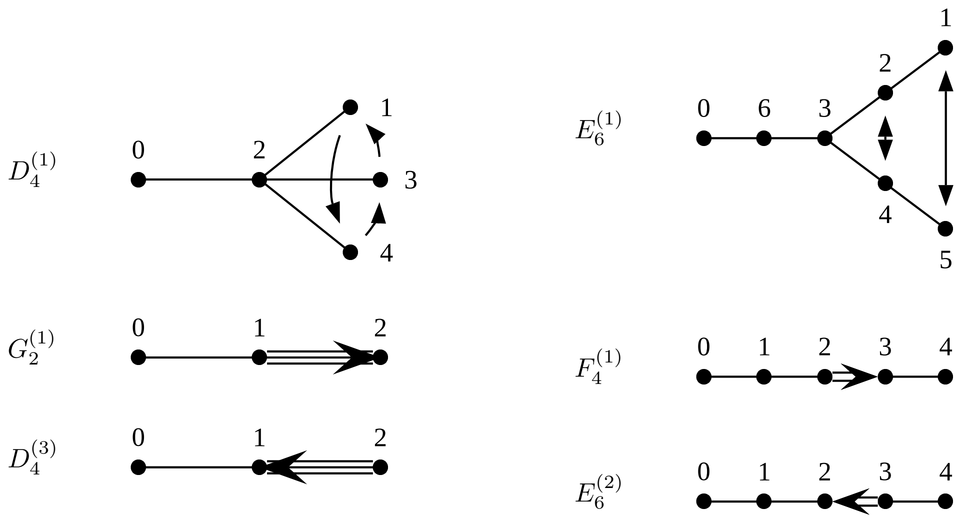


FIGURE 3. Embeddings $G_2^{(1)}, D_4^{(3)} \hookrightarrow D_4^{(1)}$ and $F_4^{(1)}, E_6^{(2)} \hookrightarrow E_6^{(1)}$

(3) There is an X-crystal B and an X-crystal isomorphism $\Psi : B \to V$ such that e_i, f_i correspond to $\widehat{e}_i, \widehat{f}_i$.

Sometimes by abuse of notation, V will be referred to as a virtual crystal.

Let us define the Y-crystal

$$\widehat{V}^{r,s} = \bigotimes_{j \in \iota(r)} B_Y^{j,\gamma_r s}$$

except for $A_{2n}^{(2)}$ and $r = n$ in which case $\widehat{V}^{n,s} = B_Y^{n,s} \otimes B_Y^{n,s}$. Denote by $u(\widehat{V}^{r,s})$ the extremal vector of weight $\Psi(s\Lambda_r)$ in $\widehat{V}^{r,s}$.

DEFINITION 2.3. Let $V^{r,s}$ be the subset of $\widehat{V}^{r,s}$ generated from $u(\widehat{V}^{r,s})$ using the virtual crystal operators $\widehat{e}_i$ and $\widehat{f}_i$ for $i \in I^X$.

CONJECTURE 2.4. [**30**, Conjecture 3.7] *There is an isomorphism of X-crystals $\Psi : B_X^{r,s} \cong V^{r,s}$ such that e_i and f_i correspond to $\widehat{e}_i$ and $\widehat{f}_i$ respectively, for all $i \in I^X$.*

In [**29**] Conjecture 2.4 is proved for embeddings $C_n^{(1)}, A_{2n}^{(2)}, D_{n+1}^{(2)} \hookrightarrow A_{2n-1}^{(1)}$ and $s = 1$. In [**30**] Conjecture 2.4 is proved for all nonexceptional types when $r = 1$.

3. Rigged configurations

In this section we define rigged configurations for all affine Kac–Moody algebras. Type $A_{2n}^{(2)}$ requires some special treatment. We need the variant $\widetilde{\gamma}_a$ of the multiplication factor γ_a which is $\widetilde{\gamma}_a = \gamma_a$ except for $A_{2n}^{(2)}$ and $a = n$ when $\widetilde{\gamma}_n = 1$. Also set $\widetilde{\alpha}_a = \alpha_a$ for all $a \in I$ except for type $A_{2n}^{(2)}$ in which case $\widetilde{\alpha}_a$ are the simple roots of type B_n.

Let $L = (L_i^{(a)})_{(a,i)\in\mathcal{H}}$ be an array of nonnegative integers where $\mathcal{H} = \{1, 2, \dots, n\} \times \mathbb{Z}_{>0}$, called the multiplicity array, where n is the rank of the underlying algebra and Λ a weight. Then an (L, Λ)-configuration is an array $m = (m_i^{(a)})_{(a,i)\in\mathcal{H}}$ such that

$$\sum_{(a,i)\in\mathcal{H}} i\, m_i^{(a)} \widetilde{\alpha}_a = \sum_{(a,i)\in\mathcal{H}} i\, L_i^{(a)} \Lambda_a - \Lambda \tag{3.1}$$

except for type $A_{2n}^{(2)}$. In this case the right hand side should be replaced by $\iota(\mathrm{r.h.s})$ where ι is a $\mathbb{Z}$-linear map from the weight lattice of type C_n to the weight lattice of type B_n such that

$$\iota(\Lambda_a^C) = \begin{cases} \Lambda_a^B & \text{for } 1 \le a < n \\ 2\Lambda_a^B & \text{for } a = n. \end{cases}$$

The vacancy numbers of a given configuration are defined as

$$p_i^{(a)} = \sum_{(b,j)\in\mathcal{H}} -\frac{2(\alpha_a \mid \alpha_b)}{\gamma_b(\alpha_a \mid \alpha_a)} \min(\widetilde{\gamma}_a i, \widetilde{\gamma}_b j) m_j^{(b)} + \sum_{j\ge 0} \min(i,j) L_j^{(a)}. \tag{3.2}$$

An (L,Λ)-configuration is called **admissible** if $p_i^{(a)} \ge 0$ for all $(a,i)\in\mathcal{H}$. The set of admissible (L,Λ)-configurations is denoted by $\overline{\mathrm{C}}(L,\Lambda)$.

A **rigged configuration** is a pair (m,J) where $m=(m_i^{(a)})_{(a,i)\in\mathcal{H}}$ is an admissible (L,Λ)-configuration and $J=(J_i^{(a)})_{(a,i)\in\mathcal{H}}$ is a matrix of partitions such that the partition $J_i^{(a)}$ is contained in a rectangle of size $m_i^{(a)}\times p_i^{(a)}$. The set of rigged configurations for fixed L and Λ is denoted by $\overline{\mathrm{RC}}(L,\Lambda)$.

Rigged configurations can also be represented as a sequence of partitions such that each part of each partition is labeled or "rigged" by a number. Let $\nu=(\nu^{(1)},\nu^{(2)},\dots,\nu^{(n)})$ be the sequence of partitions obtained from $m=(m_i^{(a)})$ as follows. Let $m_i^{(a)}(\nu)$ be the number of parts in $\nu^{(a)}$ of size i. Then ν is determined by requiring that

$$m_{\widetilde{\gamma}_a i}^{(a)}(\nu) = m_i^{(a)} \quad \text{and} \quad m_j^{(a)}(\nu)=0 \quad \text{for } j\notin \widetilde{\gamma}_a\mathbb{Z}.$$

The vacancy number $P_i^{(a)}(\nu)$ for each part i of $\nu^{(a)}$ is then

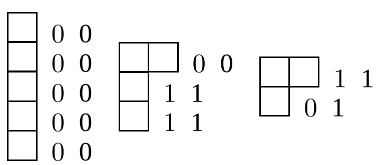
$$P_i^{(a)}(\nu) = \sum_{b\in I} -\frac{2(\alpha_a \mid \alpha_b)}{\gamma_b(\alpha_a \mid \alpha_a)} Q_i(\nu^{(b)}) + \sum_{j\ge 0} \min(\frac{i}{\widetilde{\gamma}_a}, j) L_j^{(a)},$$

where $Q_i(\rho)$ is the number of boxes in the first i columns of the partition ρ. The relation to $p_i^{(a)}$ is

$$p_i^{(a)} = P_{\widetilde{\gamma}_a i}^{(a)}(\nu).$$

A tuple (i,x) where i is a part of $\nu^{(a)}$ and x is a part of $J_i^{(a)}$ is called a string of the rigged partition $(\nu,J)^{(a)}$. Here i is the length and x the label of the string. The **colabel** of a string (i,x) of $(\nu,J)^{(a)}$ is $P_i^{(a)}(\nu)-x$.

EXAMPLE 3.1. Let $\Lambda=\Lambda_1+\Lambda_3$ of type $A_6^{(2)}$, $L_1^{(1)}=7$ and all other $L_i^{(a)}=0$. Then

$(\nu, J) =$ 

$\in \overline{\mathrm{RC}}(L,\Lambda)$,

where the first number behind each part is the label and the second one is the vacancy number.

There is also a statistic called **cocharge** defined on rigged configurations. Set $t_a^\vee = \frac{|\iota(a)|\gamma_a}{\gamma_0}$. The cocharge is given by

$$\begin{aligned} \mathrm{cc}(\nu) &= \sum_{(i,a),(b,j)\in\mathcal{H}} \frac{t_a^\vee}{\gamma_b}\cdot\frac{(\alpha_a\mid\alpha_b)}{(\alpha_a\mid\alpha_a)} \min(\widetilde{\gamma}_a i,\widetilde{\gamma}_b j) m_i^{(a)} m_j^{(b)} \\ &= \frac{1}{2}\sum_{(a,i)\in\mathcal{H}} t_a^\vee m_i^{(a)} \big(\sum_{j\ge 0}\min(i,j)L_j^{(a)} - p_i^{(a)}\big) \end{aligned} \tag{3.3}$$

for a configuration ν and $\mathrm{cc}(\nu,J)=\mathrm{cc}(\nu)+|J|$ where $|J|=\sum_{(a,i)\in\mathcal{H}} t_a^\vee |J_i^{(a)}|$ is the sum of the sizes of all partitions $J_i^{(a)}$ weighted by $t_a^\vee$.

As mentioned in the introduction, rigged configurations correspond to highest weight crystal elements. Let $B^{r,s}$ be a Kirillov–Reshetikhin crystal for $(r,s)\in\mathcal{H}$ and $B=B^{r_k,s_k}\otimes B^{r_{k-1},s_{k-1}}\otimes\cdots\otimes B^{r_1,s_1}$. Associate to B the multiplicity array $L=(L_s^{(r)})_{(r,s)\in\mathcal{H}}$ where $L_s^{(r)}$ counts the number of tensor factors $B^{r,s}$ in B. Denote by

$$\overline{\mathcal{P}}(B,\Lambda) = \{b\in B \mid \mathrm{wt}(b)=\Lambda,\, e_i(b) \text{ undefined for all } i\in I\}$$

the set of all highest weight elements of weight Λ in B. There is a natural statistics defined on B, called energy function or more precisely tail coenergy function $D: B\to\mathbb{Z}$ (see [**35**, Eq. (5.1)] for a precise definition).

The following theorem was proven in [**25**] for type $A_{n-1}^{(1)}$ and general $B=B^{r_k,s_k}\otimes\cdots\otimes B^{r_1,s_1}$, in [**32**] for type $D_n^{(1)}$ and $B=B^{r_k,1}\otimes\cdots\otimes B^{r_1,1}$ and in [**35**] for type $D_n^{(1)}$ and $B=B^{1,s_k}\otimes\cdots\otimes B^{1,s_1}$.

THEOREM 3.2. [**25, 32, 35**] *For Λ a dominant weight, B as above and L the corresponding multiplicity array, there is a bijection $\overline{\Phi} : \overline{\mathcal{P}}(B,\Lambda) \to \overline{\mathrm{RC}}(L,\Lambda)$ which preserves the statistics, that is, $D(b) = \mathrm{cc}(\overline{\Phi}(b))$ for all $b \in \overline{\mathcal{P}}(B,\Lambda)$.*

Defining the generating functions

$$\overline{X}(B,\Lambda) = \sum_{b \in \overline{\mathcal{P}}(B,\Lambda)} q^{D(b)} \qquad \text{and} \qquad \overline{M}(L,\Lambda) = \sum_{(\nu,J)\in\overline{\mathrm{RC}}(L,\Lambda)} q^{\mathrm{cc}(\nu,J)}, \tag{3.4}$$

we get the immediate corollary of Theorem 3.2.

COROLLARY 3.3. [**25, 32, 35**] *Let Λ, B and L as in Theorem 3.2. Then $\overline{X}(B,\Lambda) = \overline{M}(L,\Lambda)$.*

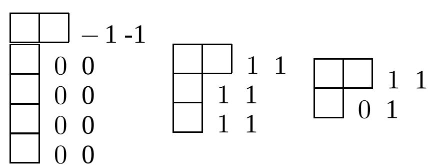
4. Crystal structure on rigged configurations

The rigged configurations of section 3 correspond to highest weight crystal elements. In this section we introduce the set of unrestricted rigged configurations $\mathrm{RC}(L)$ by defining a crystal structure generated from highest weight vectors given by elements in $\overline{\mathrm{RC}}(L) = \bigcup_{\Lambda\in P^+} \overline{\mathrm{RC}}(L,\Lambda)$ by the Kashiwara operators e_a, f_a. For simply-laced algebras the following definition was given in [**33**, Definition 3.3]. The multiplication factors γ_a for the simply-laced case are equal to 1.

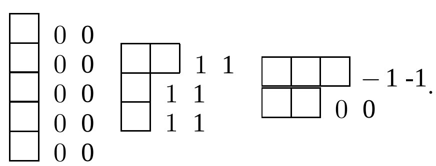
DEFINITION 4.1. Let L be a multiplicity array. Define the set of **unrestricted rigged configurations** $\mathrm{RC}(L)$ as the set generated from the elements in $\overline{\mathrm{RC}}(L)$ by the application of the operators f_a, e_a for $1 \le a \le n$ defined as follows:

(1) Define $e_a(\nu,J)$ by removing γ_a boxes from a string of length k in $(\nu,J)^{(a)}$ leaving all colabels fixed and increasing the new label by one. Here k is the length of the string with the smallest negative rigging of smallest length. If no such string exists, $e_a(\nu,J)$ is undefined.
(2) Define $f_a(\nu,J)$ by adding γ_a boxes to a string of length k in $(\nu,J)^{(a)}$ leaving all colabels fixed and decreasing the new label by one. Here k is the length of the string with the smallest nonpositive rigging of largest length. If no such string exists, add a new string of length one and label -1. If the result is not a valid unrestricted rigged configuration $f_a(\nu,J)$ is undefined.

EXAMPLE 4.2. For (ν,J) of Example 3.1 we have

$f_1(\nu,J) \quad =$ 

and

$f_3(\nu,J) \quad =$ 

THEOREM 4.3. *The operators e_a, f_a of Definition 4.1 are the Kashiwara crystal operators.*

For simply-laced algebras Theorem 4.3 was proven in [**33**] by using the local characterization of simply-laced crystals given by Stembridge [**37**]. In the following we show that, assuming that the virtual crystal embeddings of section 2.2 hold, Theorem 4.3 is also true for the nonsimply-laced algebras.

We define **virtual rigged configurations** in analogy to virtual crystals. Here $B = B^{r_k,s_k} \otimes \cdots \otimes B^{r_1,s_1}$ is a tensor product of Kirillov-Reshetikhin crystals and $L = (L_i^{(a)})$ the corresponding multiplicity array.

DEFINITION 4.4. Let $X \hookrightarrow Y$ be one of the algebra embeddings of section 2.2, Λ a weight and B a crystal for type X. Let $(V,\widehat{V})$ be the virtual Y-crystal corresponding to B. Then $\mathrm{RC}^v(L,\Lambda)$ is the set of elements $(\widehat{\nu},\widehat{J}) \in \mathrm{RC}(\widehat{L},\Psi(\Lambda))$ such that:

(1) For all $i \in \mathbb{Z}_{>0}$, $\widehat{m}_i^{(a)} = \widehat{m}_i^{(b)}$ and $\widehat{J}_i^{(a)} = \widehat{J}_i^{(b)}$ if a and b are in the same σ-orbit in I^Y.
(2) For all $i \in \mathbb{Z}_{>0}$, $a \in I^X$, and $b \in \iota(a) \subset I^Y$, we have $\widehat{m}_j^{(b)} = 0$ if $j \notin \widetilde{\gamma}_a\mathbb{Z}$ and the parts of $\widehat{J}_i^{(b)}$ are multiples of γ_a.

THEOREM 4.5. [**30**, Theorem 4.2] *There is a bijection* $\mathrm{RC}(L,\Lambda) \to \mathrm{RC}^v(L,\Lambda)$ *sending* $(\nu, J) \mapsto (\widehat{\nu}, \widehat{J})$ *given as follows. For all* $a \in I^X$, $b \in \iota(a) \subset I^Y$, *and* $i \in \mathbb{Z}_{>0}$,

$$\widehat{m}^{(b)}_{\widetilde{\gamma}_a i} = m^{(a)}_i \qquad \text{and} \qquad \widehat{J}^{(b)}_{\widetilde{\gamma}_a i} = \gamma_a J^{(a)}_i.$$

The cocharge changes by $\mathrm{cc}(\widehat{\nu}, \widehat{J}) = \gamma_0 \, \mathrm{cc}(\nu, J)$.

PROOF OF THEOREM 4.3. Theorem 4.3 was proved in [**33**] for the simply-laced algebras. Hence, assuming that the virtual crystal embeddings of section 2.2 hold, it suffices to check that e_a, f_a of Definition 4.1 satisfy (2.1). By Theorem 4.5 this reduces to checking that $\widehat{f}_a$ and $\widehat{e}_a$ preserve the conditions of Definition 4.4. We demonstrate this for $\widehat{f}_a$; the arguments for $\widehat{e}_a$ are analogous. Let $(\widehat{\nu}, \widehat{J}) \in \mathrm{RC}^v(L,\Lambda)$. Since f_a and f_b of Definition 4.1 for simply-laced algebras commute if $b \in \iota(a)$, point (1) of Definition 4.4 follows for $\widehat{f}_a(\widehat{\nu}, \widehat{J})$. To prove that point (2) holds, it suffices to check that if $\gamma_a > 1$, then the various applications of f_a in $\widehat{f}_a$ select the same string γ_a times. Note that for simply-laced algebras the application of f_a changes the vacancy number $\widehat{p}^{(b)}_i$ by

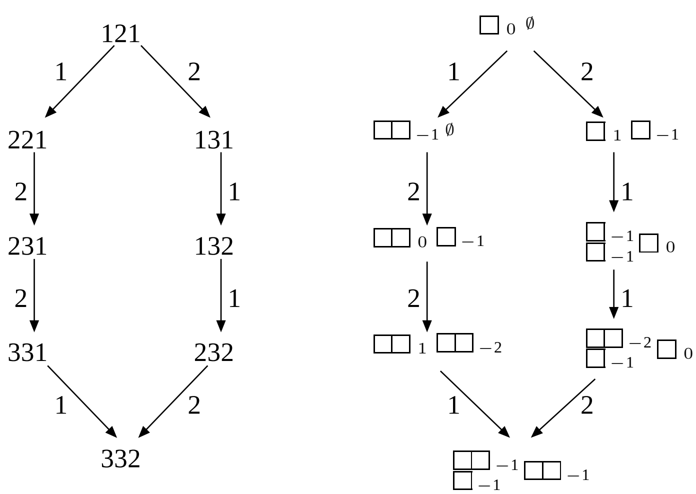
$$\widehat{p}^{(b)}_i \mapsto \widehat{p}^{(b)}_i - (\alpha_a \mid \alpha_b)\chi(i > k), \tag{4.1}$$

where k is the length of the selected string. By the definition of k (see Definition 4.1) and the fact that all riggings in the a-th rigged partition have parity γ_a by point (2) of Definition 4.4, all riggings of strings of length $i > k$ in $(\widehat{\nu}, \widehat{J})^{(a)}$ are greater or equal to $-s + \gamma_a$, where $-s$ is the smallest rigging appearing in $(\widehat{\nu}, \widehat{J})^{(a)}$. By (4.1) the riggings of length $i > k$ in $(\widehat{\nu}, \widehat{J})^{(a)}$ change by -2. Hence the smallest j such that $-s + \gamma_a - 2j \le -s - j$ is $j = \gamma_a$. This shows that γ_a applications of f_a select the same string, which in turn proves that $\widehat{f}_a(\widehat{\nu}, \widehat{J})$ satisfies the conditions of Definition 4.4. □

THEOREM 4.6. *With the same assumptions as in Theorem 3.2, the graph generated from* $(\overline{\nu}, \overline{J}) \in \overline{\mathrm{RC}}(L,\Lambda)$ *and the crystal operators* e_a, f_a *of Definition 4.1 is isomorphic to the crystal graph* $B(\Lambda)$ *of highest weight* Λ.

PROOF. For simply-laced types this was proven in [**33**, Theorem 3.7]. For nonsimply-laced types this follows from Theorems 4.3 and 4.5. □

EXAMPLE 4.7. Consider the crystal $B(\,\begin{smallmatrix}\square\square\\ \square\end{smallmatrix}\,)$ of type A_2 in $B = (B^{1,1})^{\otimes 3}$. Here is the crystal graph in the usual labeling and the rigged configuration labeling:



THEOREM 4.8. *The cocharge* cc *as defined in* (3.3) *is constant on connected crystal components.*

PROOF. For simply-laced types this was proved in [**33**, Theorem 3.9]. For nonsimply-laced types this follows from Theorems 4.3 and 4.5. □

EXAMPLE 4.9. The cocharge of the connected component in Example 4.7 is 1.

For $B = B^{r_k,s_k} \otimes \cdots \otimes B^{r_1,s_1}$ and $\Lambda \in P$ let

$$\mathcal{P}(B,\Lambda) = \{ b \in B \mid \mathrm{wt}(b) = \Lambda \}.$$

THEOREM 4.10. *Let $\Lambda \in P$, B be as in Theorem 3.2 and L the corresponding multiplicity array. Then there is a bijection $\Phi : \mathcal{P}(B,\Lambda) \to \mathrm{RC}(L,\Lambda)$ which preserves the statistics, that is, $D(b) = \mathrm{cc}(\Phi(b))$ for all $b \in \mathcal{P}(B,\Lambda)$.*

PROOF. By Theorem 3.2 there is such a bijection for the maximal elements $b \in \overline{\mathcal{P}}(B)$. By Theorems 4.6 and 4.8 this extends to all of $\mathcal{P}(B,\Lambda)$. □

Extending the definitions of (3.4) to

$$X(B,\Lambda) = \sum_{b\in\mathcal{P}(B,\Lambda)} q^{D(b)} \qquad \text{and} \qquad M(L,\Lambda) = \sum_{(\nu,J)\in\mathrm{RC}(L,\Lambda)} q^{\mathrm{cc}(\nu,J)}, \tag{4.2}$$

we obtain the corollary:

COROLLARY 4.11. *With all hypotheses of Theorem 4.10, we have $X(B,\Lambda) = M(L,\Lambda)$.*

5. Unrestricted rigged configurations for type $A_{n-1}^{(1)}$

In this section we give an explicit description of the elements in $\mathrm{RC}(L,\lambda)$ for type $A_{n-1}^{(1)}$. Generally speaking, the elements are rigged configurations where the labels lie between the vacancy number and certain lower bounds defined explicitly. This characterization will be used in section 5.2 to write down an explicit fermionic formula $M(L,\lambda)$ for the unrestricted configuration sum $X(B,\lambda)$. Section 5.3 is devoted to the affine crystal structure of $\mathrm{RC}(L,\lambda)$.

5.1. Characterization of unrestricted rigged configurations. Let $L = (L_i^{(a)})_{(a,i)\in\mathcal{H}}$ be a multiplicity array and $\lambda = (\lambda_1,\ldots,\lambda_n)$ be the n-tuple of nonnegative integers. The set of (L,λ)-configurations $\mathrm{C}(L,\lambda)$ is the set of all sequences of partitions $\nu = (\nu^{(a)})_{a\in I}$ such that (3.1) holds. As discussed in Section 3, in the usual setting a rigged configuration $(\nu,J) \in \overline{\mathrm{RC}}(L,\lambda)$ consists of a configuration $\nu \in \overline{\mathrm{C}}(L,\lambda)$ together with a double sequence of partitions $J = \{J_i^{(a)} \mid (a,i)\in\mathcal{H}\}$ such that the partition $J_i^{(a)}$ is contained in a $m_i^{(a)} \times p_i^{(a)}$ rectangle. In particular this requires that $p_i^{(a)} \geq 0$. The unrestricted rigged configurations $(\nu,J) \in \mathrm{RC}(L,\lambda)$ can contain labels that are negative, that is, the lower bound on the parts in $J_i^{(a)}$ can be less than zero.

To define the lower bounds we need the following notation. Let $\lambda' = (c_1, c_2, \ldots, c_{n-1})^t$, where $c_k = \lambda_{k+1} + \lambda_{k+2} + \cdots + \lambda_n$ is the length of the k-th column of λ', and let $\mathcal{A}(\lambda')$ be the set of tableaux of shape λ' such that the entries are strictly decreasing along columns, and the letters in column k are from the set $\{1,2,\ldots,c_{k-1}\}$ with $c_0 = c_1$.

EXAMPLE 5.1. For $n = 4$ and $\lambda = (0,1,1,1)$, the set $\mathcal{A}(\lambda')$ consists of the following tableaux

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

REMARK 5.2. Denote by $t_{j,k}$ the entry of $t \in \mathcal{A}(\lambda')$ in row j and column k. Note that $c_k - j + 1 \leq t_{j,k} \leq c_{k-1} - j + 1$ since the entries in column k are strictly decreasing and lie in the set $\{1,2,\ldots,c_{k-1}\}$. This implies $t_{j,k} \leq c_{k-1} - j + 1 \leq t_{j,k-1}$, so that the rows of t are weakly decreasing.

Given $t \in \mathcal{A}(\lambda')$, we define the **lower bound** as

$$M_i^{(a)}(t) = -\sum_{j=1}^{c_a} \chi(i \geq t_{j,a}) + \sum_{j=1}^{c_{a+1}} \chi(i \geq t_{j,a+1}),$$

where recall that $\chi(S) = 1$ if the the statement S is true and $\chi(S) = 0$ otherwise.

Let $M, p, m \in \mathbb{Z}$ such that $m \geq 0$. A (M,p,m)-quasipartition μ is a tuple of integers $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ such that $M \leq \mu_m \leq \mu_{m-1} \leq \cdots \leq \mu_1 \leq p$. Each μ_i is called a part of μ. Note that for $M = 0$ this would be a partition with at most m parts each not exceeding p.

The following theorem shows that the set of unrestricted rigged configurations can be characterized via the lower bounds.

THEOREM 5.3. [**33**, Theorem 4.6] *Let $(\nu,J) \in \mathrm{RC}(L,\lambda)$. Then $\nu \in \mathrm{C}(L,\lambda)$ and $J_i^{(a)}$ is a $(M_i^{(a)}(t), p_i^{(a)}, m_i^{(a)})$-quasipartition for some $t \in \mathcal{A}(\lambda')$. Conversely, every (ν,J) such that $\nu \in \mathrm{C}(L,\lambda)$ and $J_i^{(a)}$ is a $(M_i^{(a)}(t), p_i^{(a)}, m_i^{(a)})$-quasipartition for some $t \in \mathcal{A}(\lambda')$ is in $\mathrm{RC}(L,\lambda)$.*

EXAMPLE 5.4. Let $n=4$, $\lambda=(2,2,1,1)$, $L_1^{(1)}=6$ and all other $L_i^{(a)}=0$. Then

$$(\nu,J) \;=\; \begin{array}{l} \square\square\square \;\; -2\;0 \\ \square \;\; 0\;\;3 \end{array} \qquad \square\square \;\; 0\;\;0 \qquad \square \;\; -1\;{-1}$$

is an unrestricted rigged configuration in $\mathrm{RC}(L,\lambda)$, where we have written the parts of $J_i^{(a)}$ next to the parts of length i in partition $\nu^{(a)}$. The second number is the corresponding vacancy number $p_i^{(a)}$. This shows that the labels are indeed all weakly below the vacancy numbers. For

$$\begin{array}{|c|c|c|} \hline 4 & 4 & 1 \\ \hline 3 & 3 & \multicolumn{1}{c}{} \\ \cline{1-2} 2 & \multicolumn{2}{c}{} \\ \cline{1-1} 1 & \multicolumn{2}{c}{} \\ \cline{1-1} \end{array} \in \mathcal{A}(\lambda')$$

we get the lower bounds

$$\begin{array}{l} \square\square\square \;\; -2 \\ \square \;\; -1 \end{array} \qquad \square\square \;\; 0 \qquad \square \;\; -1,$$

which are less or equal to the riggings in (ν,J).

For type A_1 we have $\lambda=(\lambda_1,\lambda_2)$ so that $\mathcal{A}=\{t\}$ contains just the single one-column tableau of height λ_2 filled with the numbers $1,2,\ldots,\lambda_2$. In this case $M_i(t)=-\sum_{j=1}^{\lambda_2}\chi(i\ge t_{j,1})=-i$, which agrees with the findings of [**38**].

The characterization of unrestricted rigged configurations is similar to the characterization of level-restricted rigged configurations [**34**, Definition 5.5]. Whereas the unrestricted rigged configurations are characterized in terms of lower bounds, for level-restricted rigged configurations the vacancy number has to be modified according to tableaux in a certain set.

5.2. Fermionic formula. With the explicit characterization of the unrestricted rigged configurations of Section 5.1, it is possible to derive an explicit formula for the polynomials $M(L,\lambda)$ of (4.2).

Let $\mathcal{SA}(\lambda')$ be the set of all nonempty subsets of $\mathcal{A}(\lambda')$ and set

$$M_i^{(a)}(S)=\max\{M_i^{(a)}(t)\mid t\in S\}\qquad \text{for } S\in\mathcal{SA}(\lambda').$$

By inclusion-exclusion the set of all allowed riggings for a given $\nu\in\mathrm{C}(L,\lambda)$ is

$$\bigcup_{S\in\mathcal{SA}(\lambda')}(-1)^{|S|+1}\{J\mid J_i^{(a)} \text{ is a } (M_i^{(a)}(S),p_i^{(a)},m_i^{(a)})\text{-quasipartition}\}.$$

The q-binomial coefficient $\begin{bmatrix} m+p \\ m\end{bmatrix}$, defined as

$$\begin{bmatrix} m+p \\ m\end{bmatrix}=\frac{(q)_{m+p}}{(q)_m(q)_p},$$

where $(q)_n=(1-q)(1-q^2)\cdots(1-q^n)$, is the generating function of partitions with at most m parts each not exceeding p. Hence the polynomial $M(L,\lambda)$ may be rewritten as

$$M(L,\lambda)=\sum_{S\in\mathcal{SA}(\lambda')}(-1)^{|S|+1}\sum_{\nu\in\mathrm{C}(L,\lambda)}q^{\mathrm{cc}(\nu)+\sum_{(a,i)\in\mathcal{H}}m_i^{(a)}M_i^{(a)}(S)}\prod_{(a,i)\in\mathcal{H}}\begin{bmatrix} m_i^{(a)}+p_i^{(a)}-M_i^{(a)}(S) \\ m_i^{(a)}\end{bmatrix} \tag{5.1}$$

called **fermionic formula**. By Corollary 4.11 this is also a formula for the unrestricted configuration sum $X(B,\lambda)$. This formula is different from the fermionic formulas of [**13, 18**] which exist in the special case when L is the multiplicity array of $B=B^{1,s_k}\otimes\cdots\otimes B^{1,s_1}$ or $B=B^{r_k,1}\otimes\cdots\otimes B^{r_1,1}$.

5.3. The Kashiwara operators e_0 and f_0. The Kirillov–Reshetikhin crystals $B^{r,s}$ are affine crystals and admit the Kashiwara operators e_0 and f_0. It was shown in [**36**] that for type $A_{n-1}^{(1)}$ they can be defined in terms of the **promotion operator** pr as

$$e_0=\mathrm{pr}^{-1}\circ e_1\circ\mathrm{pr}\quad\text{and}\quad f_0=\mathrm{pr}^{-1}\circ f_1\circ\mathrm{pr}.$$

The promotion operator is a bijection $\mathrm{pr} : B \to B$ such that the following diagram commutes for all $a \in I$

$$\begin{tikzcd} B \arrow[r, "\mathrm{pr}"] \arrow[d, "f_a"'] & B \arrow[d, "f_{a+1}"] \\ B \arrow[r, "\mathrm{pr}"'] & B \end{tikzcd} \tag{5.2}$$

and such that for every $b \in B$ the weight is rotated

$$\langle h_{a+1} \, , \, \mathrm{wt}(pr(b)) \rangle = \langle h_a \, , \, \mathrm{wt}(b) \rangle. \tag{5.3}$$

Here subscripts are taken modulo n.

We are now going to define the promotion operator on unrestricted rigged configurations.

DEFINITION 5.5. Let $(\nu, J) \in \mathrm{RC}(L, \lambda)$. Then $\mathrm{pr}(\nu, J)$ is obtained as follows:

(1) Set $(\nu', J') = f_1^{\lambda_1} f_2^{\lambda_2} \cdots f_n^{\lambda_n}(\nu, J)$ where f_n acts on $(\nu, J)^{(n)} = \emptyset$.
(2) Apply the following algorithm ρ to (ν', J') λ_n times: Find the smallest singular string in $(\nu', J')^{(n)}$. Let the length be $\ell^{(n)}$. Repeatedly find the smallest singular string in $(\nu', J')^{(k)}$ of length $\ell^{(k)} \geq \ell^{(k+1)}$ for all $1 \leq k < n$. Shorten the selected strings by one and make them singular again.

EXAMPLE 5.6. Let $B = B^{2,2}$, L the corresponding multiplicity array and $\lambda = (1, 0, 1, 2)$. Then

$$(\nu, J) = \quad \ydiagram{1}\ 0 \qquad \begin{array}{ll} \ydiagram{2} & -1 \\ \ydiagram{1} & -1 \end{array} \qquad \ydiagram{2}\ -1 \qquad \in \mathrm{RC}(L, \lambda)$$

corresponds to the tableau $b = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 4 & 4 \\ \hline \end{array} \in \mathcal{P}(B, \lambda)$. After step (1) of Definition 5.5 we have

$$(\nu', J') = \quad \ydiagram{2}\ -1 \qquad \begin{array}{ll} \ydiagram{2} & 1 \\ \ydiagram{1} & 0 \end{array} \qquad \begin{array}{ll} \ydiagram{2} & -1 \\ \ydiagram{1} & -1 \end{array} \qquad \ydiagram{2}\ -1.$$

Then applying step (2) yields

$$\mathrm{pr}(\nu, J) = \quad \emptyset \quad \ydiagram{1}\ 0 \quad \ydiagram{1}\ -1$$

which corresponds to the tableau $\mathrm{pr}(b) = \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & 4 \\ \hline \end{array}$.

LEMMA 5.7. [**33**, Lemma 4.10] *The map* pr *of Definition 5.5 is well-defined and satisfies* (5.2) *for* $1 \leq a \leq n-2$ *and* (5.3) *for* $0 \leq a \leq n-1$.

Lemma 7 of [**36**] states that for a single Kirillov–Reshetikhin crystal $B = B^{r,s}$ the promotion operator pr is uniquely determined by (5.2) for $1 \leq a \leq n-2$ and (5.3) for $0 \leq a \leq n-1$. Hence by Lemma 5.7 pr on $\mathrm{RC}(L)$ is indeed the correct promotion operator when L is the multiplicity array of $B = B^{r,s}$.

THEOREM 5.8. [**33**, Theorem 4.11] *Let L be the multiplicity array of $B = B^{r,s}$. Then* $\mathrm{pr} : \mathrm{RC}(L) \to \mathrm{RC}(L)$ *of Definition 5.5 is the promotion operator on rigged configurations.*

CONJECTURE 5.9. [**33**, Conjecture 4.12] *Theorem 5.8 is true for any $B = B^{r_k,s_k} \otimes \cdots \otimes B^{r_1,s_1}$.*

Unfortunately, the characterization [**36**, Lemma 7] does not suffice to define pr uniquely on tensor products $B = B^{r_k,s_k} \otimes \cdots \otimes B^{r_1,s_1}$. In [**8**] a bijection $\Phi : \mathcal{P}(B, \lambda) \to \mathrm{RC}(L, \lambda)$ is defined via a direct algorithm. It is expected that Conjecture 5.9 can be proven by showing that pr and Φ commute. Alternatively, an independent characterization of pr on tensor factors would give a new, more conceptual way of defining the bijection Φ between paths and (unrestricted) rigged configurations. A proof that the crystal operators f_a and e_a commute with Φ for $a = 1, 2, \ldots, n-1$ is given in [**8**].

References

[1] G. E. Andrews, *Multiple series Rogers–Ramanujan type identities*, Pacific J. Math. **114** (1984) 267–283.
[2] G. Albertini, S. Dasmahapatra, B. M. McCoy, *Spectrum and completeness of the integrable* 3*-state Potts model: a finite size study*, Adv. Ser. Math. Phys. **16** (1992) 1–53.
[3] G.E. Andrews, A. Schilling, S.O. Warnaar, *An A_2 Bailey lemma and Rogers-Ramanujan-type identities*, J. Amer. Math. Soc. **12** (1999), no. 3, 677–702.

[4] W. N. Bailey, *Identities of the Rogers–Ramanujan type*, Proc. London Math. Soc. (2) **50** (1949) 1–10.
[5] R. J. Baxter, *Exactly solved models in statistical mechanics*, Academic Press, London, 1982.
[6] H. A. Bethe, *Zur Theorie der Metalle, I. Eigenwerte und Eigenfunktionen der linearen Atomkette*, Z. Physik **71** (1931) 205–231.
[7] L. Deka, A. Schilling, *New explicit expression for $A_n^{(1)}$ supernomials*, Extended Abstract, 17th International Conference on Formal Power Series and Algebraic Combinatorics 2005, University of Messina, Italy, June 2005.
[8] L. Deka, A. Schilling, *New fermionic formula for unrestricted Kostka polynomials*, J. Combinatorial Theory, Series A, to appear (math.CO/0509194).
[9] J. Haglund, M. Haiman, N. Loehr, *A combinatorial formula for Macdonald polynomials*, J. Amer. Math. Soc. **18** (2005), no. 3, 735–761.
[10] K. Hikami, R. Inoue, Y. Komori, *Crystallization of the Bogoyavlensky lattice*, J. Phys. Soc. Japan **68** (1999), no. 7, 2234–2240.
[11] G. Hatayama, A. Kuniba, T. Takagi, *Soliton cellular automata associated with crystal bases*, Nuclear Phys. B **577** (2000), no. 3, 619–645.
[12] G. Hatayama, K. Hikami, R. Inoue, A. Kuniba, T. Takagi, T. Tokihiro, *The $A_M^{(1)}$ automata related to crystals of symmetric tensors*, J. Math. Phys. **42** (2001), no. 1, 274–308.
[13] G. Hatayama, A.N. Kirillov, A. Kuniba, M. Okado, T. Takagi, Y. Yamada, *Character formulae of $\widehat{\mathrm{sl}}_n$-modules and inhomogeneous paths*, Nuclear Phys. B **536** (1999), no. 3, 575–616.
[14] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, Z. Tsuboi, *Paths, crystals and fermionic formulae*, MathPhys odyssey, 2001, 205–272, Prog. Math. Phys., **23**, Birkhäuser Boston, Boston, MA, 2002.
[15] G. Hatayama, A. Kuniba, M. Okado, T. Takagi, Y. Yamada, *Remarks on fermionic formula*, Contemp. Math. **248** (1999) 243–291.
[16] M. Kashiwara, *Crystalizing the q-analogue of universal enveloping algebras*, Comm. Math. Phys. **133** (1990), no. 2, 249–260.
[17] M. Kashiwara, *On crystal bases*, Representations of groups (Banff, AB, 1994), 155–197, CMS Conf. Proc., 16, Amer. Math. Soc., Providence, RI, 1995.
[18] A.N. Kirillov, *New combinatorial formula for modified Hall-Littlewood polynomials*, Contemp. Math. **254** (2000) 283–333.
[19] R. Kedem, T.R. Klassen, B.M. McCoy, E. Melzer, *Fermionic quasi-particle representations for characters of $(G^{(1)})_1 \times (G^{(1)})_1/(G^{(1)})_2$*, Phys. Lett. B **304** (1993), no. 3-4, 263–270.
[20] R. Kedem, T.R. Klassen, B.M. McCoy, E. Melzer, *Fermionic sum representations for conformal field theory characters*, Phys. Lett. B **307** (1993), no. 1-2, 68–76.
[21] S. V. Kerov, A. N. Kirillov, N. Y. Reshetikhin, *Combinatorics, the Bethe ansatz and representations of the symmetric group* J. Soviet Math. **41** (1988), no. 2, 916–924.
[22] A. N. Kirillov, N. Y. Reshetikhin, *The Bethe Ansatz and the combinatorics of Young tableaux*, J. Soviet Math. **41** (1988) 925-955.
[23] A. Kuniba, M. Okado, R. Sakamoto, T. Takagi, Y. Yamada, private communication.
[24] A.N. Kirillov, M. Shimozono, *A generalization of the Kostka-Foulkes polynomials*, J. Algebraic Combin. **15** (2002), no. 1, 27–69.
[25] A. N. Kirillov, A. Schilling, M. Shimozono, *A bijection between Littlewood-Richardson tableaux and rigged configurations*, Selecta Math. (N.S.) **8** (2002), no. 1, 67–135.
[26] A. Lascoux, B. Leclerc, J.-Y. Thibon, *Ribbon tableaux, Hall-Littlewood functions, quantum affine algebras, and unipotent varieties*, J. Math. Phys. **38** (1997), no. 2, 1041–1068.
[27] B. Leclerc, J.-Y. Thibon, *Littlewood-Richardson coefficients and Kazhdan-Lusztig polynomials*, Adv. Stud. Pure Math. **28** (2000) 155–220.
[28] M. Okado, A. Schilling, M. Shimozono, *A crystal to rigged configuration bijection for nonexceptional affine algebras*, "Algebraic Combinatorics and Quantum Groups", Edited by N. Jing, World Scientific (2003), 85–124.
[29] M. Okado, A. Schilling, M. Shimozono, *Virtual crystals and fermionic formulas of type $D_{n+1}^{(2)}, A_{2n}^{(2)}$, and $C_n^{(1)}$*, Represent. Theory **7** (2003) 101–163.
[30] M. Okado, A. Schilling, M. Shimozono, *Virtual crystals and Kleber's algorithm*, Comm. Math. Phys. **238** (2003), no. 1-2, 187–209.
[31] A. Schilling, *q-supernomial coefficients: from riggings to ribbons*, MathPhys odyssey, 2001, 437–454, Prog. Math. Phys., **23**, Birkhäuser Boston, Boston, MA, 2002.
[32] A. Schilling, *A bijection between type $D_n^{(1)}$ crystals and rigged configurations*, J. Algebra **285** (2005) 292–334.
[33] A. Schilling, *Crystal structure on rigged configurations*, IMRN, to appear (math.QA/0508107).
[34] A. Schilling, M. Shimozono, *Fermionic formulas for level-restricted generalized Kostka polynomials and coset branching functions*, Commun. Math. Phys. **220** (2001) 105–164.
[35] A. Schilling, M. Shimozono, *$X = M$ for symmetric powers*, J. Algebra **295** (2006) 562–610.
[36] M. Shimozono, *Affine type A crystal structure on tensor products of rectangles, Demazure characters, and nilpotent varieties*, J. Algebraic Combin. **15** (2002) 151–187.
[37] J. R. Stembridge, *A local characterization of simply-laced crystals*, Transactions of the AMS **355** (2003) 4807–4823.
[38] T. Takagi, *Inverse scattering method for a soliton cellular automaton*, Nuclear Phys. B **707** (2005) 577–601.
[39] S.O. Warnaar, *The Bailey lemma and Kostka polynomials*, J. Algebraic Combin. **20** (2004), no. 2, 131–171.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, ONE SHIELDS AVENUE, DAVIS, CA 95616-8633, U.S.A.
E-mail address: anne@math.ucdavis.edu
URL: http://www.math.ucdavis.edu/˜anne

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Nonnegativity properties of the dual canonical basis

Mark Skandera

Abstract. Using Du's characterization of the dual canonical basis of the coordinate ring $\mathcal{O}(GL_n(\mathbb{C}))$, we show that all basis elements may be expressed in terms of immanants. We then give a new factorization of permutations avoiding the patterns 3412 and 4231, which in turn yields a factorization theorem for the corresponding Kazhdan-Lusztig basis of the Hecke algebra $H_n(q)$. Using this factorization, we show that for every totally nonnegative immanant $\mathrm{Imm}_f(x)$ and its expansion $\sum d_w \mathrm{Imm}_w(x)$ with respect to the basis of Kazhdan-Lusztig immanants, the coefficient d_w must be nonnegative when w avoids the patterns 3412 and 4231.

Résumé. En utilisant les résultats de Du, nous démontrons que chacque élement du base dual canonique de $\mathcal{O}(GL_n(\mathbb{C}))$, se peut réalise en terme d'immanants. Nous factorisent les permutations qui evitent le 3412 et le 4231, et aussi les élements du base de Kazhdan-Lusztig pour l'algébre de Hecke $H_n(q)$. En utilisant cette factorisation, nous montrons que pour chacque immanant totalement nonnegatif $\mathrm{Imm}_f(x)$ et l'expression $\sum d_w \mathrm{Imm}_w(x)$ en terme de base dual canonique, le coefficient d_w est nonnegatif quand w evite le 3412 et le 4231.

1. Introduction

Searching for solutions of the quantum Yang-Baxter equation, Drinfeld [**Dri85**] and Jimbo [**Jim85**] introduced a quantization $U_q(\mathfrak{sl}_n\mathbb{C})$ of the universal enveloping algebra $U(\mathfrak{sl}_n\mathbb{C})$. An explosion of mathematical research soon led to a quantization $\mathcal{O}_q(SL_n\mathbb{C})$ of the coordinate ring $\mathcal{O}(SL_n\mathbb{C})$, related by Hopf algebra duality to $U_q(\mathfrak{sl}_n\mathbb{C})$, and to a development of the representation theory of these algebras now known as quantum groups. In particular, Kashiwara [**Kas91**] and Lusztig [**Lus90**] discovered a *canonical* (or *crystal*) *basis* of $U_q(\mathfrak{sl}_n\mathbb{C})$ which has many interesting representation theoretic properties. The corresponding dual basis of $\mathcal{O}_q(SL_n\mathbb{C})$ is known as the *dual canonical basis* and is perhaps best understood as the projection of another *dual canonical basis* of the quantum polynomial ring $\mathbb{C}_q[x_{1,1},\dots,x_{n,n}]$. (See [**Du92**].) An elementary description of the canonical and dual canonical bases has been somewhat elusive, especially in the nonquantum ($q = 1$) setting.

In [**Lus94**] Lusztig proved that when we specialize $q = 1$, the elements of the dual canonical basis of $\mathbb{C}[x_{1,1},\dots,x_{n,n}]$ are *totally nonnegative* (TNN) polynomials in the following sense. We define a matrix to be *totally nonnegative* (TNN) if each of its minors is nonnegative. (See, e.g. [**FZ00**].) We define a polynomial $p(x) \in \mathbb{C}[x_{1,1},\dots,x_{n,n}]$ to be *totally nonnegative* (TNN) if for each $n \times n$ TNN matrix $A = (a_{i,j})$, we have

$$p(A) \underset{\mathrm{def}}{=} p(a_{1,1},\dots,a_{n,n}) \geq 0.$$

While it is not true that a polynomial is TNN only if it belongs to the dual canonical cone, we will show that certain coordinates of the polynomial with respect to the dual canonical cone must be nonnegative. Our criterion involves avoidance of the patterns 3412 and 4231 in permutations and thus links total nonnegativity to smoothness in Schubert varieties.

In Section 2 we will review Du's formulation of the dual canonical basis and show that these elements can be expressed in terms of functions called immanants. In Section 3 we will state a factorization theorem

2000 *Mathematics Subject Classification.* Primary 81R50; Secondary 20C08.
Key words and phrases. dual canonical basis, Kazhdan-Lusztig basis, total nonnegativity.

for 3412-avoiding, 4231-avoiding permutations and for the corresponding Kazhdan-Lusztig basis elements. In Section 4 we will use the factorization and immanant results to prove that for each TNN homogeneous element $p(x)$ of the coordinate ring $\mathcal{O}(SL_n(\mathbb{C}))$, certain coordinates with respect to the dual canonical basis must be nonnegative.

2. Kazhdan-Lusztig immanants and the dual canonical basis

The canonical bases of $\mathcal{O}(SL_n(\mathbb{C}))$ and $\mathcal{O}(GL_n(\mathbb{C}))$ may be obtained easily from a basis of the polynomial ring $\mathbb{C}[x_{1,1},\dots,x_{n,n}]$. We will call this basis too *the dual canonical basis*.

Before explicitly describing the dual canonical basis, let us look at a multigrading of $\mathbb{C}[x_{1,1},\dots,x_{n,n}]$ in terms of multisets. The polynomial ring has a traditional grading by degree,

$$\mathbb{C}[x] = \bigoplus_{r\geq 0} \mathcal{A}_r,$$

where $\mathcal{A}_r$ is the complex span of degree-r monomials. We may refine this grading by defining a multigrading of $\mathcal{A}_r$ indexed by pairs of r-element multisets. Let $\mathcal{M}(n,r)$ be the set of r-element multisets of n. Then we have

$$\mathcal{A}_r = \bigoplus_{M,M'\in\mathcal{M}(n,r)} \mathcal{A}_r(M,M'),$$

where we define a polynomial to be homogeneous of multidegree (M,M') if in each of its monomials, the multiset of row indices is M and the multiset of column indices is M'. For example, the polynomial $x_{1,1}x_{2,1}^2x_{3,3} - x_{1,1}x_{2,1}x_{2,3}x_{3,1}$ belongs to the component $\mathcal{A}_3(1223,1113)$ of $\mathbb{C}[x_{1,1},\dots,x_{3,3}]$.

Closely related to this multigrading are *generalized submatrices* of x. Given two r-element multisets $M = m_1\cdots m_r$, $M' = m'_1\cdots m'_r$ of $[n]$ (written as weakly increasing words), define the (M,M') *generalized submatrix* of x to be the matrix

$$x_{M,M'} = \begin{bmatrix} x_{m_1,m'_1} & x_{m_1,m'_2} & \cdots & x_{m_1,m'_r} \\ x_{m_2,m'_1} & x_{m_2,m'_2} & \cdots & x_{m_2,m'_r} \\ \vdots & \vdots & & \vdots \\ x_{m_r,m'_1} & x_{m_r,m'_2} & \cdots & x_{m_r,m'_r} \end{bmatrix}.$$

Letting $y = x_{M,M'}$, we see that for every permutation w in S_r, the monomial

$$y_{1,w(1)}\cdots y_{r,w(r)} = x_{m_1,m'_{w(1)}}\cdots x_{m_n,m'_{w(r)}} \tag{2.1}$$

belongs to $\mathcal{A}_r(M,M')$. We obtain the polynomial in the preceding paragraph from the matrix $y = x_{1223,1113}$ as $y_{1,1}y_{2,2}y_{3,3}y_{4,4} - y_{1,1}y_{2,2}y_{3,4}y_{4,3}$.

The multigrading is also closely related to parabolic subgroups of S_r as follows. Associate to M a subset $\iota(M)$ of the generators $\{s_1,\dots,s_{r-1}\}$ of S_r by

$$\iota(M) = \{s_j \,|\, m_j = m_{j+1}\}.$$

Let $I = \iota(M)$ and $J = \iota(M')$ be the subsets of generators of S_r corresponding to multisets M, M'. Letting the parabolic subgroups W_I and W_J act by left and right multiplication on all $r\times r$ matrices (restricting the defining representation of S_r to the parabolic subgroups), we see that $x_{M,M'}$ is fixed by this action.

The dual canonical basis of $\mathbb{C}[x_{1,1},\dots,x_{n,n}]$ consists of homogeneous elements with respect to the multigrading above. Du gives a formula for the elements of this basis in terms of the following polynomials $\widetilde{Q}_{u,w}(q)$ which are alternating sums of (inverse) Kazhdan-Lusztig polynomials,

$$\widetilde{Q}_{u,w}(q) = \sum_{\substack{v\in W_I wW_J \\ u\leq v\leq w}} (-1)^{\ell(w)-\ell(v)} P_{w_0v,w_0u}(q),$$

where u and w are maximal representatives of cosets in $W_I\backslash W/W_J$, and $\leq$ is the Bruhat order on S_r. These are generalizations of Deodhar's q-parabolic Kazhdan-Lusztig polynomials [**Deo91**], for when $I=\emptyset$ we have

$$\widetilde{Q}_{u,w}(q) = \widetilde{P}^J_{w_0ww'_0,w_0uw'_0}(q),$$

where w_0 and w'_0 are the longest elements of W and W_J, respectively.

We will express the dual canonical basis in terms of Kazhdan-Lusztig immanants $\{\mathrm{Imm}_u(x) \,|\, u \in S_n\}$ introduced in [**RS05a**],

$$\mathrm{Imm}_u(x) = \sum_{w \geq u} (-1)^{\ell(w)-\ell(u)} P_{w_0 w, w_0 u}(1) x_{1,w(1)} \cdots x_{n,w(n)},$$

and in terms of generalized submatrices as defined above.

Theorem 2.1. *Let M, M' be two r-element multisets of $[n]$. The nonzero polynomials in the set $\{\mathrm{Imm}_v(x_{M,M'}) \,|\, v \in S_r\}$ are the dual canonical basis of $\mathcal{A}_r(M, M')$. In particular, the permutations v corresponding to nonzero polynomials are maximal length representatives of double cosets in $W_{\iota(M)} \backslash W / W_{\iota(M')}$.*

Proof. Let $I = \iota(M)$, $J = \iota(M')$. By [**Du92**, Lem. 2.2], the canonical basis elements of $\mathcal{A}_r(M, M')$ are in bijective correspondence with cosets in $W_{\iota(M)} \backslash W / W_{\iota(M')}$, and each has the form

$$Z_u = \sum_{z \geq u} (-1)^{\ell(z')-\ell(u')} \widetilde{Q}_{u',z'}(1) x_{1,1}^{\alpha(z,1,1)} \cdots x_{i,j}^{\alpha(z,i,j)} \cdots x_{n,n}^{\alpha(z,n,n)},$$

where u, z are minimal representatives of double cosets in $W_I \backslash W / W_J$, u', z' are the respective maximal coset representatives, and

$$\alpha(z, i, j) = |\{z(k) \,|\, m_k = i\} \cap \{k \,|\, m'_k = j\}|.$$

It is straightforward to show that $u \leq z$ if and only if $u' \leq z'$ for any pairs (u, z) and (u', z') of minimal coset representatives and corresponding maximal coset representatives. (See [**HS05**] and references listed there.) We may therefore rewrite Du's description by summing over only maximal coset representatives,

$$Z_u = \sum_{z' \geq u'} (-1)^{\ell(z')-\ell(u')} \widetilde{Q}_{u',z'}(1) x_{1,1}^{\alpha(z,1,1)} x_{1,2}^{\alpha(z,i,j)} \cdots x_{n,n}^{\alpha(z,n,n)}.$$

Let $y = x_{M,M'}$. Then for any function $f : S_r \to \mathbb{C}$ we have

$$\mathrm{Imm}_f(y) = \sum_{w \in S_r} f(w) y_{1,w(1)} \cdots y_{n,w(n)}. \tag{2.2}$$

Since each permutation u in the double coset $W_I w W_J$ satisfies

$$y_{1,u(1)} \cdots y_{n,u(n)} = y_{1,w(1)} \cdots y_{n,w(n)},$$

we may sum over these double cosets,

$$\mathrm{Imm}_f(y) = \sum_{D \in W_I \backslash W / W_J} \left(\sum_{v \in D} f(v) \right) y_{1,w(1)} \cdots y_{n,w(n)},$$

where w is any representative of the double coset D. Note that $y_{i,w(i)} = x_{j,\ell}$ if $m_i = j$ and $m'_{w(i)} = \ell$. Thus the exponent of $x_{j,\ell}$ in $y_{1,w(1)} \cdots y_{n,w(n)}$ is equal to the number of indices i which satisfy

$$m_i = j, \quad m'_{w(i)} = \ell.$$

Since this is just $\alpha(w, i, j)$, we have

$$y_{1,w(1)} \cdots y_{n,w(n)} = x_{1,1}^{\alpha(w,1,1)} \cdots x_{n,n}^{\alpha(w,n,n)}.$$

Now consider the function $f_u : v \mapsto (-1)^{\ell(v)-\ell(u)} P_{w_0 v, w_0 u}(1)$ and the corresponding immanant of y, $\mathrm{Imm}_u(y) = \mathrm{Imm}_{f_u}(y)$. If u is not a maximal representative of a double coset in $W_I \backslash W / W_J$, then by [**Cur85**, Thm. 1.2] we have $su > u$ for some transposition s in I, or we have $us > u$ for some transposition s in J. By [**RS05a**, Cor. 6.4] either of these conditions implies that $\mathrm{Imm}_u(y) = 0$. Suppose therefore that u' is a maximal coset representative. Then by (2.2) we have

$$\mathrm{Imm}_{u'}(y) = \sum_{D \in W_I \backslash W / W_J} \left(\sum_{\substack{v \in D \\ v \geq u'}} (-1)^{\ell(v)-\ell(u')} P_{w_0 v, w_0 u'}(1) \right) x_{1,1}^{\alpha(w',1,1)} \cdots x_{n,n}^{\alpha(w',n,n)},$$

where w' is the maximal representative of D, and we include the inequality $v \geq u'$ because the number $P_{w_0 v, w_0 u'}(1)$ is zero otherwise. For each coset D and its maximal representative w', the inner sum is equal to

$$(-1)^{\ell(w')-\ell(u')} \sum_{\substack{v \in W_I w' W_J \\ u' \leq v \leq w'}} (-1)^{\ell(w')-\ell(v)} P_{w_0 v, w_0 u'}(1) = (-1)^{\ell(w')-\ell(u')} \widetilde{Q}_{u',w'}(1),$$

and we have

$$\mathrm{Imm}_{u'}(y) = \sum_{D \in W_I \backslash W / W_J} (-1)^{\ell(w')-\ell(u')} \widetilde{Q}_{u',w'}(1) x_{1,1}^{\alpha(w,1,1)} \cdots x_{n,n}^{\alpha(w,n,n)}.$$

Note that for any double coset whose maximal representative w' satisfies $u' \leq w'$, we have $\widetilde{Q}_{u',w'}(1) = 0$ and the contribution to the sum is zero. The sum therefore may be taken over double cosets D whose maximal element w' satisfies $w' \geq u'$, and we have

$$Z_u = \mathrm{Imm}_{u'}(x_{M,M'}),$$

as desired. □

Quantizing the Kazhdan-Lusztig immanants by

$$\mathrm{Imm}_v(x;q) = \sum_{w \geq v} (-q^{-1/2})^{\ell(w)-\ell(v)} \widetilde{Q}_{v,w}(q) x_{1,w(1)} \cdots x_{n,w(n)},$$

one constructs the quantum dual canonical basis of $\mathcal{A}_r(M,M')$ by taking all of the polynomials $(q^{1/2})^{\ell(w_0^J)-\ell(w_0^I)} \mathrm{Imm}_v(x_{M,M'};q$ where $I = \iota(M)$, $J = \iota(M')$, v is a maximal length coset representative in $W_I \backslash W / W_J$, and w_0^I, w_0^J are the maximal length elements of W_I, W_J. Details will appear in [**Ska05**]. (See [**Bru05**], [**Du92**] for other descriptions of this basis.)

Letting $\mathbf{B}$ be the dual canonical basis of $\mathbb{C}[x_{1,1}, \ldots, x_{n,n}]$, we have the following formulas for the dual canonical bases of the coordinate rings of $GL_n(\mathbb{C})$ and $SL_n(\mathbb{C})$. The dual canonical basis of

$$\mathcal{O}(GL(n,\mathbb{C})) \cong \mathbb{C}[x_{1,1}, \ldots, x_{n,n}, t]/(\det(x)t - 1)$$

is obtained by dividing elements of $\mathbf{B}$ by powers of the determinant,

$$\cup_{r \geq 0} \cup_{(M,M') \in \mathcal{M}(n,r)} \{\mathrm{Imm}_w(x) \det(x)^{-k} \,|\, k \geq 0; w \text{ maximal in } W_{\iota M} \backslash S_r / W_{\iota(M')}\}.$$

The dual canonical basis of

$$\mathcal{O}(SL(n,\mathbb{C})) \cong \mathbb{C}[x_{1,1}, \ldots, x_{n,n}]/(\det(x) - 1)$$

is obtained by projecting $\mathbb{C}[x_{1,1}, \ldots, x_{n,n}]$ or $\mathcal{O}(GL(n,\mathbb{C}))$ onto $\mathcal{O}(SL(n,\mathbb{C}))$.

3. A Factorization Theorem

While each nonnegative linear combination of dual canonical basis elements is a totally nonnegative polynomial, the converse of this statement is false. Intimately related to this fact is the vector space duality between the component $\mathcal{A}_n([n],[n])$ of the polynomial ring $\mathbb{C}[x_{1,1}, \ldots, x_{n,n}]$ and and the group algebra $\mathbb{C}[S_n]$, defined by

$$\langle x_{1,u(1)} \cdots x_{n,u(n)}, T_v \rangle = \delta_{u,v}.$$

In particular, Kazhdan and Lusztig [**KL79**] defined a basis $\{C'_w(q) \,|\, w \in S_n\}$ of the Hecke algebra $H_n(q)$ by

$$C'_v(q) = q^{-1/2} \sum_{u \leq v} P_{u,v}(q) T_u,$$

where $\{P_{u,v}(q) \,|\, u, v \in S_n\}$ are certain polynomials for which no elementary formula is known. Dual to this basis is the basis of Kazhdan-Lusztig immanants,

$$\langle \mathrm{Imm}_u(x), C'_v(1) \rangle = \delta_{u,v}.$$

Since no elementary formula is known for the Kazhdan-Lusztig polynomials, it is not surprising that we also have no elementary formula for the Kazhdan-Lusztig basis of the Hecke algebra or for the Kazhdan-Lusztig immanants. Nevertheless, we can deduce certain properties of the Kazhdan-Lusztig immanants by studying Kazhdan-Lusztig basis elements which have a rather simple form and others which factor as products of these. The basis elements we shall consider correspond to permutations whose one-line notations avoid certain patterns. The factorization of these basis elements closely resembles the factorization of the corresponding permutations.

Given a word $u = u_1 \cdots u_k$ on a totally ordered alphabet and a permutation v in S_k with one-line notation $v_1 \cdots v_k$, we will say that u *matches the pattern* v if the letters of u appear in the same relative order as those of v. We will also say that u_1 *matches the* v_1, u_2 matches the v_2, etc. For example, a word $u_1u_2u_3$ with $u_2 < u_3 < u_1$ matches the pattern 312, with u_1 matching the 3, u_2 matching the 1, and u_3 matching the 2.

We will say that a permutation w in S_n *avoids the pattern* v if no subword $w_{i_1} \cdots w_{i_k}$ with $i_1 < \cdots < i_k$ matches the pattern v. We will also call such a permutation v*-avoiding.* In particular, we will be interested in permutations which avoid the patterns 3412 and 4231. Note that a permutation w avoids these patterns if and only if w^{-1} does, since the patterns are involutions. In particular, corresponding to each adjacent transposition s_i is the basis element $C'_{s_i}(q) = q^{-1/2}(T_e + T_{s_i})$, and we have the following factorization result of Billey and Warrington [**BW01**].

THEOREM 3.1. *Let $s_{i_1} \cdots s_{i_\ell}$ be a reduced expression for w. Then we have*

$$C'_w(q) = C'_{s_{i_1}}(q) \cdots C'_{s_{i_\ell}}(q)$$

if and only if the one-line notation for w avoids the patterns 321*,* 56781234*,* 46781235*,* 56718234*,* 46718235*.*

Other permutations w for which $C'_w(q)$ has a particularly nice form are known as *reversals.* Write $s_{[i,j]}$ for the permutation which fixes indices $1, \ldots, i-1, j+1, \ldots, n$ and reverses the remaining indices. Corresponding to reversals are the Kazhdan-Lusztig basis elements

$$C'_{s_{[i,j]}}(q) = (q^{-1/2})^{\binom{j-i+1}{2}} \sum_{v \leq s_{[i,j]}} T_v$$

We will show in Theorem 3.3 that permutations which avoid the patterns 3412 and 4231 factor as products of these basis elements.

To begin, we define the map $\oplus : S_n \times S_m \to S_{n+m}$, as is somewhat customary, by

$$s_{i_1} \cdots s_{i_\ell} \oplus s_{j_1} \cdots s_{j_k} = s_{i_1} \cdots s_{i_\ell} s_{j_1+n} \cdots s_{j_k+n}.$$

OBSERVATION 3.2. *If u and v are* 3412*-avoiding,* 4231*-avoiding permutations in S_m and S_n, then $u \oplus v$ is a* 3412*-avoiding,* 4231*-avoiding permutation in S_{m+n}.*

We will say that a permutation w has an *irreducible zig-zag factorization* if there exist a positive integer r, a sequence of nonnegative integers

$$j_1, \ldots, j_r, k_1, \ldots, k_r,$$

all odd except possibly for j_1 and k_r which may also be zero, and a sequence of intervals

$$\begin{aligned}(3.1) \quad & a_0, b_{1,1}, \ldots, b_{1,j_1}, a_1, c_{1,1}, \ldots, c_{1,k_1}, d_1, \\ & \qquad \ldots, b_{i,1}, \ldots, b_{i,j_i}, a_i, c_{i,1}, \ldots, c_{i,k_i}, d_i, \ldots, \\ & \qquad\qquad b_{r,1}, \ldots, b_{r,j_r}, a_r, c_{r,1}, \ldots, c_{r,k_r}, d_r,\end{aligned}$$

all nonempty except possibly for a_0, a_r, such that w is equal to the product of the reversals on these intervals in the order listed,

$$w = s_{a_0} \cdots s_{d_r},$$

and the endpoints of the intervals, which we denote by

$$a_i = [\lambda(a_i), \rho(a_i)], \quad b_{i,j} = [\lambda(b_{i,j}), \rho(b_{i,j})], \quad c_{i,k} = [\lambda(c_{i,k}), \rho(c_{i,k})], \quad d_i = [\lambda(d_i), \rho(d_i)],$$

satisfy the following conditions.

(1) $j_1 = 0$ if and only if $a_0 = s_\emptyset$.
(2) $k_r = 0$ if and only if $a_r = s_\emptyset$.
(3) For each i satisfying $a_{i-1} \neq s_\emptyset$ we have

$$\begin{aligned}\lambda(a_{i-1}) < \lambda(b_{i,1}) &= \lambda(b_{i,2}) < \lambda(b_{i,3}) = \cdots < \lambda(b_{i,j_i}) = \lambda(d_i), \\ \rho(a_{i-1}) = \rho(b_{i,1}) &< \rho(b_{i,2}) = \rho(b_{i,3}) < \cdots = \rho(b_{i,j_i}) < \rho(d_i).\end{aligned}$$

(4) For each i satisfying $a_i \neq s_\emptyset$ we have

(5) and

$$\lambda(a_i) = \lambda(c_{i,1}) > \lambda(c_{i,2}) = \lambda(c_{i,3}) > \cdots = \lambda(c_{i,k_i}) > \lambda(d_i),$$
$$a_i' > \rho(c_{i,1}) = \rho(c_{i,2}) > \rho(c_{i,3}) = \cdots > \rho(c_{i,k_i}) = \rho(d_i),$$

(6) For $i = 1, \ldots, r$ we have
$$\rho(b_{i,j_i}) < \lambda(c_{i,k_i}).$$

(7) For $i = 1, \ldots, r-1$ we have
$$\rho(c_{i,1}) < \rho(b_{i+1,1}).$$

Note that the intervals
$$\{a_i \mid 0 \leq i \leq r\} \cup \{b_{i,j} \mid 1 \leq i \leq r, j \text{ even }\} \cup \{c_{i,j} \mid 1 \leq i \leq r, j \text{ even }\} \cup \{d_i \mid 1 \leq i \leq r\}$$
have length at least two when they are nonempty.

Note that the lexicographic order on the set (3.1) of intervals (padded with zeros at the end) is

$$a_0, b_{1,1}, \ldots, b_{1,j_1}, d_1, c_{1,k_1}, \ldots, c_{1,1}, a_1, \ldots, b_{i,1}, \ldots, b_{i,j_i}, d_i, c_{i,k_i}, \ldots, c_{i,k_1}, a_i, \ldots, b_{r,1}, \ldots, b_{r,j_r}, d_r, c_{r,1}, \ldots, c_{r,k_r}, a_r, \tag{3.2}$$

If in this factorization we have $r = 1$ and $k_1 = 0$, then the only intervals to appear are
$$a_0, b_{1,1}, \ldots, b_{1,j_1}, d_1,$$
(with $a_0, b_{1,1}, \ldots, b_{1,j_1}$ not appearing if $j_1 = 0$) and we will say that the irreducible zig-zag factorization is *lexicographically increasing*. Given a reversal factorization $W_1 \cdots W_p$ in which each W_i is an irreducible zigzag factorization and the intervals don't overlap, we will call this a *zig-zag factorization*. Note that each permutation u possessing an irreducible zig-zag factorization decomposes as
$$u = e \oplus \cdots \oplus e \oplus v \oplus e \oplus \cdots \oplus e,$$
where e is the identity element of S_1, and v possesses the same irreducible zig-zag factorization as u.

Proposition 3.1. *A permutation avoids the patterns* 3412*,* 4231 *if and only if it has a zig-zag factorization.*

Proof. Omitted. □

Two examples of 3412-avoiding, 4231-avoiding permutations and zig-zag factorizations are
$$654213 = s_{[1,5]} s_{[3,5]} s_{[3,6]}, \qquad 621354 = s_{[1,3]} s_{[3,4]} s_{[4,6]}.$$

Proposition 3.2. *Let $s_{I_1} \cdots s_{I_p}$ be a zig-zag factorization of $w \in S_n$, let $(t_1, \ldots, t_p)$ be a subexpression of this factorization, and define the permutation $u = t_1 \cdots t_p$. Then $u \leq w$ in the Bruhat order.*

Proof. Omitted. □

The above factorization results for permutations translate into the following factorization result for Kazhdan-Lusztig basis elements. Given a sequence of intervals $I = (I_1, \ldots, I_r)$, define the $H_n(q)$ algebra element
$$\Phi(I_1, \ldots, I_r; q) = C'_{s_{I_1}}(q) \cdots C'_{s_{I_r}}(q).$$

Theorem 3.3. *Let w avoid the patterns* 3412 *and* 4231 *and have zig-zag factoriztion (3.1), define the sequence of intervals*

$$I = (a_0, b_{1,2}, b_{1,4}, \ldots, b_{1,j_1-1}, a_1, c_{1,2}, c_{1,4}, \ldots, c_{1,k_1-1}, d_1, \ldots, b_{i,2}, \ldots, b_{i,j_i-1}, a_i, c_{i,2}, \ldots, c_{i,k_i-1}, d_i, \ldots, b_{r,2}, \ldots, b_{r,j_r-1}, a_r, c_{r,2}, \ldots, c_{r,k_r-1}, d_r) \tag{3.3}$$

and define the number
$$\gamma = \prod_{i=1}^{r} \prod_{\substack{j=1 \\ j \text{ odd}}} |b_{i,j}|! \prod_{\substack{k=1 \\ k \text{ odd}}} |c_{i,k}|!.$$

Then the Kazhdan-Lusztig basis element $C'_w(1)$ factors as

$$C'_w(1) = \tfrac{1}{\gamma}\Phi(I;1),$$

and the Kazhdan-Lusztig basis element $C'_w(q)$ factors as

$$C'_w(q) = q^{-\ell(w)/2}C'_w(1),$$

or equivalently

$$C'_w(q) = \tfrac{1}{\gamma}q^{\delta/2}\Phi(I;q),$$

where

$$\delta = \sum_{i=0}^{r}\binom{|a_i|}{2} + \sum_{i=0}^{r}\binom{|d_i|}{2} + \sum_{\substack{j=1\\ j\ even}}\binom{|b_{i,j}|}{2} + \sum_{\substack{k=1\\ k\ even}}\binom{|c_{i,k}|}{2} - \ell(w).$$

Proof. Omitted. □

Corresponding to previoius examples of $3412-$*avoiding*,$4231-$*avoiding* permutations and zig-zag factorizations are the factorizations of Kazhdan-Lusztig basis elements,

$$654213 = s_{[1,5]}s_{[3,5]}s_{[3,6]}, \qquad 621354 = s_{[1,3]}s_{[3,4]}s_{[4,6]},$$

$$C'_{654213}(q) = C'_{s_{[1,5]}}(q)\frac{q^{3/2}}{3_q!}C'_{s[3,6]}(q), \qquad C'_{621354}(q) = C'_{s_{[1,3]}}(q)C'_{s[3,4]}(q)C'_{s[4,6]}(q).$$

4. The dual cone of total nonnegativity

In [**RS05a**, Sec. 7], we have cones of TNN and SNN elements of $\mathrm{span}_{\mathbb{C}}\{x_{1,w(1)}\cdots x_{n,w(n)}\,|\,w\in S_n\}$ were defined. Virtually all of the known TNN and SNN polynomials belong to these cones. (See [**RS05a**, Sec. 1].) Generalizing these definitions a bit, we will define the following cones of functions on $n\times n$ matrices. Define the *dual canonical cone*, the *dual cone of total nonnegativity*, and the *dual cone of Schur nonnegativity*, which we will denote by $\check{C}_B$, $\check{C}_{\mathrm{TNN}}$, and $\check{C}_{\mathrm{SNN}}$, respectively, to be the cones whose extreme rays are homogeneous elements of $\mathbb{C}[x_{1,1},\ldots,x_{n,n}]$ belonging to B, having the TNN property, and having the SNN property, respectively. Our use of the term *dual* refers to the relationship of this point of view to that of Stembridge [**Ste92**], who define the *cone of total nonnegativity* to be the smallest cone in $\mathbb{C}[S_n]$ containing all of elements of the form $\sum_{w\in S_n} a_{1,w(1)}\cdots a_{n,w(n)}$, where $A=(a_{i,j})$ is a totally nonnegative matrix.

Using this terminology, we have the following.

Corollary 4.1. *The dual canonical cone is contained in the intersection of the dual cones of total nonnegativity and Schur nonnegativity.*

Proof. The main results of [**RS05a**] and [**RS05b**] show that Kazhdan-Lusztig immanants of generalized submatrices of $x=(x_{i,j})_{i,j=1}^n$ are TNN and SNN. Since the cone generated by these functions is $\check{C}_B$, we have the desired result. □

The author and A. Zelevinsky have verified that the containment of $\check{C}_B$ in $\check{C}_{\mathrm{TNN}}$ is strict. In particular, the homogeneous element

$$\mathrm{Imm}_{3214}(x) + \mathrm{Imm}_{1432}(x) - \mathrm{Imm}_{3412}(x) \tag{4.1}$$

belongs to $\check{C}_{\mathrm{TNN}} \smallsetminus \check{C}_B$. Moreover we have used cluster algebras and Maple to show that this element is equal to a subtraction-free rational expression in matrix minors. Thus the cone of functions which have this subtraction-free rational function (SFR) property must also properly contain $\check{C}_B$. On the other hand, the element (4.1) does not belong to $\check{C}_{\mathrm{SNN}}$, for its evaluation on the Jacobi-Trudi matrix H_{2222} expands in the Schur basis as

$$2s_{62} + 2s_{53} + 2s_{521} - s_{44} + 2s_{431} + 2s_{422}.$$

Thus $\check{C}_B$ and $\check{C}_{\mathrm{SNN}}$ are not known to be different. Let us examine the difference $\check{C}_{\mathrm{TNN}} \smallsetminus \check{C}_B$ more closely.

THEOREM 4.2. *Let H be the planar network corresponding to a zig-zag factorization of a* 3412*-avoiding,* 4231*-avoiding permutation w in S_n, and let A_w be the path matrix of H. Then there exists a nonnegative integer c such that we have*

$$\mathrm{Imm}_v(A_w) = \begin{cases} c & \text{if } v = w, \\ 0 & \text{otherwise.} \end{cases}$$

PROOF. Omitted. □

The existence of the matrices specified by the previous theorem allows us to compare the dual canonical cone with the dual cone of total nonnegativity as follows.

THEOREM 4.3. *Let $\mathrm{Imm}_f(x)$ be totally nonnegative and let its expansion in terms of Kazhdan-Lusztig immanants be given by*

$$\mathrm{Imm}_f(x) = \sum_{w \in S_n} d_w \mathrm{Imm}_w(x).$$

Then c_u is nonnegative for each 3412*-avoiding,* 4231*-avoiding permutation u.*

PROOF. Let u be a 3412-avoiding, 4231-avoiding permutation in S_n, and suppose that d_u is negative. Let G_u be the planar network corresponding to the reversal factorization of u, and let A_u be the path matrix of G_u. Then we have

$$\mathrm{Imm}_f(A_u) = cd_u < 0,$$

contradicting the total nonnegativity of $\mathrm{Imm}_f(x)$. □

Theorem 4.3 suggests several problems. Recalling Lakshmibai and Sandhya's result [**LS90**] that a permutation w's avoidance of the patterns 3412 and 4231 is equivalent to smoothness of the Schubert variety Γ_w, we have the following.

PROBLEM 4.4. Find an intuitive reason for the connection between total nonnegativity, the dual canonical basis, and smoothness of Schubert varieties.

It would also be interesting to understand precisely how the cones mentioned earlier are related.

PROBLEM 4.5. Find the extremal rays of $\check{C}_{\mathrm{TNN}}$, $\check{C}_{\mathrm{SNN}}$, and the cone of SFR functions in $\mathbb{C}[x_{1,1}, \ldots, x_{n,n}]$, or describe the containments satisfied by these cones.

Since the factorizations given in Theorems 3.1 and 3.3 agree on permutations which avoid all seven of the forbidden patterns, the author believes that there is a simple generalization of the two results. It would be interesting to understand in even greater generality which elements of the Kazhdan-Lusztig basis factor as products of others.

5. Acknowledgements

The author is grateful to Francesco Brenti, Jie Du, Sergey Fomin, Christophe Hohlweg, Federico Incitti, Bernard Leclerc, and Andrei Zelevinsky for helpful conversations, to Anders Björner and Richard Stanley for organizing a special program in algebraic combinatorics at the Institut Mittag-Leffler in Djursholm, Sweden, and to the Institut staff for their hospitality.

References

[Bru05] J. Brundan, *Dual canonical bases and Kazhdan-Lusztig polynomials*, Preprint math.CO/0509700 on ArXiv, 2005.

[BW01] S. C. Billey and G. Warrington, *Kazhdan-Lusztig polynomials for 321-hexagon avoiding permutations*, J. Algebraic Combin. **13** (2001), no. 2, 111–136.

[Cur85] C. W. Curtis, *On Lusztig's isomorphism theorem for Hecke algebras*, J. Algebra **92** (1985), no. 2, 348–365.

[Deo91] V. Deodhar, *Duality in parabolic set up for questions in Kazhdan-Lusztig theory*, J. Algebra **142** (1991), no. 1, 201–209.

[Dri85] V. G. Drinfeld, *Hopf algebras and the quantum Yang-Baxter equation*, Dokl. Akad. Nauk. SSSR **283** (1985), no. 5, 1060–1064.

[Du92] J. Du, *Canonical bases for irreducible representations of quantum* GL_n, Bull. London Math. Soc. **24** (1992), no. 4, 325–334.

[FZ00] S. Fomin and A. Zelevinsky, *Total positivity: Tests and parametrizations*, Math. Intelligencer **22** (2000), no. 1, 23–33.

[HS05] C. Hohlweg and M. Skandera, *A Note on Bruhat order and double coset representatives*, Preprint math.CO/0511611 on ArXiv (2005).

[Jim85] M. Jimbo, *A q-difference analogue of $U(\mathfrak{g})$ and the Yang-Baxter equation*, Lett. Math. Phys. **10** (1985), no. 1, 63–69.

[Kas91] M. Kashiwara, *On crystal bases of the Q-analog of universal enveloping algebras*, Duke Math. J. **63** (1991), 465–516.
[KL79] D. Kazhdan and G. Lusztig, *Representations of Coxeter groups and Hecke algebras*, Inv. Math. **53** (1979), 165–184.
[LS90] V. Lakshmibai and B. Sandhya, *Criterion for smoothness of Schubert varieties in $SL(n)/B$*, Proc. Indian Acad. Sci. (Math Sci.) **100** (1990), no. 1, 45–52.
[Lus90] G. Lusztig, *Canonical bases arising from quantized enveloping algebras*, J. Amer. Math. Soc. **3** (1990), 447–498.
[Lus94] ———, *Total positivity in reductive groups*, Lie Theory and Geometry: in Honor of Bertram Kostant, Progress in Mathematics, vol. 123, Birkhäuser, Boston, 1994, pp. 531–568.
[RS05a] B. Rhoades and M. Skandera, *Kazhdan-Lusztig immanants and products of matrix minors*, To appear in *J. Algebra*, 2005.
[RS05b] ———, *Temperley-Lieb immanants*, To appear in *Ann. Comb.*, 2005.
[Ska05] M. Skandera, *Double parabolic Kazhdan-Lusztig polynomials*, In preparation, 2005.
[Ste92] J. Stembridge, *Some conjectures for immanants*, Can. J. Math. **44** (1992), no. 5, 1079–1099.

Department of Mathematics, Haverford College, 370 Lancaster Avenue, Haverford, PA, 19041
E-mail address: mark.skandera@gmail.com
URL: http://www.haverford.edu/math/skandera.html

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

On the characteristic map of finite unitary groups

Nathaniel Thiem and C. Ryan Vinroot

Abstract. In his classic book on symmetric functions, Macdonald describes a remarkable result by Green relating the character theory of the finite general linear group to transition matrices between bases of symmetric functions. This connection allows us to analyze the representation theory of the general linear group via symmetric group combinatorics. Using the work of Ennola, Kawanaka, Lusztig and Srinivasan, this paper describes the analogous setting for the finite unitary group. In particular, we explain the connection between Deligne-Lusztig theory and Ennola's efforts to generalize Green's work, and from this we deduce various representation theoretic results. Applications include finding certain sums of character degrees, and a model of Deligne-Lusztig type for the finite unitary group, which parallels results of Klyachko and Inglis and Saxl for the finite general linear group.

Résumé. Dans son livre classique sur les fonctions symétriques, Macdonald décrit un résultat remarquable dû à Green, qui relie la théorie des caractères du groupe général liéaire fini, aux matrices de transition entre bases de fonctions symétriques. Cette connexion permet d'analyser la théorie de représentation du groupe général linéaire à l'aide de combinatoires de groupes symétriques. En utilisant le travail d'Ennola, Kawanaka, Lusztig et Srinivasan, le présent article décrit le cadre analogue pour le groupe unitaire fini. En particulier, nous expliquons la connexion entre la théorie de Deligne-Lusztig et les efforts d'Ennola concernant la généralisation du travail de Green, et nous en déduisons plusieurs résultats en théorie de représentation. Parmi les applications, nous obtenons certaines sommes de degés de caractères, et un modèle du type Deligne-Lusztig pour le groupe unitaire fini, qui met en parallèle les résultats de Klyachko, Inglis et Saxl pour le groupe général linéaire fini.

1. Introduction

In his seminal work [**7**], Green described a remarkable connection between the class functions of the finite general linear group $\mathrm{GL}(n, \mathbb{F}_q)$ and a generalization of the ring of symmetric functions of the symmetric group S_n. In particular, Green defines a map, called the characteristic map, that takes irreducible characters to Schur-like symmetric functions, and recovers the character table of $\mathrm{GL}(n, \mathbb{F}_q)$ as the transition matrix between these Schur functions and Hall-Littlewood polynomials [**14**, Chapter IV]. Thus, we can use the combinatorics of the symmetric group S_n to understand the representation theory of $\mathrm{GL}(n, \mathbb{F}_q)$. Some of the implications of this approach include an indexing of irreducible characters and conjugacy classes of $\mathrm{GL}(n, \mathbb{F}_q)$ by multi-partitions and a formula for the degrees of the irreducible characters in terms of these partitions.

This paper describes the parallel story for the finite unitary group $\mathrm{U}(n, \mathbb{F}_{q^2})$ by collecting known results for this group and examining some applications of the unitary characteristic map. Inspired by Green, Ennola [**4, 5**] used results of Wall [**16**] to construct the appropriate ring of symmetric functions and characteristic map. Ennola was able to prove that the analogous Schur-like functions correspond to an orthonormal basis for the class functions, and conjectured that they corresponded to the irreducible characters. He theorized that the representation theory of $\mathrm{U}(n, \mathbb{F}_{q^2})$ should be deduced from the representation theory of $\mathrm{GL}(n, \mathbb{F}_q)$ by

2000 *Mathematics Subject Classification.* Primary 20C33; Secondary 05E05.

Key words and phrases. finite unitary group, symmetric functions, models of finite groups.

The first author would like to thank Stanford University for supplying a stimulating research environment, and, in particular, P. Diaconis for acting as a sounding board for many of these ideas.

The second author thanks the Tata Institute of Fundamental Research for an excellent environment for doing much of this work, and Rod Gow for helpful correspondence.

substituting "$-q$" for every occurrence of "q". The general phenomenon of obtaining a polynomial invariant in q for $\mathrm{U}(n,\mathbb{F}_{q^2})$ by this substitution has come to be known as "Ennola duality".

Roughly a decade after Ennola made his conjecture, Deligne and Lusztig [**2**] constructed a family of virtual characters, called Deligne-Lusztig characters, to study the representation theory of arbitrary finite reductive groups. Lusztig and Srinivasan [**13**] then computed an explicit decomposition of the irreducible characters of $\mathrm{U}(n,\mathbb{F}_{q^2})$ in terms of Deligne-Lusztig characters. Kawanaka [**11**] used this composition to demonstrate that Ennola duality applies to Green functions, thereby improving results of Hotta and Springer [**9**] and finally proving Ennola's conjecture.

This paper begins by describing some of the combinatorics and group theory associated with the finite unitary groups. Section 2 defines the finite unitary groups, outlines the combinatorics of multi-partitions, and gives a description of some of the key subgroups. Section 2.4 analyzes the conjugacy classes of $\mathrm{U}(n,\mathbb{F}_{q^2})$ and the Jordan decomposition of these conjugacy classes.

Section 3 outlines the statement and development of the Ennola conjecture from two perspectives. Both points of view define a map from a ring of symmetric functions to the character ring C of $\mathrm{U}(n,\mathbb{F}_{q^2})$. However, the first uses the multiplication for C as defined by Ennola, and the second uses Deligne-Lusztig induction as the multiplicative structure of C. This multiplicative structure on the graded ring of characters of the unitary group was studied by Digne and Michel in [**3**], where the focus is that this multiplication induces a Hopf algebra structure. While some of the results in Section 3 appear in a different form in [**3**], our approach focuses on the explicit map between characters and symmetric functions.

The main results are

I. (Theorem 3.2) The Deligne-Lusztig characters correspond to power-sum symmetric functions via the characteristic map of Ennola.

II. (Corollary 3.2) The multiplicative structure that Ennola defined on C is Deligne-Lusztig induction.

Section 4 computes the degrees of the irreducible characters, and uses this result to evaluate various sums of character degrees (see [**14**, IV.6, Example 5] for the $\mathrm{GL}(n,\mathbb{F}_q)$ analogue of this method). The main results are

III. (Theorem 4.1) An irreducible $\chi^{\boldsymbol{\lambda}}$ character of $\mathrm{U}(m,\mathbb{F}_{q^2})$ corresponds to

$$(-1)^{\lfloor m/2\rfloor+n(\boldsymbol{\lambda})}s_{\boldsymbol{\lambda}}\quad\text{and}\quad \chi^{\boldsymbol{\lambda}}(1)=q^{n(\boldsymbol{\lambda}')}\frac{\displaystyle\prod_{1\leq i\leq m}(q^i-(-1)^i)}{\displaystyle\prod_{\Box\in\boldsymbol{\lambda}}(q^{\mathbf{h}(\Box)}-(-1)^{\mathbf{h}(\Box)})},$$

where $s_{\boldsymbol{\lambda}}$ is a Schur-like function, and both $n(\boldsymbol{\lambda})$ and $\mathbf{h}(\Box)$ are combinatorial statistics on the multi-partition $\boldsymbol{\lambda}$.

IV. (Corollary 4.2) If $\mathcal{P}_n^{\Theta}$ indexes the irreducible characters $\chi^{\boldsymbol{\lambda}}$ of $\mathrm{U}(n,\mathbb{F}_{q^2})$, then

$$\sum_{\boldsymbol{\lambda}\in\mathcal{P}_n^{\Theta}}\chi^{\boldsymbol{\lambda}}(1)=|\{g\in\mathrm{U}(n,\mathbb{F}_{q^2})\ \mid\ g\text{ symmetric}\}|.$$

V. (Theorem 4.3) We give a subset $\mathcal{X}\subseteq\mathcal{P}_{2n}^{\Theta}$ such that

$$\sum_{\boldsymbol{\lambda}\in\mathcal{X}}\chi^{\boldsymbol{\lambda}}(1)=(q+1)q^2(q^3+1)\cdots q^{2n-2}(q^{2n-1}+1)=\frac{|\mathrm{U}(2n,\mathbb{F}_{q^2})|}{|\mathrm{Sp}(2n,\mathbb{F}_q)|}.$$

Section 5 uses results by Ohmori [**15**] and Henderson [**8**] to adapt a model for the general linear group, found by Klyachko [**12**] and Inglis and Saxl [**10**], to the finite unitary group. The main result is

VI. (Theorem 5.2) Let $U_m=\mathrm{U}(m,\mathbb{F}_{q^2})$, where q is odd, and let Γ_m be the Gelfand-Graev character of U_m, $\mathbf{1}$ be the trivial character of the finite symplectic group $Sp_{2r}=\mathrm{Sp}(2r,\mathbb{F}_q)$, and R_L^G be the Deligne-Lusztig induction functor. Then

$$\sum_{0\leq 2r\leq m}R_{U_{m-2r}\oplus U_{2r}}^{U_m}\left(\Gamma_{m-2r}\otimes\mathrm{Ind}_{Sp_{2r}}^{U_{2r}}(\mathbf{1})\right)=\sum_{\boldsymbol{\lambda}\in\mathcal{P}_m^{\Theta}}\chi^{\boldsymbol{\lambda}}.$$

That is, in the theorem of Klyachko, one may replace parabolic induction by Deligne-Lusztig induction to obtain a theorem for the unitary group.

These results give considerable combinatorial control over the representation theory of the finite unitary group, and there are certainly more applications to these results than what we present in this paper. Furthermore, this characteristic map gives some insight as to how a characteristic map might look in general type, using the invariant rings of other Weyl groups.

2. Preliminaries

2.1. The unitary group and its underlying field. Let $\mathbb{K} = \bar{\mathbb{F}}_q$ be the algebraic closure of the finite field with q elements and let $\mathbb{K}_m = \mathbb{F}_{q^m}$ denote the finite subfield with q^m elements. Let $\mathrm{GL}(n, \mathbb{K})$ denote the general linear group over $\mathbb{K}$, and define Frobenius maps

$$\begin{array}{rccc} F: & \mathrm{GL}(n,\mathbb{K}) & \longrightarrow & \mathrm{GL}(n,\mathbb{K}) \\ & (a_{ij}) & \mapsto & (a_{ji}^q)^{-1}, \end{array} \quad \text{and} \quad \begin{array}{rccc} F': & \mathrm{GL}(n,\mathbb{K}) & \longrightarrow & \mathrm{GL}(n,\mathbb{K}) \\ & (a_{ij}) & \mapsto & (a_{n-j,n-i}^q)^{-1}. \end{array} \tag{2.1}$$

Then the unitary group $U_n = \mathrm{U}(n, \mathbb{K}_2)$ is given by

$$U_n = \mathrm{GL}(n,\mathbb{K})^F = \{a \in \mathrm{GL}(n,\mathbb{K}) \mid F(a) = a\} \tag{2.2}$$

$$\cong \mathrm{GL}(n,\mathbb{K})^{F'} = \{a \in \mathrm{GL}(n,\mathbb{K}) \mid F'(a) = a\}. \tag{2.3}$$

We define the multiplicative groups $\mathbb{M}_m$ as

$$\mathbb{M}_m = \mathrm{GL}(1,\mathbb{K})^{F^m} = \{x \in \mathbb{K} \mid x^{q^m-(-1)^m} = 1\}.$$

Note that $\mathbb{M}_m \cong \mathbb{K}_m^\times$ only if m is even. We identify $\mathbb{K}^\times$ with the inverse limit $\varprojlim \mathbb{M}_m$ with respect to the norm maps

$$\begin{array}{rccc} N_{mr}: & \mathbb{M}_m & \longrightarrow & \mathbb{M}_r \\ & x & \mapsto & xx^{-q}\cdots x^{(-q)^{m/r-1}} \end{array}, \quad \text{where } m, r \in \mathbb{Z}_{\geq 1} \text{ with } r \mid m.$$

If $\mathbb{M}_m^*$ is the group of characters of $\mathbb{M}_m$, then the direct limit $\mathbb{K}^* = \varinjlim \mathbb{M}_m^*$ gives the group of characters of $\mathbb{K}^\times$. Let

$$\Theta = \{F\text{-orbits of } \mathbb{K}^*\}.$$

A polynomial $f(t) \in \mathbb{K}_2[t]$ is *F-irreducible* if there exists an F-orbit $\{x, x^{-q}, \ldots, x^{(-q)^d}\}$ of $\mathbb{K}^\times$ such that

$$f(t) = (t - x)(t - x^{-q}) \cdots (t - x^{(-q)^d}).$$

Let

$$\Phi = \{f \in \mathbb{K}_2[t] \mid f \text{ is } F\text{-irreducible}\} \stackrel{1-1}{\longleftrightarrow} \{F\text{-orbits of } \mathbb{K}^\times\}. \tag{2.4}$$

2.2. Combinatorics of Φ-partitions and Θ-partitions. Fix an ordering of Φ and Θ, and let

$$\mathcal{P} = \{\text{partitions}\} \quad \text{and} \quad \mathcal{P}_n = \{\nu \in \mathcal{P} \mid |\nu| = n\}.$$

Let $\mathcal{X}$ be either Φ or Θ. An *$\mathcal{X}$-partition* $\boldsymbol{\nu} = (\boldsymbol{\nu}(x_1), \boldsymbol{\nu}(x_2), \ldots)$ is a sequence of partitions indexed by $\mathcal{X}$. The *size* of an $\mathcal{X}$-partition $\boldsymbol{\nu}$ is

$$||\boldsymbol{\nu}|| = \sum_{x \in \mathcal{X}} |x||\boldsymbol{\nu}(x)|, \quad \text{where} \quad |x| = \begin{cases} |x| & \text{if } \mathcal{X} = \Theta, \\ d(x) & \text{if } \mathcal{X} = \Phi, \end{cases} \tag{2.5}$$

$|x|$ is the size of the orbit $x \in \Theta$, and $d(x)$ is the degree of the polynomial $x \in \Phi$. Let

$$\mathcal{P}_n^{\mathcal{X}} = \{\mathcal{X}\text{-partitions } \boldsymbol{\nu} \mid ||\boldsymbol{\nu}|| = n\}, \quad \text{and} \quad \mathcal{P}^{\mathcal{X}} = \bigcup_{n=1}^{\infty} \mathcal{P}_n^{\mathcal{X}}. \tag{2.6}$$

For $\boldsymbol{\nu} \in \mathcal{P}^{\mathcal{X}}$, let

$$n(\boldsymbol{\nu}) = \sum_{x \in \mathcal{X}} |x| n(\boldsymbol{\nu}(x)), \quad \text{where} \quad n(\nu) = \sum_{i=1}^{\ell(\nu)} (i-1)\nu_i. \tag{2.7}$$

The *conjugate* $\boldsymbol{\nu}'$ of $\boldsymbol{\nu}$ is the $\mathcal{X}$-partition $\boldsymbol{\nu}' = (\boldsymbol{\nu}(x_1)', \boldsymbol{\nu}(x_2)', \ldots)$, where ν' is the usual conjugate partition for $\nu \in \mathcal{P}$.

The *semisimple part* $\boldsymbol{\nu}_s$ of $\boldsymbol{\nu} = (\boldsymbol{\nu}(x_1), \boldsymbol{\nu}(x_2), \ldots) \in \mathcal{P}_n^{\mathcal{X}}$ is

$$\boldsymbol{\nu}_s = ((1^{|\boldsymbol{\nu}(x_1)|}), (1^{|\boldsymbol{\nu}(x_2)|}), \ldots) \in \mathcal{P}_n^{\mathcal{X}}, \tag{2.8}$$

and the *unipotent part* $\boldsymbol{\nu}_u$ of $\boldsymbol{\nu} \in \mathcal{P}_n^{\mathcal{X}}$ is given by

$$\boldsymbol{\nu}_u(\mathbb{1}) \quad \text{has parts} \quad \{|x|\boldsymbol{\nu}(x)_i \mid x \in \mathcal{X}, i = 1, \ldots, \ell(\boldsymbol{\nu}(x))\} \tag{2.9}$$

where

$$\mathbb{1} = \begin{cases} \{\mathbf{1}\} & \text{if } \mathcal{X} = \Theta, \\ t-1 & \text{if } \mathcal{X} = \Phi, \end{cases}$$

$\mathbf{1}$ is the trivial character in $\mathbb{K}^*$, and $\boldsymbol{\nu}_u(x) = \emptyset$ for $x \neq \mathbb{1}$.

2.3. Levi subgroups and maximal tori. Let $\mathcal{X}$ be either Φ or Θ as in Section 2.2. For $\boldsymbol{\nu} \in \mathcal{P}_n^{\mathcal{X}}$, let

$$L_{\boldsymbol{\nu}} = \bigoplus_{x \in \mathcal{X}_{\boldsymbol{\nu}}} L_{\boldsymbol{\nu}}(x), \quad \text{where} \quad \mathcal{X}_{\boldsymbol{\nu}} = \{x \in \mathcal{X} \mid \boldsymbol{\nu}(x) \neq \emptyset\}, \tag{2.10}$$

and for $x \in \mathcal{X}_{\boldsymbol{\nu}}$,

$$L_{\boldsymbol{\nu}}(x) = \begin{cases} \mathrm{U}(|\boldsymbol{\nu}(x)|, \mathbb{K}_{2|x|}) & \text{if } |x| \text{ is odd}, \\ \mathrm{GL}(|\boldsymbol{\nu}(x)|, \mathbb{K}_{|x|}) & \text{if } |x| \text{ is even}. \end{cases} \tag{2.11}$$

Then $L_{\boldsymbol{\nu}}$ is a Levi subgroup of $U_n = \mathrm{U}(n, \mathbb{K}_2)$ (though not uniquely determined by $\boldsymbol{\nu}$). The Weyl group

$$W_{\boldsymbol{\nu}} = \bigoplus_{x \in \mathcal{X}_{\boldsymbol{\nu}}} S_{|\boldsymbol{\nu}(x)|}, \tag{2.12}$$

of $L_{\boldsymbol{\nu}}$ has conjugacy classes indexed by

$$\mathcal{P}_s^{\boldsymbol{\nu}} = \{\boldsymbol{\gamma} \in \mathcal{P}^{\mathcal{X}} \mid \boldsymbol{\gamma}_s = \boldsymbol{\nu}_s\}, \tag{2.13}$$

and the size of the conjugacy class $c_{\boldsymbol{\gamma}}$ is

$$|c_{\boldsymbol{\gamma}}| = \frac{|W_{\boldsymbol{\gamma}}|}{z_{\boldsymbol{\gamma}}}, \quad \text{where} \quad z_{\boldsymbol{\gamma}} = \prod_{x \in \mathcal{X}} z_{\boldsymbol{\gamma}(x)} \quad \text{and} \quad z_{\gamma} = \prod_{i=1}^{\ell(\gamma)} i^{m_i} m_i!, \tag{2.14}$$

for $\gamma = (1^{m_1} 2^{m_2} \cdots) \in \mathcal{P}$.

For every $\nu = (\nu_1, \nu_2, \ldots, \nu_\ell) \in \mathcal{P}_n$ there exists a maximal torus (unique up to isomorphism) T_ν of U_n such that

$$T_\nu \cong \mathbb{M}_{\nu_1} \times \mathbb{M}_{\nu_2} \times \cdots \times \mathbb{M}_{\nu_\ell}.$$

For every $\boldsymbol{\gamma} \in \mathcal{P}_s^{\boldsymbol{\mu}}$, there exists a maximal torus (unique up to isomorphism) $T_{\boldsymbol{\gamma}} \subseteq L_{\boldsymbol{\nu}}$ such that

$$T_{\boldsymbol{\gamma}} = \bigoplus_{x \in \mathcal{X}_{\boldsymbol{\nu}}} T_{\boldsymbol{\gamma}}(x), \quad \text{where} \quad T_{\boldsymbol{\gamma}}(x) \cong \mathbb{M}_{|x|\boldsymbol{\gamma}(x)_1} \times \cdots \times \mathbb{M}_{|x|\boldsymbol{\gamma}(x)_\ell}. \tag{2.15}$$

Note that as a maximal torus of U_n, the torus $T_{\boldsymbol{\gamma}} \cong T_{\boldsymbol{\gamma}_u(\mathbb{1})}$.

2.4. Conjugacy classes and Jordan decomposition.

Proposition 2.1. *The conjugacy classes $c_{\boldsymbol{\mu}}$ of U_n are indexed by $\boldsymbol{\mu} \in \mathcal{P}_n^{\Phi}$.*

For $r \in \mathbb{Z}_{\geq 0}$, let $\psi_r(x) = \prod_{i=1}^r (1 - x^i)$.

Proposition 2.2 (Wall). *Let $g \in c_{\boldsymbol{\mu}}$. The order $a_{\boldsymbol{\mu}}$ of the centralizer g in U_n is*

$$a_{\boldsymbol{\mu}} = (-1)^{||\boldsymbol{\mu}||} \prod_{f \in \Phi} a_{\boldsymbol{\mu}(f)}\big((-q)^{d(f)}\big), \text{ where } a_\mu(x) = x^{|\mu| + 2n(\mu)} \prod_j \psi_{m_j}(x^{-1}),$$

for $\mu = (1^{m_1} 2^{m_2} 3^{m_3} \cdots) \in \mathcal{P}$.

For $\boldsymbol{\mu} \in \mathcal{P}^{\Phi}$, let $L_{\boldsymbol{\mu}}$ be as in (2.10). Note that $|L_{\boldsymbol{\mu}}| = a_{\boldsymbol{\mu}_s}$.

Lemma 2.1. *Suppose $g \in c_{\boldsymbol{\mu}}$ with Jordan decomposition $g = su$. Then*

(a) *$s \in c_{\boldsymbol{\mu}_s}$ and $u \in c_{\boldsymbol{\mu}_u}$, where $\boldsymbol{\mu}_s$ and $\boldsymbol{\mu}_u$ are as in (2.8) and (2.9),*
(b) *the centralizer $C_{U_n}(s)$ of s in U_n is isomorphic to $L_{\boldsymbol{\mu}}$.*

3. The Ennola Conjecture

3.1. The characteristic map. Let $X = \{X_1, X_2, \ldots\}$ be an infinite set of variables and let $\Lambda(X)$ be the graded $\mathbb{C}$-algebra of symmetric functions in the variables $\{X_1, X_2, \ldots\}$. Define the *power-sum symmetric function*, $p_\nu(X)$, and the *Schur function*, $s_\lambda(X)$, for $\nu, \lambda \in \mathcal{P}$, as they are in [**14**, Chapter I].

The irreducible characters ω^λ of S_n are indexed by $\lambda \in \mathcal{P}_n$, as in [**14**, Chapter I]. Let $\omega^\lambda(\nu)$ be the value of ω^λ on a permutation with cycle type ν. The relationship between $p_\nu(X)$ and $s_\lambda(X)$ is given by

$$s_\lambda(X) = \sum_{\nu \in \mathcal{P}_{|\lambda|}} \omega^\lambda(\nu) z_\nu^{-1} p_\nu(X), \quad \text{where} \quad z_\nu = \prod_{i \geq 1} i^{m_i} m_i! \tag{3.1}$$

is the order of the centralizer in S_n of the conjugacy class corresponding to $\nu = (1^{m_1} 2^{m_2} \cdots) \in \mathcal{P}$. Let $t \in \mathbb{C}$. For $\mu \in \mathcal{P}$, let the *Hall-Littlewood symmetric function* $P_\mu(X; t)$ be as it is defined in [**14**].

For $\nu, \mu \in \mathcal{P}_n$, the *classical Green function* $Q_\nu^\mu(t)$ is given by

$$p_\nu(X) = \sum_{\mu \in \mathcal{P}_{|\nu|}} Q_\nu^\mu(t^{-1}) t^{n(\mu)} P_\mu(X; t). \tag{3.2}$$

The $p_\nu(X)$, $s_\lambda(X)$, and $P_\mu(X; t)$, are all bases of $\Lambda(X)$ as a $\mathbb{C}$-algebra.

For every $f \in \Phi$, fix a set of independent variables $X^{(f)} = \{X_1^{(f)}, X_2^{(f)}, \ldots\}$, and for any symmetric function h, we let $h(f) = h(X^{(f)})$ denote the symmetric function in the variables $X^{(f)}$. Let

$$\Lambda = \mathbb{C}\text{-span}\{P_{\boldsymbol{\mu}} \mid \boldsymbol{\mu} \in \mathcal{P}^\Phi\}, \quad \text{where} \quad P_{\boldsymbol{\mu}} = (-q)^{-n(\boldsymbol{\mu})} \prod_{f \in \Phi} P_{\boldsymbol{\mu}(f)}(f; (-q)^{-d(f)}).$$

Then

$$\Lambda = \bigoplus_{n \geq 0} \Lambda_n, \quad \text{where} \quad \Lambda_n = \mathbb{C}\text{-span}\{P_{\boldsymbol{\mu}} \mid \ ||\boldsymbol{\mu}|| = n\},$$

makes Λ a graded $\mathbb{C}$-algebra. Define a Hermitian inner product on Λ by

$$\langle P_{\boldsymbol{\mu}}, P_{\boldsymbol{\nu}} \rangle = a_{\boldsymbol{\mu}}^{-1} \delta_{\boldsymbol{\mu}\boldsymbol{\nu}}.$$

For each $\varphi \in \Theta$ let $Y^{(\varphi)} = \{Y_1^{(\varphi)}, Y_2^{(\varphi)}, \ldots\}$ be an infinite variable set, and for a symmetric function h, let $h(\varphi) = h(Y^{(\varphi)})$. Relate symmetric functions in the X variables to symmetric functions in the Y variables via the transform

$$p_n(\varphi) = (-1)^{n|\varphi|-1} \sum_{x \in \mathbb{M}_{n|\varphi|}} \xi(x) p_{n|\varphi|/d(f_x)}(f_x), \tag{3.3}$$

where $\varphi \in \Theta$, $\xi \in \varphi$, and $f_x \in \Phi$ satisfies $f_x(x) = 0$.

Then

$$\Lambda = \mathbb{C}\text{-span}\{s_{\boldsymbol{\lambda}} \mid \boldsymbol{\lambda} \in \mathcal{P}^\Theta\}, \quad \text{where} \quad s_{\boldsymbol{\lambda}} = \prod_{\varphi \in \Theta} s_{\boldsymbol{\lambda}(\varphi)}(\varphi). \tag{3.4}$$

Let C_n denote the set of complex-valued class functions of the group U_n, and for $||\boldsymbol{\mu}|| = n$, let $\pi_{\boldsymbol{\mu}} : U_n \to \mathbb{C}$ be the class function which is 1 on $c_{\boldsymbol{\mu}}$ and 0 elsewhere. Then the $\pi_{\boldsymbol{\mu}}$ form a $\mathbb{C}$-basis for C_n. By Proposition 2.2, the usual inner product on class functions of finite groups, $\langle \cdot, \cdot \rangle : C_n \times C_n \to \mathbb{C}$, satisfies

$$\langle \pi_{\boldsymbol{\mu}}, \pi_{\boldsymbol{\lambda}} \rangle = a_{\boldsymbol{\mu}}^{-1} \delta_{\boldsymbol{\mu}\boldsymbol{\lambda}}.$$

For $\alpha_i \in C_{n_i}$, Ennola [**5**] defined a product $\alpha_1 \star \alpha_2 \in C_{n_1+n_2}$, which takes the following value on the conjugacy class $c_{\boldsymbol{\lambda}}$:

$$\alpha_1 \star \alpha_2(c_{\boldsymbol{\lambda}}) = \sum_{||\boldsymbol{\mu}_i|| = n_i} g_{\boldsymbol{\mu}_1\boldsymbol{\mu}_2}^{\boldsymbol{\lambda}} \alpha_1(c_{\boldsymbol{\mu}_1}) \alpha_2(c_{\boldsymbol{\mu}_2}),$$

where $g_{\boldsymbol{\mu}_1\boldsymbol{\mu}_2}^{\boldsymbol{\lambda}}$ is the product of Hall polynomials (see [**14**, Chapter II])

$$g_{\boldsymbol{\mu}_1\boldsymbol{\mu}_2}^{\boldsymbol{\lambda}} = \prod_{f \in \Phi} g_{\boldsymbol{\mu}_1(f)\boldsymbol{\mu}_2(f)}^{\boldsymbol{\lambda}(f)}((-q)^{d(f)}).$$

Extend the inner product to $C = \bigoplus_{n\geq 0} C_n$, by requiring the components C_n and C_m to be orthogonal for $n \neq m$. This gives C a graded $\mathbb{C}$-algebra structure. The *characteristic map* is

$$\begin{array}{rccll} \mathrm{ch}: & C & \longrightarrow & \Lambda & \\ & \pi_{\boldsymbol{\mu}} & \mapsto & P_{\boldsymbol{\mu}}, & \quad \text{for } \boldsymbol{\mu} \in \mathcal{P}^{\Phi}. \end{array}$$

PROPOSITION 3.1. *Let multiplication in the character ring C of U_n be given by $\star$. Then the characteristic map $\mathrm{ch} : C \to \Lambda$ is an isometric isomorphism of graded $\mathbb{C}$-algebras.*

Following the work of Green [**7**] on the general linear group, Ennola was able to obtain the following result. We may follow the proof in Macdonald [**14**, IV.4] on the general linear group case, making the appropriate changes.

PROPOSITION 3.2 (Ennola). *The set $\{s_{\boldsymbol{\lambda}} \mid \boldsymbol{\lambda} \in \mathcal{P}^{\Theta}\}$ is an orthonormal basis for Λ.*

Now let $\chi^{\boldsymbol{\lambda}} \in R$ be class functions so that $\chi^{\boldsymbol{\lambda}}(1) > 0$ and $\mathrm{ch}(\chi^{\boldsymbol{\lambda}}) = \pm s_{\boldsymbol{\lambda}}$. Ennola conjectured that $\{\chi^{\boldsymbol{\lambda}} \mid \boldsymbol{\lambda} \in \mathcal{P}_n^{\Theta}\}$ is the set of irreducible characters of U_n. He pointed out that if one could show that the product $\star$ takes virtual characters to virtual characters, then the conjecture would follow. There is no known direct proof of this fact, however. Significant progress on Ennola's conjecture was only made after the work of Deligne and Lusztig [**2**].

3.2. Deligne-Lusztig Induction. Let $T_\nu \cong \mathbb{M}_{\nu_1} \times \cdots \times \mathbb{M}_{\nu_\ell}$ be a maximal torus of U_n. If $t \in T_\nu$, then t is conjugate to

$$J_{(1^{m_1})}(f_1) \oplus \cdots \oplus J_{(1^{m_\ell})}(f_\ell), \quad \text{where } f_i \in \Phi,\ m_i d(f_i) = \nu_i.$$

Define $\boldsymbol{\gamma}_t \in \mathcal{P}^{\Phi}$ by

$$\boldsymbol{\gamma}_t(f) \qquad \text{has parts} \qquad \{m_i \mid f_i = f\}. \tag{3.5}$$

Note that $(\boldsymbol{\gamma}_t)_u(t-1) = \nu$, but in general $t \notin c_{\boldsymbol{\gamma}_t}$.

For $\boldsymbol{\mu} \in \mathcal{P}^{\Phi}$, let $L_{\boldsymbol{\mu}}$, $\boldsymbol{\gamma} \in \mathcal{P}_s^{\boldsymbol{\mu}}$ and $T_{\boldsymbol{\gamma}}$ be as in Section 2.3. Let θ be a character of T_ν. The *Deligne-Lusztig character* $R_\nu(\theta) = R_{T_\nu}^{U_n}(\theta)$ is the virtual character of U_n given by

$$\big(R_\nu(\theta)\big)(g) = \sum_{\substack{t \in T_\nu \\ \boldsymbol{\gamma}_t \in \mathcal{P}_s^{\boldsymbol{\mu}}}} \theta(t) Q_{T_{\boldsymbol{\gamma}_t}}^{L_{\boldsymbol{\mu}}}(u),$$

where $g \in c_{\boldsymbol{\mu}}$ has Jordan decomposition $g = su$ (thus, by Lemma 2.1 $C_{U_n}(s) \cong L_{\boldsymbol{\mu}}$), and $Q_{T_{\boldsymbol{\gamma}_t}}^{L_{\boldsymbol{\mu}}}(u)$ is a Green function for the unitary group (see, for example, [**1**]).

Deligne and Lusztig proved that the $R_\nu(\theta)$ span the class functions of U_n,

$$C_n = \mathbb{C}\text{-span}\{R_\nu(\theta) \mid \nu \in \mathcal{P}_n, \theta \in \mathrm{Hom}(T_\nu, \mathbb{C}^\times)\},$$

so we may define *Deligne-Lusztig induction* by

$$\begin{array}{rccl} R_{U_m \oplus U_n}^{U_{m+n}}: & C_m \otimes C_n & \longrightarrow & C_{m+n} \\ & R_\alpha^{U_m}(\theta_\alpha) \otimes R_\beta^{U_n}(\theta_\beta) & \mapsto & R_{T_\alpha \oplus T_\beta}^{U_{m+n}}(\theta_\alpha \otimes \theta_\beta), \end{array} \tag{3.6}$$

for $\alpha \in \mathcal{P}_m$, $\beta \in \mathcal{P}_n$, $\theta_\alpha \in \mathrm{Hom}(T_\alpha, \mathbb{C})$, and $\theta_\beta \in \mathrm{Hom}(T_\beta, \mathbb{C})$.

Let Λ and C be as in Section 3.1, except we now give C a graded $\mathbb{C}$-algebra structure using Deligne-Lusztig induction. That is, we define a multiplication $\circ$ on C by

$$\chi \circ \eta = R_{U_m \oplus U_n}^{U_{m+n}}(\chi \otimes \eta), \quad \text{for } \chi \in C_m \text{ and } \eta \in C_n.$$

We recall the characteristic map defined in Section 3.1,

$$\begin{array}{rccll} \mathrm{ch}: & C & \longrightarrow & \Lambda & \\ & \pi_{\boldsymbol{\mu}} & \mapsto & P_{\boldsymbol{\mu}} & \quad \text{for } \boldsymbol{\mu} \in \mathcal{P}^{\Phi}. \end{array}$$

It is immediate that ch is an isometric isomorphism of vector spaces, but it is not yet clear if ch is also a ring homomorphism when C has multiplication given by Deligne-Lusztig induction.

3.3. The Ennola conjecture. In this section we summarize the remaining steps that are necessary to obtain the proof of the Ennola conjecture. First, we must compute $\mathrm{ch}(R_\nu(\theta))$, and to do so we need to write the Green functions $Q_{T_\gamma}^{L_\mu}(u)$ as polynomials in q. These Green functions turn out to be those of the general linear group, except with q replaced by $-q$, which is the essence of Ennola's original idea. This fact was proven by Hotta and Springer [**9**] for the case that $p = \mathrm{char}(\mathbb{F}_q)$ is large compared to n, and was finally proven in full generality by Kawanaka [**11**].

THEOREM 3.1 (Hotta-Springer, Kawanaka). *The Green functions for the unitary group are given by* $Q_{T_\gamma}^{L_\mu}(u) = Q_{\boldsymbol{\gamma}}^{\boldsymbol{\mu}}(-q)$, *where*

$$Q_{\boldsymbol{\gamma}}^{\boldsymbol{\mu}}(-q) = \prod_{f\in\Phi_{\boldsymbol{\mu}}} Q_{\boldsymbol{\gamma}(f)}^{\boldsymbol{\mu}(f)}((-q)^{d(f)}),$$

and $Q_\gamma^\mu(q)$ *is the classical Green function as in (3.2).*

For $\nu = (\nu_1, \nu_2, \ldots, \nu_\ell) \in \mathcal{P}$ and $\theta = \theta_1 \otimes \theta_2 \otimes \cdots \otimes \theta_\ell$ a character of T_ν, define

$$p_{\boldsymbol{\mu}_{\nu\theta}} = \prod_{\varphi\in\Theta} p_{\boldsymbol{\mu}_{\nu\theta}(\varphi)}(\varphi), \quad \text{where} \quad \boldsymbol{\mu}_{\nu\theta}(\varphi) = (\nu_i/|\varphi| \ \mid \ \theta_i \in \varphi).$$

From Theorem 3.1 and the machinery of the characteristic map, we obtain the following.

THEOREM 3.2. *Let* $\nu = (\nu_1, \nu_2, \ldots, \nu_\ell) \in \mathcal{P}$, $\theta = \theta_1\otimes\theta_2\otimes\cdots\otimes\theta_\ell$ *be a character of* T_ν, *and* $\boldsymbol{\nu} = \boldsymbol{\mu}_{\nu\theta} \in \mathcal{P}^\Theta$. *Then*

$$\mathrm{ch}\big(R_\nu(\theta)\big) = (-1)^{||\boldsymbol{\nu}||-\ell(\boldsymbol{\nu})} p_{\boldsymbol{\nu}}.$$

COROLLARY 3.1. *Let multiplication in the character ring C of U_n be given by $\circ$. Then the characteristic map* $\mathrm{ch} : C \to \Lambda$ *is an isometric isomorphism of graded $\mathbb{C}$-algebras.*

An immediate consequence is that the graded multiplication that Ennola originally defined on C is exactly Deligne-Lusztig induction, or

COROLLARY 3.2. *Let $\chi \in C_m$ and $\eta \in C_n$. Then*

$$\chi \circ \eta = \chi \star \eta.$$

We therefore have the advantage of taking either definition when convenience demands.

For $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$, let $L_{\boldsymbol{\lambda}}$, $W_{\boldsymbol{\lambda}}$, and $T_{\boldsymbol{\gamma}}$, $\boldsymbol{\gamma} \in \mathcal{P}_s^{\boldsymbol{\lambda}}$, be as in Section 2.3.

Note that the combinatorics of $\boldsymbol{\gamma}$ almost specifies character θ_γ of $T_{\boldsymbol{\gamma}}$ in the sense that

$$\theta_{\boldsymbol{\gamma}}(T_{\boldsymbol{\gamma}}(\varphi)) = \theta_\varphi(T_{\boldsymbol{\gamma}}(\varphi)), \qquad \text{for some } \theta_\varphi \in \varphi.$$

In fact, we may define

$$R_{\boldsymbol{\gamma}} = R_{T_{\boldsymbol{\gamma}}}^{U_n}(\theta_{\boldsymbol{\gamma}}) = \mathrm{ch}^{-1}\big((-1)^{||\boldsymbol{\gamma}||-\ell(\boldsymbol{\gamma})} p_{\boldsymbol{\gamma}}\big), \tag{3.7}$$

where $\theta_{\boldsymbol{\gamma}}$ is any choice of the θ_φ's.

For every $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$ there exists a character $\omega^{\boldsymbol{\lambda}}$ of $W_{\boldsymbol{\lambda}}$ defined by

$$\omega^{\boldsymbol{\lambda}}(\boldsymbol{\gamma}) = \prod_{\varphi\in\Theta} \omega^{\boldsymbol{\lambda}(\varphi)}(\boldsymbol{\gamma}(\varphi)),$$

where $\omega^{\boldsymbol{\lambda}}(\boldsymbol{\gamma})$ is the value of $\omega^{\boldsymbol{\lambda}}$ on the conjugacy class $c_{\boldsymbol{\gamma}}$ corresponding to $\boldsymbol{\gamma} \in \mathcal{P}_s^\Theta$.

In [**13**], Lusztig and Srinivasan decomposed the irreducible characters of U_n as linear combinations of Deligne-Lusztig characters, as follows.

THEOREM 3.3 (Lusztig-Srinivasan). *Let* $\boldsymbol{\lambda} \in \mathcal{P}_n^\Theta$. *Then there exists* $\tau'(\boldsymbol{\lambda}) \in \mathbb{Z}_{\geq 0}$ *such that the class function*

$$R(\boldsymbol{\lambda}) = (-1)^{\tau'(\boldsymbol{\lambda})+\lfloor n/2\rfloor+\sum_{\varphi\in\Theta}|\boldsymbol{\lambda}(\varphi)|+|\varphi|\lceil|\boldsymbol{\lambda}(\varphi)|/2\rceil} \sum_{\boldsymbol{\gamma}\in\mathcal{P}_s^{\boldsymbol{\lambda}}} \frac{\omega^{\boldsymbol{\lambda}}(\boldsymbol{\gamma})}{z_{\boldsymbol{\gamma}}} R_{\boldsymbol{\gamma}}$$

is an irreducible character of U_n ($z_{\boldsymbol{\gamma}}$ is as in (2.14)).

Finally, we obtain the Ennola Conjecture.

COROLLARY 3.3 (Ennola Conjecture). *For $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$, there exists $\tau(\boldsymbol{\lambda}) \in \mathbb{Z}_{\geq 0}$ such that*

$$\left\{ \mathrm{ch}^{-1}\Big((-1)^{\tau(\boldsymbol{\lambda})} s_{\boldsymbol{\lambda}} \Big) \ \mid\ \boldsymbol{\lambda} \in \mathcal{P}_n^\Theta \right\}$$

is the set of irreducible characters of U_n.

PROOF. By Theorem 3.3 and Theorem 3.2,

$$\begin{aligned}
\mathrm{ch}(R(\boldsymbol{\lambda})) &= (-1)^{\tau'(\boldsymbol{\lambda})+\lfloor n/2\rfloor+\sum_{\varphi\in\Theta}|\boldsymbol{\lambda}(\varphi)|+|\varphi|\lceil|\boldsymbol{\lambda}(\varphi)|/2\rceil} \sum_{\boldsymbol{\gamma}\in\mathcal{P}_s^{\boldsymbol{\lambda}}} \frac{\omega^{\boldsymbol{\lambda}}(\boldsymbol{\gamma})}{z_{\boldsymbol{\gamma}}} (-1)^{n-\ell(\boldsymbol{\gamma})} p_{\boldsymbol{\gamma}} \\
&= (-1)^{\tau'(\boldsymbol{\lambda})+\lfloor n/2\rfloor+n+\sum_{\varphi\in\Theta}|\varphi|\lceil|\boldsymbol{\lambda}(\varphi)|/2\rceil} \sum_{\boldsymbol{\gamma}\in\mathcal{P}_s^{\boldsymbol{\lambda}}} \frac{\omega^{\boldsymbol{\lambda}}(\boldsymbol{\gamma})}{z_{\boldsymbol{\gamma}}} (-1)^{\sum_{\varphi\in\Theta}|\boldsymbol{\lambda}(\varphi)|-\ell(\boldsymbol{\gamma})} p_{\boldsymbol{\gamma}}
\end{aligned}$$

Note that the sign character $\omega^{\boldsymbol{\lambda}_s}$ of $W_{\boldsymbol{\lambda}}$ acts by

$$\omega^{\boldsymbol{\lambda}_s}(\boldsymbol{\gamma}) = (-1)^{\sum_{\varphi\in\Theta}|\boldsymbol{\gamma}(\varphi)|-\ell(\boldsymbol{\gamma})},$$

and that $\omega^{\boldsymbol{\lambda}} \otimes \omega^{\boldsymbol{\lambda}_s} = \omega^{\boldsymbol{\lambda}'}$, so since $\boldsymbol{\gamma} \in \mathcal{P}_s^{\boldsymbol{\lambda}}$,

$$\begin{aligned}
\mathrm{ch}(R(\boldsymbol{\lambda})) &= (-1)^{\tau'(\boldsymbol{\lambda})+\lfloor n/2\rfloor+n+\sum_{\varphi\in\Theta}|\varphi|\lceil|\boldsymbol{\lambda}(\varphi)|/2\rceil} \sum_{\boldsymbol{\gamma}\in\mathcal{P}_s^{\boldsymbol{\lambda}}} \frac{(\omega^{\boldsymbol{\lambda}}\otimes\omega^{\boldsymbol{\lambda}_s})(\boldsymbol{\gamma})}{z_{\boldsymbol{\gamma}}} p_{\boldsymbol{\gamma}} \\
&= (-1)^{\tau'(\boldsymbol{\lambda})+\lfloor n/2\rfloor+n+\sum_{\varphi\in\Theta}|\varphi|\lceil|\boldsymbol{\lambda}(\varphi)|/2\rceil} \sum_{\boldsymbol{\gamma}\in\mathcal{P}_s^{\boldsymbol{\lambda}}} \frac{\omega^{\boldsymbol{\lambda}'}(\boldsymbol{\gamma})}{z_{\boldsymbol{\gamma}}} p_{\boldsymbol{\gamma}},
\end{aligned}$$

and by applying (3.1) to a product over Θ,

$$= (-1)^{\tau'(\boldsymbol{\lambda})+\lfloor n/2\rfloor+n+\sum_{\varphi\in\Theta}|\varphi|\lceil|\boldsymbol{\lambda}(\varphi)|/2\rceil} s_{\boldsymbol{\lambda}'}. \qquad \square$$

Remark. There are at least two natural ways to index the irreducible characters of U_n by Θ-partitions: Theorem 3.3 gives a natural indexing by Θ-partitions, but Corollary 3.3 indicates that the conjugate choice is equally natural. Since we like to think of Schur functions as irreducible characters, we have chosen the latter indexing. However, several references, including Ennola [**5**], Ohmori [**15**], and Henderson [**8**], make use of the former.

4. Characters degrees

In this section, we calculate the degrees of the irreducible characters of the finite unitary group and find several character degree sums.

Let $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$, and suppose $\square \in \boldsymbol{\lambda}$ is in position (i,j) in $\boldsymbol{\lambda}(\varphi)$ for some $\varphi \in \Theta$. The *hook length* $\mathbf{h}(\square)$ of $\square$ is

$$\mathbf{h}(\square) = |\varphi| h(\square), \quad \text{where} \quad h(\square) = \boldsymbol{\lambda}(\varphi)_i - \boldsymbol{\lambda}(\varphi)'_j - i - j + 1,$$

is the usual hook length for partitions.

For $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$, let

$$\eta^{\boldsymbol{\lambda}} = \mathrm{ch}^{-1}(s_{\boldsymbol{\lambda}}).$$

Adapting the computations in [**14**, IV.6], we obtain the following result.

THEOREM 4.1. *Let $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$ and let 1 be the identity in $U_{||\boldsymbol{\lambda}||}$. Then*

$$\eta^{\boldsymbol{\lambda}}(1) = (-1)^{\tau(\boldsymbol{\lambda})} q^{n(\boldsymbol{\lambda}')} \frac{\prod_{1\leq i\leq ||\boldsymbol{\lambda}||} (q^i - (-1)^i)}{\prod_{\square\in\boldsymbol{\lambda}} (q^{\mathbf{h}(\square)} - (-1)^{\mathbf{h}(\square)})},$$

where $\tau(\boldsymbol{\lambda}) = ||\boldsymbol{\lambda}||(||\boldsymbol{\lambda}||+3)/2 + n(\boldsymbol{\lambda}) \equiv \lfloor ||\boldsymbol{\lambda}||/2 \rfloor + n(\boldsymbol{\lambda}) \pmod 2$. So for each $\boldsymbol{\lambda}$, we have $\chi^{\boldsymbol{\lambda}} = (-1)^{\tau(\boldsymbol{\lambda})}\eta^{\boldsymbol{\lambda}}$.

The following result follows from Theorem 4.1 and the Littlewood-Richardson rule.

Corollary 4.1. *Let $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{P}^\Theta$. Then $\chi^{\boldsymbol{\mu}} \circ \chi^{\boldsymbol{\nu}}$ is a character if and only if every $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$ such that $c^{\boldsymbol{\lambda}}_{\boldsymbol{\mu}\boldsymbol{\nu}} > 0$ satisfies*

$$n(\boldsymbol{\mu}) + n(\boldsymbol{\nu}) \equiv n(\boldsymbol{\lambda}) + ||\boldsymbol{\mu}||\, ||\boldsymbol{\nu}|| \pmod 2$$

Following a similar approach to [**14**, IV.6, Example 5], we consider the coefficient of t^m in the series

$$S = \sum_{\boldsymbol{\lambda} \in \mathcal{P}^\Theta} (-1)^{n(\boldsymbol{\lambda}) + ||\boldsymbol{\lambda}||} \delta(s_{\boldsymbol{\lambda}}) t^{||\boldsymbol{\lambda}||},$$

where $n(\boldsymbol{\lambda})$ is as in (2.7), and we obtain the following result.

Theorem 4.2. *The sum of the degrees of the complex irreducible characters of U_m is given by*

$$\sum_{||\boldsymbol{\lambda}||=m} \chi^{\boldsymbol{\lambda}}(1) = (q+1)q^2(q^3+1)q^4(q^5+1)\cdots\big(q^m + \frac{(1-(-1)^m)}{2}\big).$$

Write $f_{U_m}(q) = \sum_{||\boldsymbol{\lambda}||=m} \chi^{\boldsymbol{\lambda}}(1)$. The polynomial $f_{G_m}(q)$ expressing the sum of the degrees of the complex irreducible characters of $G_m = \mathrm{GL}(m, \mathbb{F}_q)$, was computed in [**6**] for odd q and in [**12**] and Example 6 of [**14**, IV.6] for general q. From these results we see that

$$f_{U_m}(q) = (-1)^{m(m+1)/2} f_{G_m}(-q),$$

another example of Ennola duality.

Gow [**6**] and Klyachko [**12**] proved that the sum of the degrees of the complex irreducible characters of G_n is equal to the number of symmetric matrices in G_n. We obtain the same result for U_n by applying Theorem 4.2 and a counting argument.

Corollary 4.2. *The sum of the degrees of the complex irreducible characters of $\mathrm{U}(n, \mathbb{F}_{q^2})$ is equal to the number of symmetric matrices in $\mathrm{U}(n, \mathbb{F}_{q^2})$.*

A Θ-partition $\boldsymbol{\lambda}$ is *even* if every part of $\boldsymbol{\lambda}(\varphi)$ is even for every $\varphi \in \Theta$.

Theorem 4.3. *The sum of the degrees of the complex irreducible characters of U_{2m} corresponding to $\boldsymbol{\lambda}$ such that $\boldsymbol{\lambda}'$ is even is given by*

$$\sum_{\substack{||\boldsymbol{\lambda}||=2m \\ \boldsymbol{\lambda}' \text{ even}}} \chi^{\boldsymbol{\lambda}}(1) = (q+1)q^2(q^3+1)\cdots q^{2m-2}(q^{2m-1}+1) = \frac{|\mathrm{U}(2m, \mathbb{F}_{q^2})|}{|\mathrm{Sp}(2m, \mathbb{F}_q)|}.$$

Write $g_{U_m}(q) = \sum_{||\boldsymbol{\lambda}||=2m, \boldsymbol{\lambda}' \text{ even}} \chi^{\boldsymbol{\lambda}}(1)$, and let $g_{G_m}(q)$ denote the corresponding sum for G_m. The polynomial $g_{G_m}(q)$ was calculated in Example 7 of [**14**, IV.6], and similar to the previous example, we see that we have

$$g_{U_m}(q) = (-1)^m g_{G_m}(-q).$$

In the case that q is odd, Proposition 4.3 follows from the following stronger result obtained by Henderson [**8**]. Let $Sp_{2n} = \mathrm{Sp}(2n, \mathbb{F}_q)$ be the symplectic group over the finite field $\mathbb{F}_q$.

Theorem 4.4 (Henderson). *Let q be odd. The decomposition of $\mathrm{Ind}_{Sp_{2n}}^{U_{2n}}(\mathbf{1})$ into irreducibles is given by*

$$\mathrm{Ind}_{Sp_{2n}}^{U_{2n}}(\mathbf{1}) = \sum_{\substack{||\boldsymbol{\lambda}||=2n \\ \boldsymbol{\lambda}' \text{even}}} \chi^{\boldsymbol{\lambda}}.$$

The fact that Proposition 4.3 holds for all q suggests that Theorem 4.4 should as well.

5. A Deligne-Lusztig model

A *model* of a finite group G is a representation ρ, which is a direct sum of representations induced from one-dimensional representations of subgroups of G, such that every irreducible representation of G appears as a component with multiplicity 1 in the decomposition of ρ.

Klyachko [**12**] and Inglis and Saxl [**10**] obtained a model for $\mathrm{GL}(n, \mathbb{F}_q)$, where the induced representations can be written as a Harish-Chandra product of Gelfand-Graev characters and the permutation character of the finite symplectic group.

In this section we show that the same result is true for the finite unitary group, except the Harish-Chandra product is replaced by Deligne-Lusztig induction. The result is therefore not a model for $\mathrm{U}(n, \mathbb{F}_q)$

in the finite group character induction sense, but rather from the Deligne-Lusztig point of view.

Let $U'_n = \mathrm{GL}(n, \mathbb{K})^{F'}$ as in (2.3), and let

$$B_< = \{u \in U'_n \mid u \text{ unipotent and uppertriangular}\} \subseteq U'_n.$$

Fix a nontrivial character $\psi : \mathbb{K}_2^+ \to \mathbb{C}^\times$ of the additive group of the field $\mathbb{K}_2$ such that the restriction to the subgroup $\{x \in \mathbb{K}_2 \mid x^q + x = 0\}$ is also nontrivial. The map $\psi_{(n)} : B_< \to \mathbb{C}$ given by

$$\psi_{(n)}(u) = \psi\big(u_{12} + \cdots + u_{\lfloor n/2\rfloor -1, \lfloor n/2\rfloor} + u_{\lfloor n/2\rfloor, \lceil n/2\rceil+1}\big), \text{ for } u = (u_{ij}) \in B_<,$$

is a linear character of $B_<$. Then

$$\Gamma'_{(n)} = \mathrm{Ind}_{B_<}^{U'_n}(\psi_{(n)})$$

is the *Gelfand-Graev character* of U'_n. Let $\Gamma_{(n)}$ be the corresponding Gelfand-Graev character of $U_n = \mathrm{GL}(n, \mathbb{K})^F$. For $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$, define

$$\mathrm{ht}(\boldsymbol{\lambda}) = \max\{\ell(\boldsymbol{\lambda}(\varphi)) \mid \varphi \in \Theta\}.$$

The following appears in Section 5.2 of [**15**], but we can also give a proof using the characteristic map.

Theorem 5.1. *The decomposition of* $\Gamma_{(m)}$ *into irreducibles is given by*

$$\Gamma_{(m)} = \sum_{\substack{\boldsymbol{\lambda} \in \mathcal{P}_m^\Theta \\ \mathrm{ht}(\boldsymbol{\lambda}) = 1}} \chi^{\boldsymbol{\lambda}}.$$

For a partition λ, let $o(\lambda)$ denote the number of odd parts of λ, and for $\boldsymbol{\lambda} \in \mathcal{P}^\Theta$, let $o(\boldsymbol{\lambda}) = \sum_{\varphi \in \Theta} |\varphi| o(\boldsymbol{\lambda}(\varphi))$.

Theorem 5.2. *Let q be odd. For each r such that $0 \le 2r \le m$,*

$$\Gamma_{m-2r} \circ \mathrm{Ind}_{Sp_{2r}}^{U_{2r}}(\mathbf{1}) = \sum_{o(\boldsymbol{\lambda}') = m-2r} \chi^{\boldsymbol{\lambda}}.$$

Furthermore,

$$\sum_{0 \le 2r \le m} \Gamma_{m-2r} \circ \mathrm{Ind}_{Sp_{2r}}^{U_{2r}}(\mathbf{1}) = \sum_{||\boldsymbol{\lambda}|| = m} \chi^{\boldsymbol{\lambda}}$$

Proof. Suppose $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{P}^\Theta$, such that $\mathrm{ht}(\boldsymbol{\mu}) = 1$ and $\boldsymbol{\nu}'$ is even. From the characteristic map, Corollary 3.3, and Pieri's formula [**14**, I.5.16],

$$\chi^{\boldsymbol{\mu}} \circ \chi^{\boldsymbol{\nu}} = (-1)^{\tau(\boldsymbol{\mu}) + \tau(\boldsymbol{\nu})} \sum_{\boldsymbol{\lambda}} \chi^{\boldsymbol{\lambda}}, \tag{5.1}$$

where the sum is taken over all $\boldsymbol{\lambda}$ such that for every $\varphi \in \Theta$, $\boldsymbol{\lambda}(\varphi) - \boldsymbol{\nu}(\varphi)$ is a horizontal $|\boldsymbol{\mu}(\varphi)|$-strip.

We now use Corollary 4.1 to show that $\chi^{\boldsymbol{\mu}} \circ \chi^{\boldsymbol{\nu}}$ is a character. As $\boldsymbol{\lambda}(\varphi) - \boldsymbol{\nu}(\varphi)$ is a horizontal $|\boldsymbol{\mu}(\varphi)|$-strip, the part $\boldsymbol{\lambda}(\varphi)'_i$ is either $\boldsymbol{\nu}(\varphi)'_i$ or $\boldsymbol{\nu}(\varphi)'_i + 1$ for every $i = 1, 2, \ldots, \ell(\boldsymbol{\lambda}(\varphi))$. By assumption, $\boldsymbol{\nu}'$ is even, so $\boldsymbol{\nu}(\varphi)'_i$ is even for every $\varphi \in \Theta$, and so

$$\binom{\boldsymbol{\nu}(\varphi)'_i + 1}{2} = \boldsymbol{\nu}(\varphi)'_i + \binom{\boldsymbol{\nu}(\varphi)'_i}{2} \equiv \binom{\boldsymbol{\nu}(\varphi)'_i}{2} \qquad (\text{mod } 2).$$

Thus, $n(\boldsymbol{\lambda}(\varphi)) = \sum_i \binom{\boldsymbol{\lambda}(\varphi)'_i}{2} \equiv n(\boldsymbol{\nu}(\varphi))$ (mod 2). The assumption $\mathrm{ht}(\boldsymbol{\mu}) = 1$ implies $n(\boldsymbol{\mu}(\varphi)) = 0$, and since $||\boldsymbol{\nu}||$ is even,

$$n(\boldsymbol{\mu}) + n(\boldsymbol{\nu}) \equiv n(\boldsymbol{\lambda}) + ||\boldsymbol{\mu}||\,||\boldsymbol{\nu}|| \quad (\text{mod } 2).$$

By Corollary 4.1, $\chi^{\boldsymbol{\mu}} \circ \chi^{\boldsymbol{\nu}}$ is a character.

Use the decompositions of Theorem 4.4 and Theorem 5.1 in the product (5.1) to observe that the irreducible characters $\chi^{\boldsymbol{\lambda}}$ in the decomposition of $\Gamma_{m-2r} \circ \mathrm{Ind}_{Sp_{2r}}^{U_{2r}}(\mathbf{1})$ are indexed by $\boldsymbol{\lambda} \in \mathcal{P}_m^\Theta$ such that for every φ, $\boldsymbol{\lambda}(\varphi) - \boldsymbol{\nu}(\varphi)$ is a horizontal $|\boldsymbol{\mu}(\varphi)|$-strip, where $||\boldsymbol{\mu}|| = m - 2r$, for some $\boldsymbol{\nu}(\varphi)$ such that $\boldsymbol{\nu}(\varphi)'$ is even. Then the number of odd parts of $\boldsymbol{\lambda}(\varphi)'$ is exactly $|\boldsymbol{\mu}(\varphi)|$, and so the $\boldsymbol{\lambda}$ in the decomposition must satisfy $\sum_{\varphi \in \Theta} |\varphi| o(\boldsymbol{\lambda}(\varphi)') = ||\boldsymbol{\mu}|| = m - 2r$. $\square$

References

1. R. Carter, Finite groups of Lie type: conjugacy classes and complex characters. John Wiley and Sons, 1985.
2. P. Deligne and G. Lusztig, Representations of reductive groups over finite fields, *Ann. of Math.* (2) **103** (1976), no. 1, 103–161.
3. F. Digne and J. Michel, Foncteurs de Lusztig et caractères des groupes linéaires et unitaires sur un corps fini, *J. Algebra* **107** (1987), no. 1, 217–255.
4. V. Ennola, On the conjugacy classes of the finite unitary groups, *Ann. Acad. Sci. Fenn. Ser. A I No.* **313** (1962), 13 pages.
5. V. Ennola, On the characters of the finite unitary groups, *Ann. Acad. Sci. Fenn. Ser. A I No.* **323** (1963), 35 pages.
6. R. Gow, Properties of the characters of the finite general linear group related to the transpose-inverse involution, *Proc. London Math. Soc. (3)* **47** (1983), no. 3, 493–506.
7. J.A. Green, The characters of the finite general linear groups, *Trans. Amer. Math. Soc.* **80** (1955), no. 2, 402–447.
8. A. Henderson, Symmetric subgroup invariants in irreducible representations of G^F, when $G = GL_n$, *J. Algebra* **261** (2003), no. 1, 102–144.
9. R. Hotta and T.A. Springer, A specialization theorem for certain Weyl group representations and an application to the Green polynomials of unitary groups, *Invent. Math.* **41** (1977), no. 2, 113–127.
10. N.F.J. Inglis and J. Saxl, An explicit model for the complex representations of the finite general linear groups, *Arch. Math. (Basel)* **57** (1991), no. 5, 424–431.
11. N. Kawanaka, Generalized Gel'fand-Graev representations and Ennola duality, In *Algebraic groups and related topics (Kyoto/Nagoya, 1983)*, 175–206, *Adv. Stud. Pure Math.*, 6, North-Holland, Amsterdam, 1985.
12. A. A. Klyachko, Models for complex representations of the groups $\mathrm{GL}(n, q)$, *Mat. Sb. (N.S.)* **120(162)** (1983), no. 3, 371–386.
13. G. Lusztig and B. Srinivasan, The characters of the finite unitary groups, *J. Algebra* **49** (1977), no. 1, 167–171.
14. I.G. Macdonald, Symmetric functions and Hall polynomials. Second edition. With Contributions by A. Zelevinsky. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1995.
15. Z. Ohmori, On a Zelevinsky theorem and the Schur indices of the finite unitary groups, *J. Math. Sci. Univ. Tokyo* **2** (1997), no. 2, 417–433.
16. G. E. Wall, On the conjugacy classes in the unitary, orthogonal and symplectic groups, *J. Austral. Math. Soc.* **3** (1962), 1–62.

Department of Mathematics, Stanford University, Stanford, CA 94305-2125
E-mail address: thiem@math.stanford.edu

Department of Mathematics, The University of Arizona, 617 N Santa Rita Ave, Tucson, AZ 85721-0089
E-mail address: vinroot@math.arizona.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A Bijection for Unicellular Partitioned Bicolored Maps

E. Vassilieva and G. Schaeffer

Abstract. In the present paper we construct a bijection that relates a set $C_{N,p,q}$ of unicellular partitioned bicolored maps to a set of couples (t,σ) of ordered bicolored trees and partial permutations. This bijection allows us to derive an elegant formula for the enumeration of unicellular bicolored maps, an analogue of the well-known Harer-Zagier result for unicolored one-face maps.

Résumé. Dans cet article nous construisons une bijection mettant en relation l'ensemble $C_{N,p,q}$ des cartes bicolores unicellulaires partitionnées et l'ensemble des couples (t,σ) d'arbres bicolores ordonnés et de permutations partielles. Cette bijection nous permet de dériver une formule élégante pour l'énumération des cartes bicolores unicellulaires, analogue au résultat de Harer et Zagier pour les cartes unicolores monofaces.

1. Introduction

Maps are graphs embedded in orientable surfaces. More precisely, a *map* is a $2-$cell decomposition of a compact, connected, orientable surface into vertices ($0-$cells), edges ($1-$cells) and faces ($2-cells$) homeomorphic to open discs. Loops and multiple edges are allowed. A detailed description of these objects as well as examples of their numerous applications in various branches of mathematics and physics can be found in the survey [**2**] and in [**8**]. One face (unicellular) maps represent an object of special interest. In particular, Harer and Zagier enumerated unicellular maps of genus g with prescribed number of edges in order to calculate the Euler characteristics of the moduli spaces (see [**6**]). Numerous proofs of this well-known formula have been proposed. As a rule they are technically complicated and up to recently no elementary proof was known. A first purely combinatorial method was given by Lass in [**7**]. Another one involving a direct bijection was developped by Goulden and Nica in [**4**].

This paper is focused on unicellular bicolored maps, i.e. one-face maps with white and black vertices verifying the property that each edge is joining a black and a white vertices. Formally, a unicellular bicolored map of N edges, m white and n black vertices is equivalent to a couple of permutations $(\alpha,\beta)\in\Sigma_N$ such that α has m independent cycles, β has n independent cycles and $\alpha\beta=\gamma$ where γ is the long cycle $(1\,2\,3\,\ldots\,N)$. As a first approach to this question, in [**5**], Goupil and Schaeffer derived a formula to count the number of factorizations (α,β) of γ with α of cycle type λ and β of cycle type μ, for any pair (λ,μ) of partitions of N. Summing over all λ with m parts and μ with n parts allows to recover a complicated formula for this counting problem. Independently, a more elegant formula for the enumeration of these objects has been calculated by Adrianov in [**1**]. His method involves characters on the symmetric group and the resulting formula, expressed in terms of Gauss hypergeometric function, leaves little room for simple combinatorial interpretation. In this paper, we derive a new formula solving the same enumeration problem. To this end we construct a bijection having some aspects similar to the one of Goulden and Nica in [**4**] for unicolored maps.

Throughout this paper, we adopt the following notations. We denote by $BT(p,q)$ the set of ordered bicolored (black and white)trees with p white and q black vertices. We assume that all the trees in this set

2000 *Mathematics Subject Classification.* Primary 05A15; Secondary 05C10.
Key words and phrases. unicellular bicolored maps, partial permutations, bicolored trees, Harer-Zagier formula.

have a white root. The cardinality of $BT(p,q)$ (see e.g. [**3**]) is given by:

$$|BT(p,q)| = \frac{p+q-1}{pq}\binom{p+q-2}{p-1}^2 \tag{1.1}$$

We also denote by $PP(X,Y,A)$ the set of partial permutations from any subset of X of cardinality A to any subset of Y (of the same cardinality). The cardinality of this set is given by:

$$|PP(X,Y,A)| = \binom{|X|}{A}\binom{|Y|}{A}A! \tag{1.2}$$

For the sake of simplicity, in all what follows, if X or Y is equal to $[M]$ (all the integers between 1 and M), we will note M instead of $[M]$. Now let us turn to our main result:

THEOREM 1.1. *The numbers $B(m,n,N)$ of unicellular bicolored maps with m white vertices, n black vertices and N edges verify:*

$$\sum_{m,n\geq 1} B(m,n,N)y^m z^n = N! \sum_{p,q\geq 1}\binom{N-1}{p-1,q-1}\binom{y}{p}\binom{z}{q} \tag{1.3}$$

In the following section we give a bijective proof for this formula. To this extent we introduce a new class of objects, the *unicellular partitioned bicolored maps*.

2. Unicellular Partitioned Bicolored Maps

2.1. Definition. Let $C_{N,p,q}$ be the set of triples (π_1,π_2,α) such that π_1 et π_2 are partitions of $[N]$ into p and q blocks and such that α is a permutation of $[N]$ verifying the following properties :

- Each block of π_1 is the union of cycles of α.
- Each block of π_2 is the union of cycles of $\beta = \alpha^{-1}\gamma$, where $\gamma = (1\,2\,\ldots\,N)$.

2.2. Geometrical Interpretation. The set $C_{N,p,q}$ can be viewed as a set of unicellular partitioned bicolored maps. A triple (π_1,π_2,α) corresponds to a unicellular bicolored map with N edges where :

- The cycles of α describe the white vertices of the map.
- The cycles of $\beta = \alpha^{-1}\gamma$ describe the black vertices.
- π_1 partitions the white vertices into p subsets
- π_2 partitions the black vertices into q subsets

EXAMPLE 2.1. Figure (1) gives a representation of the triple $(\pi_1,\pi_2,\alpha) \in C_{9,3,2}$, defined by $\alpha = (1)(24)(3)(57)(6)(89)$, $\beta = (1479)(23)(56)(8)$, $\pi_1 = \{\pi_1^{(1)},\pi_1^{(2)},\pi_1^{(3)}\}$, $\pi_2 = \{\pi_2^{(1)},\pi_2^{(2)}\}$ with :

$$\pi_1^{(1)} = \{2,4,6\}, \quad \pi_1^{(2)} = \{8,9\}, \quad \pi_1^{(3)} = \{1,3,5,7\},$$
$$\pi_2^{(1)} = \{2,3,5,6\}, \quad \pi_2^{(2)} = \{1,4,7,8,9\}$$

where the numbering of the blocks is purely arbitrary.
To visualise it better we also assume that each block is associated with some particular shape: $\pi_1^{(1)}$ with square, $\pi_1^{(2)}$ with circle, $\pi_1^{(3)}$ with triangle, $\pi_2^{(1)}$ with rhombus and $\pi_2^{(2)}$ with pentagon. Therefore each vertex of our partitioned map will have a shape corresponding to its block.

2.3. Connection with Unicellular Bicolored Maps. Let $c_{N,p,q} =\mid C_{N,p,q}\mid$. Using the Stirling number of the second kind $S(a,b)$ enumerating the partitions of a set of a elements into b non-empty, unordered subsets, we have:

$$c_{N,p,q} = \sum_{m\geq p, n\geq q} S(m,p)S(n,q)B(m,n,N) \tag{2.1}$$

Then, since $\sum_{b=1}^{a} S(a,b)(x)_b = x^a$ (see e.g [**9**]) where the falling factorial $(x)_b = \prod_{i=0}^{b-1}(x-i)$:

$$\sum_{m,n\geq 1} B(m,n,N)y^m z^n = \sum_{p,q\geq 1} c_{N,p,q}(y)_p(z)_q \tag{2.2}$$

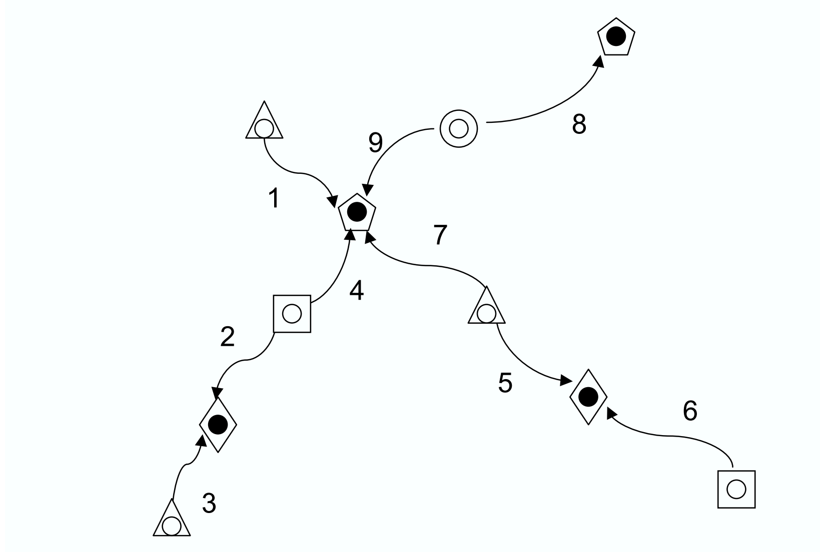


FIGURE 1. Example of a Partitioned Bicolored Map

3. Bijective Description of Unicellular Partitioned Bicolored Maps

3.1. Combinatorial Interpretation of the Main Formula. Combining equations 1.3 and 2.2 gives:

$$c_{N,p,q} = \frac{N!}{p!q!}\binom{N-1}{p-1,q-1} \tag{3.1}$$

Now if we rearrange the above formula, we get :

$$c_{N,p,q} = \left[\frac{p+q-1}{pq}\binom{p+q-2}{p-1}^2\right]\left[\binom{N}{N+1-(p+q)}\binom{N-1}{N+1-(p+q)}(N+1-(p+q))!\right] \tag{3.2}$$

We finally have :

$$|C_{N,p,q}| = |BT(p,q)| \times |PP(N,N-1,N+1-(p+q))| \tag{3.3}$$

In order to prove our main theorem we simply need to show that the number of unicellular partitioned bicolored maps with N edges, p white blocks and q black blocks is equal to the number of bicolored trees with p white vertices and q black vertices (with a white root) times the number of partial permutations from any subset of $[N]$ containing $N+1-(p+q)$ elements to any subset of $[N-1]$ (containing $N+1-(p+q)$ elements). To this purpose we use a bijection between the appropriate sets.

3.2. Construction of the Bijection. In this section, we construct a bijective mapping that associates to a triple $(\pi_1, \pi_2, \alpha) \in C_{N,p,q}$ an ordered bicolored tree in $BT(p,q)$ and a partial permutation in $PP(N, N-1, N+1-(p+q))$.

Ordered Bicolored Tree

Let $\pi_1^{(1)}, \dots, \pi_1^{(p)}$ and $\pi_2^{(1)}, \dots, \pi_2^{(q)}$ be the blocks of the partitions π_1 and π_2 respectively. Denote by $m_1^{(i)}$ the maximal element of the block $\pi_1^{(i)}$ $(1 \le i \le p)$ and by $m_2^{(j)}$ the maximal element of $\pi_2^{(j)}$ $(1 \le j \le q)$. We attribute the index p to the block of partition π_1 containing the element 1. Suppose that the indexation of all other blocks is arbitrary and doesn't respect any supplementary constraints. We create a labelled ordered bicolored tree T on the set of p white and q black vertices, such that white vertices have black descendants and vice versa. The root of T is the white vertex p. For every $j = 1, \dots, q$ we set that a black vertex j is a descendant of a white vertex i if the element $\beta(m_2^{(j)})$ belongs to the white block $\pi_1^{(i)}$. Similarly, for every

$i = 1, \ldots, p-1$ a white vertex i is a descendant of a black vertex j if the element $m_1^{(i)}$ belongs to the black block $\pi_2^{(j)}$. If black vertices j, k are both descendants of a white vertex i, then j is to the left of k when $\beta(m_2^{(j)}) < \beta(m_2^{(k)})$; if white vertices i, l are both descendants of a black vertex j, then i is to the left of l when $\beta^{-1}(m_1^{(i)}) < \beta^{-1}(m_1^{(l)})$. It can be proved that the previous construction allows to specify a unique path from any vertex i to the root vertex and thus, the tree T is well defined.

Removing the labels from T we obtain the bicolored ordered tree t.

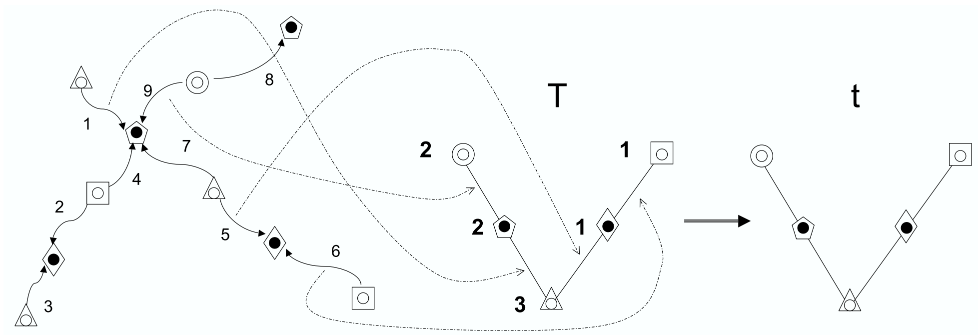


Figure 2. Construction of the Ordered Bicolored Tree

Example 3.1. Let us go back to example 2.1. We keep the previous numbering of the blocks since it verifies the condition $1 \in \pi_1^{(p)}$. For this example, $\beta(m_2^{(1)}) = \beta(6) = 5 \in \pi_1^{(3)}$ and $\beta(m_2^{(2)}) = \beta(9) = 1 \in \pi_1^{(3)}$ the black rhombus 1 and the black pentagon 2 are both descendants of the white triangle 3. Moreover, as $\beta(m_2^{(1)}) < \beta(m_2^{(2)})$ the vertex 2 is to the left of vertex 1. Further, $m_1^{(1)} = 6 \in \pi_2^{(1)}$, $m_1^{(2)} = 9 \in \pi_2^{(2)}$ and hence the white circle 1 is descendant of the black pentagon 1, while the white square 2 is descendant of the black rhombus 2. Thus, we construct first the tree T then, removing the labels, get the tree t (see Figure 3).

Partial Permutation

The construction of the partial permutation contains two main steps.

(i) *Relabelling permutations.* Consider the *reverse-labelled* bicolored tree t' resulting from the labelling of t, based on two independant reverse-labelling procedures for white and black vertices. The root is labelled p, the white vertices at level 2 are labelled from right to left, beginning with $p-1$, proceeding by labelling from right to left white vertices at level 4 and all the other even levels until reaching the leftmost white vertex at the top even level labelled by 1. The black vertices at level 1 are labelled from right to left, beginning with q, and following by labelling of the black vertices at all the other odd levels from left to right until reaching the leftmost vertex of the top odd level labelled by 1. Trees T and t' give two, possibly different, labellings of t. Suppose that the (black or white) vertex of t labelled i in T is labelled j in t'. Then define $\pi_1^j = \pi_1^{(i)}$ for white vertex and $\pi_2^j = \pi_2^{(i)}$ for black vertex, repeat this re-indexing for all white and all black vertices. We obtain a different indexing $\pi_1^1, \ldots, \pi_1^p$ of the white blocks of partition π_1; and a different indexing $\pi_2^1, \ldots, \pi_2^q$ of the black blocks of partition π_2. The reader can easily see, that $\pi_1^p = \pi_1^{(p)}$. Let ω^i and υ^j be the strings given by writing the elements of π_1^i and π_2^j in increasing order. Denote $\omega = \omega^1 \ldots \omega^p$ and $\upsilon = \upsilon^1 \ldots \upsilon^q$ concatenations of $\omega^1, \ldots, \omega^p$ and $\upsilon^1, \ldots, \upsilon^q$ respectively. We define $\lambda \in S_N$ by setting ω the first line and $[N]$ the second line of λ in the two-line representation of λ. Similarly, we define $\nu \in S_N$ by setting υ the first line and $[N]$ the second line of ν in the two-line representation of ν.

Example 3.2. Let us continue Example 3.1 by constructing relabelling permutations λ and ν. Figure 3 put the tree T and the reversed-labelled tree t' side by side that gives quite a natural

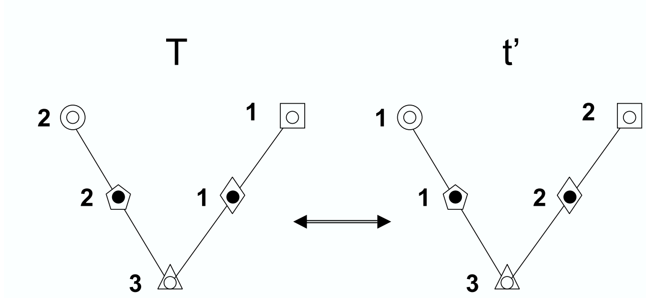


FIGURE 3. Relabelling of the Blocks

illustration of the block relabelling:

$$\begin{aligned}\pi_1^1 &= \pi_1^{(2)}, \pi_1^2 = \pi_1^{(1)}, \pi_1^3 = \pi_1^{(3)},\\ \pi_2^1 &= \pi_2^{(2)}, \pi_2^2 = \pi_2^{(1)}\end{aligned}$$

The strings ω^i and υ^j are given by :

$$\begin{aligned}\omega^1 &= 89, \quad \omega^2 = 246, \ \omega^3 = 1357,\\ \upsilon^1 &= 14789, \upsilon^2 = 2356\end{aligned}$$

We construct now the relabelling permutations λ and ν.

$$\lambda = \left(\begin{array}{cc|ccc|cccc} 8 & 9 & 2 & 4 & 6 & 1 & 3 & 5 & 7\\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\end{array}\right) \qquad \nu = \left(\begin{array}{ccccc|cccc} 1 & 4 & 7 & 8 & 9 & 2 & 3 & 5 & 6\\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\end{array}\right)$$

Figures 4 depicts this two new labellings on our example.

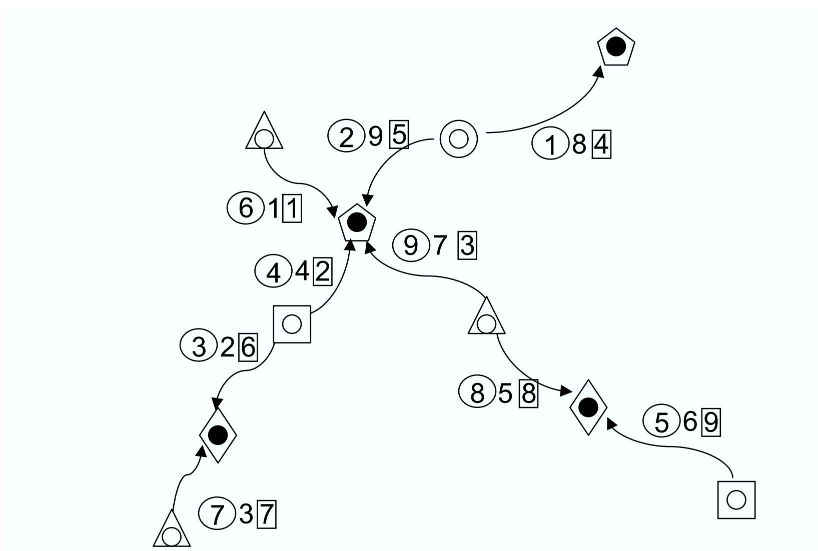


FIGURE 4. Relabellings of the Partitioned Bicolored Map

(ii) Partial permutation We can now introduce a partial permutation that gives an insight both on the connexion between the λ and ν relabelling and on the structure of the partitioned bicolored map. Let S be the subset of $[N]$ containing all the edges of the map that were not used to construct the bicolored tree. Namely :

$$S = [N] \setminus \{m_1^1, m_1^2, \ldots, m_1^{p-1}, \beta(m_2^1), \ldots, \beta(m_2^q)\} \tag{3.4}$$

We define the partial permutation σ on set $[N]$ as the composition of previously defined permutations applied to relabelling of S by λ:

$$\sigma = \nu \circ \beta^{-1} \circ \lambda^{-1} |_{\lambda(S)} \tag{3.5}$$

In Lemmae 3.4 and 3.5 we show that σ is a bijection between two subsets of $N+1-(p+q)$ elements and that its image set is included in $[N-1]$.

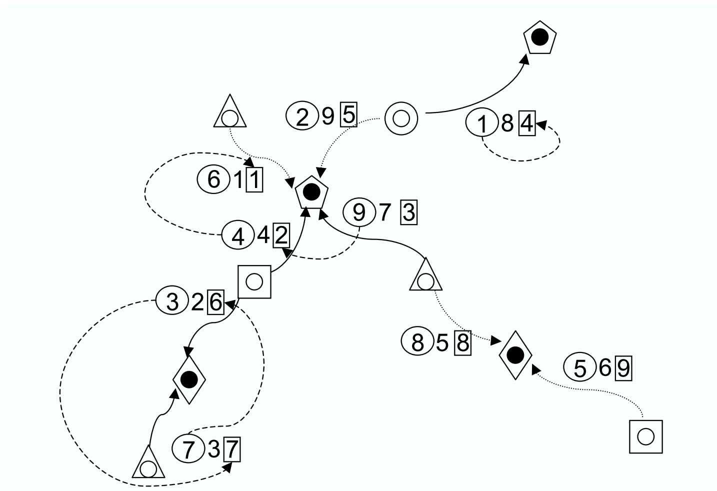


Figure 5. Connections through σ between λ and ν relabelling

Example 3.3. On the example previously described the set S is equal to :

$$S = \{2, 3, 4, 7, 8\}$$

The partial permutation σ is defined by :

$$\sigma = \begin{pmatrix} 1 & 3 & 4 & 7 & 9 \\ 4 & 7 & 1 & 6 & 2 \end{pmatrix}$$

The set of vertices that were not used to construct the tree and their connections to the map through σ can be viewed on Figure 5.

Lemma 3.4. *The cardinal of the set S defined above verifies $\mid S \mid = N + 1 - (p + q)$*

Proof. To prove the assertion of this lemma we will show the equivalent statement :

$$\{m_1^1, m_1^2, \ldots, m_1^{p-1}\} \cap \{\beta(m_2^1), \ldots, \beta(m_2^q)\} = \emptyset$$

Assume that there exist i, j, $i = 1, \ldots, p-1$, $j = 1, \ldots, q$, such that

$$\beta(m_2^j) = m_1^i \tag{3.6}$$

Then as the blocks of π_2 are stable by β we have $m_1^i \in \pi_2^j$ and $m_1^i \leq m_2^j$. As the blocks of π_1 are stable by α, the assumption (3.6) also implies that $\alpha\beta(m_2^j) = \gamma(m_2^j) \in \pi_1^i$. Hence, $\gamma(m_2^j) \leq m_1^i$. Combining these two inequalities, we have $\gamma(m_2^j) \leq m_2^j$ that occurs only if $m_2^j = N$. In this case, $\gamma(m_2^j) = 1$ and $1 \in \pi_1^i$, i.e. $i = p$ which is a contradiction ■

Lemma 3.5. *The element N does not belong to the image of permutation σ.*

Proof. Let us remark that according to the construction of the relabelling permutation ν, we have $N = \nu(m_2^q)$. Besides,

$$\nu(m_2^q) = \nu \circ \beta^{-1} \circ \lambda^{-1}(\lambda(\beta(m_2^q)))$$

Thus, as $\lambda(\beta(m_2^q))$ does not belong to $\lambda(S)$, the element N does not belong to the image of permutation σ
■

Bijective Mapping

Let us denote by $\Theta_{N,p,q}$ the mapping defined by :

$$\begin{array}{lllll} \Theta_{N,p,q} & : & C_{N,p,q} & \longrightarrow & BT(p,q) \times PP(N, N-1, N+1-(p+q)) \\ & & (\pi_1, \pi_2, \alpha) & \longmapsto & (t, \sigma) \end{array} \tag{3.7}$$

We prove in the following section that the mapping $\Theta_{N,p,q}$ is actually a bijection.

4. Proof of the Bijection

4.1. Injectivity. Let (t, σ) be the image of some triple $(\pi_1, \pi_2, \alpha) \in C_{N,p,q}$ by $\Theta_{N,p,q}$. We show in a constructive fashion that (π_1, π_2, α) is actually uniquely determined by (t, σ).

First we use the tree t and the integers lacking in the two lines of σ to find out the number of elements in each block and the extension of σ to the whole set $[N]$. By construction of σ, the integers lacking in its first line $\lambda(S)$ are the elements

$$\lambda(m_1^1), \ldots, \lambda(m_1^{p-1}), \lambda(\beta(m_2^1)), \ldots, \lambda(\beta(m_2^q))$$

Now, if black vertex j is a descendant of a white vertex i in the reversed-labelled tree t' then $\beta(m_2^j) \in \pi_1^i$ for $j = 1, \ldots, q$. Due to this property we know exactly how the elements $\beta(m_2^1), \ldots, \beta(m_2^q)$ are distributed between the white blocks. Moreover, by definition of λ any element of $\lambda(\pi_1^i)$ is strictly less than any element of $\lambda(\pi_1^j)$ for all $i < j$. Thus, to recover the exact order on the elements lacking in the first line of σ, it remains to establish the order on the lacking elements belonging to the same block $\lambda(\pi_1^i)$ that are not the maximum one (obviously the greatest) if any. These elements correspond to the set of descendants of the white vertex i in t'. As we have defined that a black vertex j_1 is on the left of a black vertex j_2, descendant of the same vertex i, if and only if $\beta(m_2^{j_1}) \le \beta(m_2^{j_2})$ and the restriction of λ to any block of π_1 is an increasing function, their order is naturally induced by the left to right order on the set of descendants of i in the reversed-labelled tree t'.

Consider the set $\nu \circ \beta^{-1}(S)$ in the second line of σ. The lacking elements are

$$\nu(m_2^1), \ldots, \nu(m_2^q), \nu(\beta)^{-1}(m_1^1), \ldots, \nu(\beta)^{-1}(m_1^{p-1})$$

Similarly to the first line of σ, we use the structure of t', the relation between ν and t' as well as the fact that $\nu(\beta)^{-1}(m_1^{i_1}) \le \nu(\beta)^{-1}(m_1^{i_2})$ if i_1 and i_2 are descendant of the same black vertex and i_1 is on the left of i_2 to order these elements. Once the order on both of the sets of lacking elements is established, the lacking integers can be uniquely identified with these elements. Hence, the extension $\overline{\sigma} = \nu \circ \beta^{-1} \circ \lambda^{-1}$ of the partial permutation σ to the whole set $[N]$ is uniquely determined since

$$\forall i \in [p-1], \qquad \overline{\sigma}(\lambda(m_1^i)) = \nu(\beta^{-1}(m_1^i)) \tag{4.1}$$

$$\forall j \in [q], \qquad \overline{\sigma}(\lambda(\beta(m_2^j))) = \nu(m_2^j) \tag{4.2}$$

Now, the knowledge of $\lambda(m_1^1), \ldots, \lambda(m_1^{p-1})$ and $\nu(m_2^1), \ldots, \nu(m_2^q)$ allows us to determine the number of elements in each of the blocks of partitions $\lambda(\pi_1) = \lambda(\pi_1^1), \ldots, \lambda(\pi_1^p)$ and $\nu(\pi_2) = \nu(\pi_2^1), \ldots, \nu(\pi_2^q)$. Indeed, the blocks of the above partitions are intervals:

$$\begin{array}{rcl} \lambda(\pi_1^1) & = & [\lambda(m_1^1)] \\ \lambda(\pi_1^i) & = & [\lambda(m_1^i)] \setminus [\lambda(m_1^{i-1})] \text{ for } 2 \le i \le p-1 \\ \lambda(\pi_1^p) & = & [N] \setminus [\nu(m_1^{p-1})] \\ \nu(\pi_2^1) & = & [\nu(m_1^1)] \\ \nu(\pi_2^i) & = & [\nu(m_1^i)] \setminus [\lambda(m_1^{i-1})] \text{ for } 2 \le i \le q \end{array}$$

Hence, $\lambda(\pi_1)$ and $\nu(\pi_2)$ are uniquely determined by (t, σ). Besides, since π_2 is stable by β, we can use $\overline{\sigma}$ to recover $\lambda(\pi_2)$. Indeed:

$$
\begin{aligned}
\overline{\sigma}^{-1}(\nu(\pi_2)) &= \lambda \circ \beta \circ \nu^{-1}(\nu(\pi_2)) \\
\overline{\sigma}^{-1}(\nu(\pi_2)) &= \lambda \circ \beta(\pi_2) \\
\overline{\sigma}^{-1}(\nu(\pi_2)) &= \lambda(\pi_2) \qquad (4.3)
\end{aligned}
$$

Then as $\overline{\sigma}$ and $\nu(\pi_2)$ are uniquely determined, so is $\lambda(\pi_2)$.

Example 4.1. We give here an illustration of the first steps of the injectivity proof. Let us suppose that we are given the parameters $N = 10, p = 3, q = 2$, the following partial permutation

and the bicolored ordere

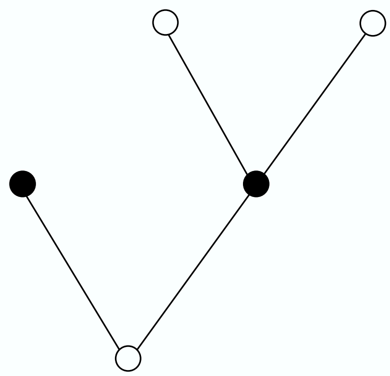


Figure 6. A Bicolored Tree

Consider the set $\lambda(S)$ in the first line of σ. Assuming a reverse labelling of the tree, the elements lacking in $\lambda(S)$ are

$$\lambda(m_1^1), \lambda(m_1^2), \lambda(\beta(m_2^1), \lambda(\beta(m_2^2)$$

The numbers lacking in $\lambda(S)$ to complete it up to $\lambda([N])$ are $1, 2, 7, 9$. According to the previous remarks, we can identify all these numbers in the following way:

$$\lambda(m_1^1) = 1, \lambda(m_1^2) = 2, \lambda(\beta(m_2^1)) = 7, \lambda(\beta(m_2^2)) = 9$$

Consider the set $(\nu \circ \beta)^{-1}(S)$ in the second line of σ. The elements lacking are

$$\nu(m_2^1), \nu(m_2^2), \nu(\beta)^{-1}(m_1^1), \nu(\beta)^{-1}(m_2^2)$$

We have

$$\nu(m_2^1) = 2, \nu(\beta)^{-1}(m_1^1) = 3, \nu(\beta)^{-1}(m_2^2) = 9, \nu(m_2^2) = 10$$

Now we can extend σ to the permutation $\overline{\sigma}$ on the set $[N]$:

$$\overline{\sigma} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 9 & 4 & 6 & 3 & 1 & 2 & 8 & 10 & 7 \end{pmatrix}$$

Note, that as $\lambda(m_1^1) = 1, \lambda(m_1^2) = 2, \nu(m_2^1) = 2, \nu(m_2^2) = 10$, we also can identify the images of white blocks by λ and images of black blocks by ν:

$$
\begin{aligned}
\lambda(\pi_1^1) &= \{1\}, \quad \lambda(\pi_1^2) = \{2\}, \lambda(\pi_1^3) = \{3, 4, 5, 6, 7, 8, 9, 10\}, \\
\nu(\pi_2^1) &= \{1, 2\}, \nu(\pi_2^2) = \{3, 4, 5, 6, 7, 8, 9, 10\}
\end{aligned}
$$

Using (4.3) we obtain the relabelling of partition π_2:

$$
\begin{aligned}
\lambda(\pi_2^1) &= \{6, 7\}, \qquad (4.4) \\
\lambda(\pi_2^2) &= \{1, 2, 3, 4, 5, 8, 9, 10\}
\end{aligned}
$$

Now let us show that λ and ν are uniquely determined as well. As $1 \in \pi_1^p$ and λ is an increasing function on each block of π_1, $\lambda(1)$ is necessarily the least element of $\lambda(\pi_1^p)$. Let then $\lambda(\pi_2^k)$ be the block of $\lambda(\pi_2)$ such that $\lambda_1 \in \lambda(\pi_2^k)$. As ν is an increasing function on each block of π_2, necessarily $\nu(1)$ is the least element of $\nu(\pi_2^k)$.

Now assume that for a given i in $[N-1]$, $\lambda(1), \dots, \lambda(i)$ and $\nu(1), \dots, \nu(i)$ have been determined. As π_1 is stable by α, necessarily $\beta(i)$ and $i+1 = \gamma(i) = \alpha \circ \beta(i)$ belong to the same block of π_1. Hence $\lambda(i+1)$ and $\lambda(\beta(i))$ belong to the same block of $\lambda(\pi_1)$. But:

$$\lambda(\beta(i)) = \lambda \circ \beta \circ \nu^{-1}(\nu(i)) = \overline{\sigma}^{-1}(\nu(i)) \tag{4.5}$$

As a consequence, $\lambda(i+1)$ and $\overline{\sigma}^{-1}(\nu(i))$ belong to the same block of $\lambda(\pi_1)$. Finally, as λ is an increasing function on each block of π_1, $\lambda(i+1)$ is necessarily the least element of the block of $\lambda(\pi_1)$ containing $\overline{\sigma}^{-1}(\nu(i))$ that has not been used yet to identify $\lambda(1), \dots, \lambda(i)$.

Let us denote by $\lambda(\pi_2^l)$ the block of $\lambda(\pi_2)$ containing $\lambda(i+1)$. Since ν is an increasing function on each block of π_2, $\nu(i+1)$ is uniquely determined as being the least element of the block $\nu(\pi_2^l)$ that has not already been used to identify $\nu(1), \dots, \nu(i)$. By iterating the above procedure for all the integers in $[N-1]$ we see that λ and ν are uniquely determined.

To end this proof, we remark that :

$$\begin{aligned}
\pi_1 &= \lambda^{-1}(\lambda(\pi_1)) \\
\pi_2 &= \nu^{-1}(\nu(\pi_2)) \\
\alpha &= \gamma \circ \beta^{-1} = \gamma \circ \nu^{-1} \circ \overline{\sigma} \circ \lambda
\end{aligned}$$

As a result, at most one triple (π_1, π_2, α) can be associated by $\Theta_{N,p,q}$ to (t, σ). Moreover, if such a triple exists, it can be computed using the description of $\Theta_{N,p,q}^{-1}$ given by the above proof.

Example 4.2. We apply the iterative reconstruction of λ and ν to the previous example. A table of three lines and N columns will be used to sum up the available information on λ and ν on each step of the reconstruction: the first line is given by $[N]$ and represents the initial labelling γ of the edges of the partitioned map, the second and third lines represent the relabellings of the same edges by λ and ν. We initialize the procedure by putting $3 = \min(\lambda(\pi_1^3))$ at the first position of the line for λ:

$$\begin{array}{cccccccccccc}
\gamma & : & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\lambda & : & 3 & * & * & * & * & * & * & * & * & * \\
\nu & : & * & * & * & * & * & * & * & * & * & *
\end{array}$$

Now, looking at equations (4.4) we establish, that the element 3 belongs to the second black block $\lambda(\pi_2^2)$. As the least element of $\nu(\pi_2^2)$ is 3, we put 3 in the first position of the third line of our table:

$$\begin{array}{cccccccccccc}
\gamma & : & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\lambda & : & 3 & * & * & * & * & * & * & * & * & * \\
\nu & : & 3 & * & * & * & * & * & * & * & * & *
\end{array}$$

Let us establish now what is the next white block in our partitioned map. For this goal we take the image of the last discovered element $\nu(1)$ by $\overline{\sigma}^{-1}$:

$$\overline{\sigma}^{-1}(\nu(1)) = \overline{\sigma}^{-1}(3) = 5$$

Thus $\overline{\sigma}^{-1}(\nu(1))$ belongs to π_1^3. We then deduce that $\lambda(2)$ is the least element of $\lambda(\pi_1^3)$ which has not been met yet, i.e 4. We write 4 at the second position on the line for λ:

$$\begin{array}{cccccccccccc}
\gamma & : & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\lambda & : & 3 & 4 & * & * & * & * & * & * & * & * \\
\nu & : & 3 & * & * & * & * & * & * & * & * & *
\end{array}$$

We iterate the process until λ and ν are fully reconstructed :

$$\begin{array}{cccccccccccc}
\gamma & : & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\lambda & : & 3 & 4 & 5 & 1 & 6 & 7 & 8 & 9 & 10 & 2 \\
\nu & : & 3 & 4 & 5 & 6 & 1 & 2 & 7 & 8 & 9 & 10
\end{array}$$

Once λ and ν are known, we have the partitioned map reconstructed:

$$\begin{aligned}
\pi_1 &= \{\{4\}\{10\}\{1,2,3,5,6,7,8,9\}\} \\
\pi_2 &= \{\{5,6\}\{1,2,3,4,7,8,9,10\}\} \\
\alpha &= (13256798)(4)(10)
\end{aligned}$$

The bicolored map represented by this permutation is the unicellular map of genus 2 on Figure 7.

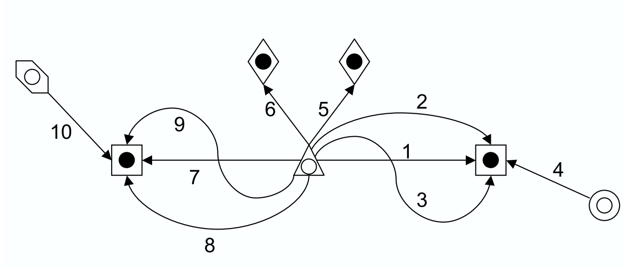


Figure 7. The Partitioned Bicolored Map Once Reconstructed

4.2. Surjectivity. Let us now proceed by showing that $\Theta_{N,p,q}$ is a surjection. Clearly, up to the reconstruction of λ and ν the first steps of the procedure described in the previous section can be applied to any couple (t,σ) belonging to $BT(p,q)\times PP(N,N-1,N-1-(p+q))$. Namely we can define the extension $\overline{\sigma}$ of σ to the whole set $[N]$ as well as the partitions $\lambda(\pi_1)$, $\lambda(\pi_2)$ and $\nu(\pi_2)$. Then we use lemma 4.3 to show that the reconstruction of λ and ν can also always be succesfully completed.

Lemma 4.3. *Given any couple (t,σ) belonging to $BT(p,q)\times PP(N,N-1,N-1-(p+q))$, the iterative procedure for the reconstruction of λ and ν can always be performed and gives a valid output in any case.*

Proof. First of all, we notice that only two reasons can prevent the procedure from being performed until its end. Either for a given i in $[N-1]$, $\overline{\sigma}^{-1}(\nu(i))$ belongs to a block of $\lambda(\pi_1)$ that has all its elements already used for the construction of $\lambda(1),\dots,\lambda(i)$ so that we cannot define $\lambda(i+1)$; or $\lambda(i+1)$ belongs to a block of $\lambda(\pi_2)$ such that the corresponding block of $\nu(\pi_2)$ has all its elements already been used for the construction of $\nu(1),\dots,\nu(i)$ and we are not able to define $\nu(i+1)$. We show by induction that this situation never occurs.

Assume that we have already successfully iterated the procedure up to $i\le N-1$. Also assume that we cannot define $\lambda(i+1)$ due to the reason stated above. We note $\lambda(\pi_1^k)$ the block containing $\overline{\sigma}^{-1}(\nu(i))$. Then:

(i) If $\lambda(\pi_1^k)$ does not contain $\lambda(1)$, the last assumption implies that $|\pi_1^k|+1$ different integers, including $\nu(i)$, used for the construction of ν have their image by $\overline{\sigma}^{-1}$ in $\lambda(\pi_1^k)$. This is of course a contradiction with the fact that $\overline{\sigma}^{-1}$ is a bijection.

(ii) If $\lambda(1)$ belongs to $\lambda(\pi_1^k)$ (thus $k=p$), we still have a contradiction. In this particular case, we only know that $|\pi_1^k|$ different integers have their image by $\overline{\sigma}^{-1}$ in $\lambda(\pi_1^p)$. However, according to our definition of $\overline{\sigma}$, $\lambda(\pi_1)$ and $\lambda(\pi_2)$, if the white vertex in t corresponding to a given block π_1^a is the direct descendant of the black vertex associated with π_2^b then :

$$\lambda(m_1^a)\in\lambda(\pi_2^b) \tag{4.6}$$

In other words we cannot have used all the elements of $\nu(\pi_2^b)$ for the construction of ν until the maximum element of $\lambda(\pi_1^a)$ (and henceforth all the elements of $\lambda(\pi_1^a)$) has been used for the construction of λ. In a similar fashion, if the black vertex associated to π_2^c is the direct descendant of the white one corresponding to π_1^d, we have:

$$\overline{\sigma}^{-1}(\nu(m_2^c))\in\lambda(\pi_1^d) \tag{4.7}$$

And all the elements of $\nu(\pi_2^c)$ must be used for the reconstruction of ν before all the elements of $\lambda(\pi_1^d)$ are used for the reconstruction of λ. To summarize, all the elements of a block associated to a vertex x (either black or white) are not used for the construction of λ and ν until all the elements of the blocks associated with vertices that are descendant of x are used for the same construction. As π_1^p is associated with the root of t, if all the elements of $\lambda(\pi_1^p)$ have already been used for the construction of λ and ν, it means that all the elements of all the other blocks of $\lambda(\pi_1)$ and $\nu(\pi_2)$ have been already used as well. The reconstruction is hence completed and $i = N$. That is a contradiction with our assumption $i \leq N-1$.

Once $\lambda(i+1)$ is found, we notice that $\nu(i+1)$ can always be defined. Indeed, if $\lambda(i+1)$ belonged to a block $\lambda(\pi_2^l)$ such that all the elements of $\nu(\pi_2^l)$ have been already used to construct ν, it would mean that $|\pi_2^l|+1$ different integers belong to $\lambda(\pi_2^l)$, which is a contradiction. Our induction is completed by an obvious remark that $\lambda(1)$ and $\nu(1)$ can always be defined. ∎

For the final step of this proof we need to show that once λ and ν are constructed the permutation α defined by

$$\alpha = \gamma \circ \nu^{-1} \circ \overline{\sigma} \circ \lambda \tag{4.8}$$

verifies the two following conditions:

$$\alpha(\pi_1) = \pi_1 \tag{4.9}$$

$$\alpha^{-1}\gamma(\pi_2) = \pi_2 \tag{4.10}$$

Condition (4.10) comes from the fact that we have defined:

$$\lambda(\pi_2) = \overline{\sigma}^{-1}(\pi_2) \tag{4.11}$$

Thus,

$$\pi_2 = \lambda^{-1} \circ \overline{\sigma}^{-1} \circ \nu \circ \gamma^{-1} \circ \gamma(\pi_2) \tag{4.12}$$

and by consequence

$$\pi_2 = \alpha^{-1} \circ \gamma(\pi_2) \tag{4.13}$$

Condition (4.9) can be shown using the fact that for all i in $[N]$, $\lambda(i)$ and $\overline{\sigma}^{-1} \circ \nu \circ \gamma^{-1}(i)$ belong to the same block of $\lambda(\pi_1)$. Hence, $\lambda^{-1} \circ \overline{\sigma}^{-1} \circ \nu \circ \gamma^{-1}(i)$ and i belong to the same block of π_1. Finally, the blocks of π_1 are stable by α^{-1} and henceforth by α.

References

[1] N.M. Adrianov, *An analogue of the Harer-Zagier formula for unicellular two-color maps*, Funct. Anal. Appl., **31(3)**, 149–155, 1998

[2] R. Cori, A. Machi, *Maps, hypermaps and their automorphisms: a survey I, II, III*, Expositiones Math., **10**, 403–427, 429–447, 449–467, 1992

[3] I.P. Goulden and D.M. Jackson, *Combinatorial enumeration*, John Wiley & Sons Inc., New York, 1983

[4] I.P. Goulden and A. Nica, *A direct bijection for the Harer-Zagier formula*, J. Combinatorial Theory (A) **111**, 224–238, 2005

[5] A. Goupil, G. Schaeffer, *Factoring n-cycles and counting maps of given genus*, Europ. J. Combinatorics, **19(7)**, 819-834, 1998

[6] J. Harer and D. Zagier, *The Euler characteristic of the moduli space of curves*, Inventiones Mathematicae, **85**, 457–486, 1986

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

[8] G. Schaeffer, *Conjugaison d'arbres et cartes combinatoires aléatoires*, Ph.D. Thesis, l'Université Bordeaux I, 1998

[9] D. Stanton and D. White, *Constructive combinatorics*, Springer-Verlag, New York, 1986

LIX – Laboratoire d'Informatique de l' Ecole Polytechnique, 91128 Palaiseau Cedex, FRANCE
E-mail address: ekaterina.vassilieva@lix.polytechnique.fr

LIX – Laboratoire d'Informatique de l' Ecole Polytechnique, 91128 Palaiseau Cedex, FRANCE
E-mail address: schaeffe@lix.polytechnique.fr

Part III
Posters
Affiches

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Strong Descent Numbers and Turán Type Theorems (Extended Abstract)

Ron M. Adin and Yuval Roichman

Abstract. For a permutation π in the symmetric group S_n let the *total degree* be its valency in the Hasse diagram of the strong Bruhat order on S_n, and let the *down degree* be the number of permutations which are covered by π in the strong Bruhat order. The maxima of the total degree and the down degree and their values at a random permutation are computed. Proofs involve variants of a classical theorem of Turán from extremal graph theory.

Résumé. Pour une permutation π du groupe symétrique S_n, on définit le degré total de π comme sa valence dans le diagramme de Hasse de l'ordre de Bruhat fort sur S_n, et le degré bas de]*pi* comme le nombre de permutations couvertes par π dans l'ordre de Bruhat fort. Nous calculons les valeurs maximales pour le degré total et le degré bas, ainsi que leur valeurs pour une permutation aléatoire. Nos démonstrations utilisent des variantes d'un théorème de Turán provenant de la théorie des graphes extrémaux.

1. The Down, Up and Total Degrees

Definition 1.1. *For a permutation $\pi \in S_n$ let the* down degree $d_-(\pi)$ *be the number of permutations in S_n which are covered by π in the strong Bruhat order. Let the* up degree $d_+(\pi)$ *be the number of permutations which cover π in this order. The* total degree *of π is the sum*

$$d(\pi) := d_-(\pi) + d_+(\pi),$$

i.e., the valency of π in the Hasse diagram of the strong Bruhat order.

Explicitly, for $1 \le a < b \le n$ let $t_{a,b} = t_{b,a} \in S_n$ be the transposition interchanging a and b, and for $\pi \in S_n$ let

$$\ell(\pi) := \min\{k \mid \pi = s_{i_1} s_{i_2} \cdots s_{i_k}\}$$

be the *length* of π with respect to the standard Coxeter generators $s_i = t_{i,i+1}$ $(1 \le i < n)$ of S_n. Then

$$\begin{aligned} d_-(\pi) &= \#\{t_{a,b} \mid \ell(t_{a,b}\pi) = \ell(\pi) - 1\} \\ d_+(\pi) &= \#\{t_{a,b} \mid \ell(t_{a,b}\pi) = \ell(\pi) + 1\} \\ d(\pi) = d_-(\pi) + d_+(\pi) &= \#\{t_{a,b} \mid \ell(t_{a,b}\pi) = \ell(\pi) \pm 1\} \end{aligned}$$

For the general definitions and other properties of the weak and strong Bruhat orders see, e.g., [**10**, Ex. 3.75] and [**3**, §§2.1, 3.1].

2000 *Mathematics Subject Classification.* Primary 20B30; Secondary 05C35.

Key words and phrases. Bruhat order, descent number, symmetric group, Turan graph.

Research of both authors was supported in part by the Israel Science Foundation, founded by the Israel Academy of Sciences and Humanities, and by the EC's IHRP Programme, within the Research Training Network "Algebraic Combinatorics in Europe", grant HPRN-CT-2001-00272.

We shall describe $\pi \in S_n$ by its sequence of values $[\pi(1), \ldots, \pi(n)]$.

Observation 1.2. *π covers σ in the strong Bruhat order on S_n if and only if there exist $1 \le i < k \le n$ such that*

(1) $b := \pi(i) > \pi(k) =: a$.
(2) $\sigma = t_{a,b}\pi$, *i.e.*, $\pi = [\ldots, b, \ldots, a, \ldots]$ *and* $\sigma = [\ldots, a, \ldots, b, \ldots]$.
(3) *There is no* $i < j < k$ *such that* $a < \pi(j) < b$.

Corollary 1.3. *For every* $\pi \in S_n$

$$d_-(\pi) = d_-(\pi^{-1}).$$

Example 1.4. *In* S_3, $d_-[123] = 0$, $d_-[132] = d_-[213] = 1$, *and* $d_-[321] = d_-[231] = d_-[312] = 2$. *On the other hand,* $d[321] = d[123] = 2$ *and* $d[213] = d[132] = d[312] = d[231] = 3$.

Remark 1.5. *The classical descent number of a permutation π in the symmetric group S_n is the number of permutations in S_n which are covered by π in the (right)* weak *Bruhat order. Thus, the down degree may be considered as a "strong descent number".*

Definition 1.6. *For* $\pi \in S_n$ *denote*

$$D_-(\pi) := \{t_{a,b} \,|\, \ell(t_{a,b}\pi) = \ell(\pi) - 1\},$$

the strong descent set *of* π.

Example 1.7. *The strong descent set of* $\pi = [7, 9, 5, 2, 3, 8, 4, 1, 6]$ *is*

$$D_-(\pi) = \{t_{1,2}, t_{1,3}, t_{1,4}, t_{2,5}, t_{3,5}, t_{4,5}, t_{4,8}, t_{5,7}, t_{5,9}, t_{6,7}, t_{6,8}, t_{8,9}\}.$$

Remark 1.8. *Generalized pattern avoidance, involving strong descent sets, was applied by Yong and Woo* [**12**] *to determine which Schubert varieties are Gorenstein.*

Proposition 1.9. *The strong descent set $D_-(\pi)$ uniquely determines the permutation π.*

Proof. By induction on n. The claim clearly holds for $n = 1$.

Let π be a permutation in S_n, and let $\bar\pi \in S_{n-1}$ be the permutation obtained by deleting the value n from π. Note that, by Observation 1.2,

$$D_-(\bar\pi) = D_-(\pi) \setminus \{t_{a,n} \,|\, 1 \le a < n\}.$$

By the induction hypothesis $\bar\pi$ is uniquely determined by this set. Hence it suffices to determine the position of n in π.

Now, if $j := \pi^{-1}(n) < n$ then clearly $t_{\pi(j+1),n} \in D_-(\pi)$. Moreover, by Observation 1.2, $t_{a,n} \in D_-(\pi) \Longrightarrow a \ge \pi(j+1)$. Thus $D_-(\pi)$ determines

$$\bar\pi(j) = \pi(j+1) = \min\{a \,|\, t_{a,n} \in D_-(\pi)\},$$

and therefore determines j. Note that this set of a's is empty if and only if $j = n$. This completes the proof.

□

2. Maximal Down Degree

In this section we compute the maximal value of the down degree on S_n and find all the permutations achieving the maximum. We prove

Proposition 2.1. *For every positive integer* n

$$\max\{d_-(\pi) |\ \pi \in S_n\} = \lfloor n^2/4 \rfloor.$$

Remark 2.2. *The same number appears as the order dimension of the strong Bruhat poset* [**8**]. *An upper bound on the maximal down degree for finite Coxeter groups appears in* [**5**, Prop. 3.4].

For the proof of Proposition 2.1 we need a classical theorem of Turán.

DEFINITION 2.3. *Let $r \le n$ be positive integers. The* Turán graph $T_r(n)$ *is the complete r-partite graph with n vertices and all parts as equal in size as possible, i.e., each size is either $\lfloor n/r \rfloor$ or $\lceil n/r \rceil$. Denote by $t_r(n)$ the number of edges of $T_r(n)$.*

THEOREM 2.4. [**11**] [**4**, IV, Theorem 8] **(Turán's Theorem)**

(1) *Every graph of order n with more than $t_r(n)$ edges contains a complete subgraph of order $r+1$.*
(2) *$T_r(n)$ is the unique graph of order n with $t_r(n)$ edges that does not contain a complete subgraph of order $r+1$.*

We shall apply the special case $r = 2$ (due to Mantel) of Turán's theorem to the following graph.

DEFINITION 2.5. *The* strong descent graph *of $\pi \in S_n$, denoted $\Gamma_-(\pi)$, is the undirected graph whose set of vertices is $\{1, \ldots, n\}$ and whose set of edges is*

$$\{\{a, b\} \mid t_{a,b} \in D_-(\pi)\}.$$

By definition, the number of edges in $\Gamma_-(\pi)$ equals $d_-(\pi)$.

REMARK 2.6. *Permutations for which the strong descent graph is connected are called* indecomposable. *Their enumeration was studied in* [**6**]*; see* [**7**, pp. 7–8]. *The number of components in $\Gamma_-(\pi)$ is equal to the number of* global descents *in πw_0 (where $w_0 := [n, n-1, \ldots, 1]$), which were introduced and studied in* [**2**, Corollaries 6.3 and 6.4].

LEMMA 2.7. *For every $\pi \in S_n$, the strong descent graph $\Gamma_-(\pi)$ is triangle-free.*

PROOF. Assume that $\Gamma_-(\pi)$ contains a triangle. Then there exist $1 \le a < b < c \le n$ such that $t_{a,b}, t_{a,c}, t_{b,c} \in D_-(\pi)$. By Observation 1.2,

$$t_{a,b}, t_{b,c} \in D_-(\pi) \Longrightarrow \pi^{-1}(c) < \pi^{-1}(b) < \pi^{-1}(a) \Longrightarrow t_{a,c} \notin D_-(\pi).$$

This is a contradiction. $\square$

PROOF. (OF PROPOSITION 2.1) By Theorem 2.4(1) together with Lemma 2.7, for every $\pi \in S_n$

$$d_-(\pi) \le t_2(n) = \lfloor n^2/4 \rfloor.$$

Equality holds since

$$d_-([\lfloor n/2 \rfloor + 1, \lfloor n/2 \rfloor + 2, \ldots, n, 1, 2, \ldots, \lfloor n/2 \rfloor]) = \lfloor n^2/4 \rfloor.$$

$\square$

Next we classify (and enumerate) the permutations which achieve the maximal down degree.

LEMMA 2.8. *Let $\pi \in S_n$ be a permutation with maximal down degree. Then π has no decreasing subsequence of length 4.*

PROOF. Assume that $\pi = [\ldots d \ldots c \ldots b \ldots a \ldots]$ with $d > c > b > a$ and $\pi^{-1}(a) - \pi^{-1}(d)$ minimal. Then $t_{a,b}, t_{b,c}, t_{c,d} \in D_-(\pi)$ but, by Observation 1.2, $t_{a,d} \notin D_-(\pi)$. It follows that $\Gamma_-(\pi)$ is not a complete bipartite graph, since $\{a, b\}$, $\{b, c\}$, and $\{c, d\}$ are edges but $\{a, d\}$ is not. By Lemma 2.7, combined with Theorem 2.4(2), the number of edges in $\Gamma_-(\pi)$ is less than $\lfloor n^2/4 \rfloor$. $\square$

PROPOSITION 2.9. *For every positive integer n*

$$\#\{\pi \in S_n \mid d_-(\pi) = \lfloor n^2/4 \rfloor\} = \begin{cases} n, & \text{if } n \text{ is odd;} \\ n/2, & \text{if } n \text{ is even.} \end{cases}$$

Each such permutation has the form

$$\pi = [t+m+1, t+m+2, \ldots, n, t+1, t+2, \ldots, t+m, 1, 2, \ldots, t],$$

where $m \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$ and $1 \le t \le n-m$. Note that $t = n-m$ (for m) gives the same permutation as $t = 0$ (for $n-m$ instead of m).

PROOF. It is easy to verify the claim for $n \le 3$. Assume $n \ge 4$.

Let $\pi \in S_n$ with $d_-(\pi) = \lfloor n^2/4 \rfloor$. By Theorem 2.4(2), $\Gamma_-(\pi)$ is isomorphic to the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. Since $n \ge 4$, each side of the graph contains at least two vertices. Let $1 = a < b$ be two vertices on one side, and $c < d$ two vertices on the other side of the graph. Since $t_{b,c}, t_{b,d} \in D_-(\pi)$, there are three possible cases:

(1) $b < c$, and then $\pi = [\ldots c \ldots d \ldots b \ldots]$ (since $\pi = [\ldots d \ldots c \ldots b \ldots]$ contradicts $t_{b,d} \in D_-$).
(2) $c < b < d$, and then $\pi = [\ldots d \ldots b \ldots c \ldots]$.
(3) $d < b$, and then $\pi = [\ldots b \ldots c \ldots d \ldots]$ (since $\pi = [\ldots b \ldots d \ldots c \ldots]$ contradicts $t_{b,c} \in D_-$).

The same also holds for a instead of b, but then cases 2 and 3 are impossible since $a = 1 < c$. Thus necessarily c appears before d in π, and case 2 is therefore impossible for *any* b on the same side as $a = 1$. In other words: no vertex on the same side as $a = 1$ is intermediate, either in position (in π) or in value, to c and d.

Assume now that n is even. The vertices not on the side of 1 form (in π) a block of length $n/2$ of numbers which are consecutive in value as well in position. They also form an increasing subsequence of π, since $\Gamma_-(\pi)$ is bipartite. The numbers preceding them are all larger in value, and are increasing; the numbers succeeding them are all smaller in value, are increasing, and contain 1. It is easy to check that each permutation π of this form has maximal $d_-(\pi)$. Finally, π is completely determined by the length $1 \le t \le n/2$ of the last increasing subsequence.

For n odd one obtains a similar classification, except that the length of the side not containing 1 is either $\lfloor n/2 \rfloor$ or $\lceil n/2 \rceil$. This completes the proof.

□

3. Maximal Total Degree

Obviously, the maximal value of the total degree $d = d_- + d_+$ cannot exceed $\binom{n}{2}$, the total number of transpositions in S_n. This is slightly better than the bound $2\lfloor n^2/4 \rfloor$ obtainable from Proposition 2.1. The actual maximal value is smaller.

THEOREM 3.1. *For $n \ge 2$, the maximal total degree in the Hasse diagram of the strong Bruhat order on S_n is*

$$\lfloor n^2/4 \rfloor + n - 2.$$

In order to prove this result, associate with each permutation $\pi \in S_n$ a graph $\Gamma(\pi)$, whose set of vertices is $\{1, \ldots, n\}$ and whose set of edges is

$$\{\{a,b\} \mid \ell(t_{a,b}\pi) - \ell(\pi) = \pm 1\}.$$

This graph has many properties; e.g., it is K_5-free and is the edge-disjoint union of two triangle-free graphs on the same set of vertices. However, these properties are not strong enough to imply the above result. A property which does imply it is the following bound on the minimal degree.

LEMMA 3.2. *There exists a vertex in $\Gamma(\pi)$ with degree at most $\lfloor n/2 \rfloor + 1$.*

PROOF. Assume, on the contrary, that each vertex in $\Gamma(\pi)$ has at least $\lfloor n/2 \rfloor + 2$ neighbors. This applies, in particular, to the vertex $\pi(1)$. Being the first value of π, the neighborhood of $\pi(1)$ in $\Gamma(\pi)$, viewed as a subsequence of $[\pi(2), \ldots, \pi(n)]$, consists of a shuffle of a decreasing sequence of numbers larger than $\pi(1)$ and an increasing sequence of numbers smaller than $\pi(1)$. Let a be the rightmost neighbor of $\pi(1)$. The intersection of the neighborhood of a with the neighborhood of $\pi(1)$ is of cardinality at most two. Thus the degree of a is at most

$$n - (\lfloor n/2 \rfloor + 2) + 2 = \lceil n/2 \rceil \leq \lfloor n/2 \rfloor + 1,$$

which is a contradiction.

$\square$

PROOF. (OF THEOREM 3.1) First note that, by definition, the total degree of $\pi \in S_n$ in the Hasse diagram of the strong Bruhat order is equal to the number of edges in $\Gamma(\pi)$. We will prove that this number $e(\Gamma(\pi)) \leq \lfloor n^2/4 \rfloor + n - 2$, by induction on n.

The claim is clearly true for $n = 2$. Assume that the claim holds for $n-1$, and let $\pi \in S_n$. Let a be a vertex of $\Gamma(\pi)$ with minimal degree, and let $\bar{\pi} \in S_{n-1}$ be the permutation obtained from π by deleting the value a (and decreasing by 1 all the values larger than a). Then

$$e(\Gamma(\bar{\pi})) \geq e(\Gamma(\pi) \setminus a),$$

where the latter is the number of edges in $\Gamma(\pi)$ which are not incident with the vertex a. By the induction hypothesis and Lemma 3.2,

$$\begin{aligned} e(\Gamma(\pi)) &= e(\Gamma(\pi) \setminus a) + d(a) \leq e(\Gamma(\bar{\pi})) + d(a) \\ &\leq \lfloor (n-1)^2/4 \rfloor + (n-1) - 2 + \lfloor n/2 \rfloor + 1 \\ &= \lfloor n^2/4 \rfloor + n - 2. \end{aligned}$$

Equality holds since, letting $m := \lfloor n/2 \rfloor$,

$$e(\Gamma([m+1, m+2, \ldots, n, 1, 2, \ldots, m])) = \lfloor n^2/4 \rfloor + n - 2.$$

$\square$

THEOREM 3.3.

$$\#\{\pi \in S_n \,|\, d(\pi) = \lfloor n^2/4 \rfloor + n - 2\} = \begin{cases} 2, & \text{if } n = 2; \\ 4, & \text{if } n = 3 \text{ or } n = 4; \\ 8, & \text{if } n \geq 6 \text{ is even;} \\ 16, & \text{if } n \geq 5 \text{ is odd.} \end{cases}$$

The extremal permutations have one of the following forms:

$$\pi_0 := [m+1, m+2, \ldots, n, 1, 2, \ldots, m] \qquad (m \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}),$$

and the permutations obtained from π_0 by one or more of the following operations:

$$\begin{aligned} \pi &\mapsto \pi^r := [\pi(n), \pi(n-1), \ldots, \pi(2), \pi(1)] \qquad \text{(reversing } \pi\text{)}, \\ \pi &\mapsto \pi^s := \pi \cdot t_{1,n} \qquad \text{(interchanging } \pi(1) \text{ and } \pi(n)\text{)}, \\ \pi &\mapsto \pi^t := t_{1,n} \cdot \pi \qquad \text{(interchanging 1 and } n \text{ in } \pi\text{)}. \end{aligned}$$

PROOF. It is not difficult to see that all the specified permutations are indeed extremal, and their number is as claimed (for all $n \geq 2$).

The claim that there are no other extremal permutations will be proved by induction on n. For small values of n (say $n \leq 4$) this may be verified directly. Assume now that the claim holds for some $n \geq 4$, and let $\pi \in S_{n+1}$ be extremal. Following the proof of Lemma 3.2, let a be a vertex of $\Gamma(\pi)$ with degree at most $\lfloor (n+1)/2 \rfloor + 1$, which is either $\pi(1)$ or its rightmost neighbor. As in the proof of Theorem 3.1, let $\bar{\pi} \in S_n$ be the permutation obtained from π by deleting the value a

(and decreasing by 1 all the values larger than a). All the inequalities in the proof of Theorem 3.1 must hold as equalities, namely: $e(\Gamma(\pi)\setminus a) = e(\Gamma(\bar\pi))$, $d(a) = \lfloor (n+1)/2 \rfloor + 1$, and $\bar\pi$ is extremal in S_n. By the induction hypothesis, $\bar\pi$ must have one of the prescribed forms. In all of them, $\{\bar\pi(1), \bar\pi(n)\} = \{m, m+1\}$ is an edge of $\Gamma(\bar\pi)$. Therefore the corresponding edge $\{\pi(1), \pi(n+1)\}$ (or $\{\pi(2), \pi(n+1)\}$ if $a = \pi(1)$, or $\{\pi(1), \pi(n)\}$ if $a = \pi(n+1)$) is an edge of $\Gamma(\pi)\setminus a$, namely of $\Gamma(\pi)$. If $a \ne \pi(1), \pi(n+1)$ then $\pi(n+1)$ is the rightmost neighbor of $\pi(1)$, contradicting the choice of a. If $a = \pi(n+1)$ we may use the operation $\pi \mapsto \pi^r$. Thus we may assume from now on that $a = \pi(1)$.

Let $N(a)$ denote the set of neighbors of a in $\Gamma(\pi)$. Assume first that

$$\bar\pi = \pi_0 = [m+1, m+2, \ldots, n, 1, 2, \ldots, m] \qquad (m \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}).$$

Noting that $\lceil n/2 \rceil = \lfloor (n+1)/2 \rfloor$ and keeping in mind the decrease in certain values during the transition $\pi \mapsto \bar\pi$, we have the following cases:

(1) $a > m+1$: in this case $1, \ldots, m \notin N(a)$, so that

$$d(a) \le n - m \le \lceil n/2 \rceil = \lfloor (n+1)/2 \rfloor < \lfloor (n+1)/2 \rfloor + 1.$$

Thus π is not extremal.

(2) $a < m$: in this case $m+3, \ldots, n+1, m+1 \notin N(a)$, so that

$$d(a) \le 1 + (m-1) \le \lceil n/2 \rceil < \lfloor (n+1)/2 \rfloor + 1.$$

Again, π is not extremal.

(3) $a \in \{m, m+1\}$: in this case

$$d(a) = 1 + m \le \lfloor (n+1)/2 \rfloor + 1,$$

with equality if and only if $m = \lfloor (n+1)/2 \rfloor$. This gives $\pi \in S_{n+1}$ of the required form (either π_0 or π_0^s).

A similar analysis for $\bar\pi = \pi_0^s$ gives extremal permutations only for $a \in \{m+1, m+2\}$ and $d(a) = 3$, so that $n = 4$ and $\bar\pi = [2413] \in S_4$. The permutations obtained are $\pi = [32514]$ and $\pi = [42513]$, which are $\pi_0^{rt}, \pi_0^{rst} \in S_5$, respectively.

The other possible values of $\bar\pi$ are obtained by the $\pi \mapsto \pi^r$ and $\pi \mapsto \pi^t$ operations from the ones above, and yield analogous results.

□

4. Expectation

Theorem 4.1. *For every positive integer n, the expected down degree of a random permutation in S_n is*

$$E_{\pi\in S_n} d_-(\pi) = \sum_{i=2}^{n} \sum_{j=2}^{i} \sum_{k=2}^{j} \frac{1}{i\cdot(k-1)} = (n+1)\sum_{i=1}^{n} \frac{1}{i} - 2n.$$

For a proof see [**1**].

Corollary 4.2. *As $n \to \infty$,*

$$E_{\pi\in S_n} d_-(\pi) = n \ln n + O(n)$$

and

$$E_{\pi\in S_n} d(\pi) = 2n \ln n + O(n).$$

5. Generalized Down Degrees

Definition 5.1. *For $\pi \in S_n$ and $1 \leq r < n$ let*

$$D_-^{(r)}(\pi) := \{t_{a,b} | \ \ell(\pi) > \ell(t_{a,b}\pi) > \ell(\pi) - 2r\}$$

the r-th strong descent set *of* π*.*

Define the r-th down degree *as*

$$d_-^{(r)}(\pi) := \# D_-^{(r)}(\pi).$$

Example 5.2. *The first strong descent set and down degree are those studied in the previous section; namely, $D_-^{(1)}(\pi) = D_-(\pi)$ and $d_-^{(1)}(\pi) = d_-(\pi)$.*

The $(n-1)$-st strong descent set is the set of inversions*:*

$$D_-^{(n-1)}(\pi) = \{t_{a,b} \,|\, a < b,\, \pi^{-1}(a) > \pi^{-1}(b)\}.$$

Thus

$$d_-^{(n-1)}(\pi) = inv(\pi),$$

the inversion number of π*.*

Observation 5.3. *For every $\pi \in S_n$ and $1 \leq a < b \leq n$, $t_{a,b} \in D_-^{(r)}(\pi)$ if and only if $\pi = [\ldots, b, \ldots, a, \ldots]$ and there are less than r letters between the positions of b and a in π whose value is between a and b.*

Definition 5.4. *The* r-th strong descent graph *of $\pi \in S_n$, $\Gamma_r(\pi)$ is the graph whose set of vertices is $\{1, \ldots, n\}$ and whose set of edges is*

$$\{\{a, b\} | \ t_{a,b} \in D_-^{(r)}(\pi)\}.$$

The following lemma generalizes Lemma 2.7.

Lemma 5.5. *For every $\pi \in S_n$, the graph $\Gamma_r(\pi)$ contain no subgraph isomorphic to the complete graph K_{r+2}.*

For a proof see [**1**].

Corollary 5.6. *For every $1 \leq r < n$,*

$$max\{d_-^{(r)}(\pi) | \ \pi \in S_n\} \leq t_{r+1}(n) \leq \binom{r+1}{2} \left(\frac{n}{r+1}\right)^2.$$

Proof. Combining Turán's Theorem together with Lemma 5.5.

□

Acknowledgements. The concept of strong descent graph came up during conversations with Francesco Brenti. Its name and certain other improvements were suggested by Christian Krattenthaler. Thanks also to Nathan Reading, Amitai Regev, Alexander Yong, and an anonymous referee.

References

[1] R. M. Adin and Y. Roichman, *On Degrees in the Hasse Diagram of the Strong Bruhat Order*, arXiv:math.CO/0505020.

[2] M. Aguiar and F. Sottile, *Structure of the Malvenuto-Reutenauer Hopf algebra of permutations*, Adv. Math. 191 (2005), 225–275.

[3] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics 231, Springer, New York, 2005.

[4] B. Bollobás, *Modern Graph Theory*, Graduate Texts in Mathematics 184, Springer, New York, 1998.

[5] F. Brenti, *Upper and lower bounds for Kazhdan-Lusztig polynomials*, Europ. J. Combin. 19 (1998), 283–297.

[6] L. Comtet, *Sur les coefficients del l'inverse de la série formelle* $\sum n!t^n$, Compt. Rend. Acad. Sci. Paris A-B 275 (1972), A569–A572.
[7] I. M. Gessel and R. P. Stanley, *Algebraic Enumeration*, in: Handbook of Combinatorics, Vol. 2, Eds. R. Graham et al., M.I.T. Press, 1995.
[8] N. Reading, *Order dimension, strong Bruhat order and lattice properties for posets*, Order 19 (2002), 73–100.
[9] A. Regev and Y. Roichman, *Generalized statistics on S_n and pattern avoidance*, Europ. J. Combin. 26 (2005), 29–57.
[10] R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, Wadsworth and Brooks/Cole, Monterey, CA, 1986.
[11] P. Turán, *An extremal problem in graph theory* (Hungarian), Mat. Fiz. Lapok 48 (1941), 436–452.
[12] A. Woo and A. Yong, *When is a Schubert variety Gorenstein?*, arXiv:math.AG/0409490.

Department of Mathematics, Bar-Ilan University, Ramat-Gan 52900, Israel
E-mail address: `radin@math.biu.ac.il`

Department of Mathematics, Bar-Ilan University, Ramat-Gan 52900, Israel
E-mail address: `yuvalr@math.biu.ac.il`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

On bar partitions and spin character zeros

Christine Bessenrodt

Abstract. The main combinatorial result in this article is a classification of bar partitions of n which are of maximal p-bar weight for all odd primes $p \leq n$. As a consequence, we show that apart from very few exceptions any irreducible spin character of the double covers of the symmetric and alternating groups vanishes on some element of odd prime order.

Résumé. Notre résultat principal est une classification des partages barrés de n qui ont un poids p-barré maximal pour tous les nombres prémiers p impairs inférieur à n. Comme conséquence, on a que, à quelques exceptions près, tout caractère spin irréductible d'une couverture double des groupes symètriques et groupes alternants s'annule sur un élément d'ordre premier.

1. Introduction

A well known result by Burnside states that any non-linear irreducible character of a finite group vanishes on some element of the group. This was refined in [**9**], where it was shown that such a character always has a zero at an element of prime power order; it had also been noticed in [**9**] that any non-linear irreducible character of a finite simple group except possibly the alternating groups even vanishes on some element of prime order. This was complemented in [**5**] where it was shown that this character property also holds for the symmetric and the alternating groups. Indeed, this vanishing property was a consequence of a combinatorial result on the weights of partitions.

Here, we deal with the corresponding result on bar weights of partitions into distinct parts (which we call bar partitions). This then yields a vanishing property for irreducible spin characters of the double covers of the symmetric and alternating groups on elements of odd prime order.

In the next section we collect together some combinatorial preliminaries; we then briefly recall the results from [**5**] in the case of partitions and ordinary characters of the symmetric and alternating groups. In Section 4 we discuss the case of bar partitions and spin characters of the double cover groups; in the main result, Theorem 4.1, the bar partitions of n are classified which are of maximal p-bar weight for all primes $p \leq n$. These then give rise to the desired spin character zeros; see Theorem 4.2.

A detailed paper with full proofs will appear elsewhere.

2. Preliminaries

We refer to [**8**], [**12**], [**7**] for details about partitions, Young diagrams, hooks and bar partitions, shifted diagrams and bars, respectively.

Consider a partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ of the integer n. Thus $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_l > 0$ and $\lambda_1 + \lambda_2 + \ldots + \lambda_l = n$, with integer parts λ_i; $l = l(\lambda)$ is the length of λ. The Young diagram of λ consists of n boxes with λ_i boxes in the ith row. We refer to the boxes in matrix notation, i.e. the (i, j)-box is the

---

2000 *Mathematics Subject Classification.* Primary 05A17; Secondary 20C30.

Key words and phrases. partitions, bar partitions, symmetric groups and their double covers, spin characters, vanishing property .

The author thanks the Isaac Newton Institute for Mathematical Sciences for its hospitality during a stay there in the frame of the programme *Symmetric functions and Macdonald polynomials*, where an early part of the work for this article was done. Thanks go also to Peking University for its hospitality in the final phase of the work.

jth box in the ith row. The (i,j)-hook consists of the boxes in the Young diagram to the right of and below the (i,j)-box, and including this box. The number of boxes in this hook is its hook length, denoted by h_{ij}.

For $n \in \mathbb{N}$, we denote by $D(n)$ the set of partitions of n into distinct parts, and we set $D = \bigcup_n D(n)$. We call the partitions in D also bar partitions. A partition $\lambda \in D(n)$ is in $D^+(n)$ (or $D^-(n)$, respectively) if $n - l(\lambda)$ is even (or odd, respectively).
We denote by $\mathcal{O}(n)$ the set of partitions of n into odd parts; elements of the double cover groups $\tilde{S}_n$ which correspond to elements of S_n of cycle type $\alpha \in \mathcal{O}$ are said to be of type $\mathcal{O}$.
For $\lambda \in D$, we consider the corresponding shifted diagram, where in the ith row we start on the diagonal at (i,i) rather than at the box $(i,1)$. By flipping over the diagonal we obtain the shift symmetric diagram $S(\lambda)$. The bar lengths in λ correspond to the hook lengths in the λ-boxes of $S(\lambda)$; the bar length at position (i,j) is then denoted b_{ij}; we abbreviate the bar lengths in the first row by $b_{1i} = b_i$.

Example. Take $\lambda = (4,3,1)$. In the shift symmetric diagram below the bar lengths are filled into the corresponding boxes of λ.

$$\begin{array}{ccccc} . & 7 & 5 & 4 & 2 \\ . & . & 4 & 3 & 1 \\ . & . & . & 1 & \\ . & . & & & \end{array}$$

The removal of a p-bar from $\lambda \in D(n)$ corresponds to taking a part p or two parts summing to p out of λ, or subtracting p from a part of λ if possible (i.e., if the resulting partition is in $D(n-p)$). Doing this as long as possible gives the $\bar{p}$-core $\lambda_{(\bar{p})}$ of λ; the number of p-bars removed is then the p-bar weight $\bar{w}_p(\lambda)$ of λ (see [**7**] or [**12**] for details). These operations may also be performed on a suitable $\bar{p}$-abacus.

Example. Take $p = 3$, $\lambda = (7,3,2,1)$. Removing a bar of length 3 from λ can be achieved by removing the parts 2 and 1 from λ, or by removing the part 3, or by replacing 7 by 4. When we do this in succession, we have reached the bar partition (4), from which we can remove a further 3-bar and thus obtain $(1) = \lambda_{(\bar{3})}$; the $\bar{3}$-weight of λ is thus 4.

We will often make use of the following property of the p-bar weight of a partition (see [**11**], [**12**]); the Lemma may easily be proved by considering the $\bar{p}$-abacus (see [**12**]).

Lemma 2.1. *Let p be an odd prime. If λ is a bar partition of $\bar{p}$-weight $\bar{w}_p(\lambda) = w$, then λ has exactly w bars of length divisible by p. In particular, if λ has a bar of length divisible by p, then it has a bar of length p.*

This is used to prove some easy but crucial results about bar lengths (compare this with [**4**] where a similar Lemma for hook lengths is used).

For $p = 2$, a suitable parameter to consider is the $\bar{4}$-core of λ which is computed using the $\bar{4}$-abacus with one runner for the even parts, and two conjugate runners for the parts $\equiv 1,3 \mod 4$; in contrast to the $\bar{p}$-abacus for odd p, here we are allowed to subtract 2 from the even parts (so these will be removed when computing the $\bar{4}$-core).

3. Partitions and ordinary characters of S_n and A_n

Before stating the new results on bar partitions and spin characters in the next section, we recall here the recent results from [**5**]. Towards the refinement of Burnside's Theorem for S_n and A_n the following main combinatorial result was proved there:

Theorem 3.1. [**5**] *Let λ be a partition of $n \in \mathbb{N}$. Then the following holds:*
(i) *λ is of maximal p-weight for all primes $p \le n$, if and only if one of the following occurs:*

$$\lambda = (n)\,, (1^n) \text{ or } (2^2)\,.$$

(ii) *λ is of maximal p-weight for all odd primes $p \leq n$, if and only if λ is one of the partitions in* (i)*, one of* $(n-1,1)$*,* $(2,1^{n-2})$*, where $n = 2^a + 1$ for some $a \in \mathbb{N}$, or one of the following occurs:*

$$\begin{array}{rrl} n = & 6: & \lambda = (3,2,1) \\ n = & 8: & \lambda = (5,2,1) \ \textit{or} \ (3,2,1^3) \\ n = & 9: & \lambda = (6,3) \ \textit{or} \ (2^3,1^3) \\ n = & 10: & \lambda = (4,3,2,1) \end{array}$$

This has the desired consequence:

Theorem 3.2. [**5**] *Let $n \in \mathbb{N}$. Let χ be any non-linear irreducible character of the symmetric group S_n or the alternating group A_n. Then χ vanishes on some element of prime order. If $\chi(1)$ is not a 2-power, then χ is zero on some element of odd prime order.*

Theorem 3.1 also has a consequence for the distribution into p-blocks; this was recently taken up in more detail in [**2**].

We refer to [**8**, section 2.5] for the labelling of the irreducible characters of A_n. A simple relation between the p-weight of a partition λ and the defect of the p-block containing the irreducible character labelled by λ is given in [**8**, 6.2.45]. The principal p-block of a finite group is the block containing the trivial character.

Theorem 3.3. [**5**] (i) *The characters $[n], [1^n]$ and $[2^2]$ are the only irreducible characters of S_n which are in p-blocks of maximal defect for all primes p.*
Apart from $[1^2], [1^3], [1^4], [1^6], [2^2]$, the trivial character of S_n is the only irreducible character which is in the principal p-block for all primes $p \leq n$.
(ii) *The characters $\{n\}$, $\{2,1\}_\pm$ and $\{2^2\}_\pm$ are the only irreducible characters of A_n which are in p-blocks of maximal defect for all primes p.*
They belong to the principal p-block for all primes $p \leq n$, except for the characters $\{2,1\}_\pm$ at $p = 2$.

We will see that our main result on bar partitions is of a similar type as Theorem 3.1 above, and it has similar consequences for character zeros of spin characters and for the distribution of spin characters into spin p-blocks, for odd primes p.

4. Bar partitions and spin characters

In our main result we present a classification of the bar partitions of n which have maximal $\bar{p}$-weight $\left\lfloor \frac{n}{p} \right\rfloor$ for all odd primes $p \leq n$; equivalently, the $\bar{p}$-core of these bar partitions is small in the sense that it is of size smaller than p. (Here $\lfloor \cdot \rfloor$ denotes the floor function. Thus $\lfloor x \rfloor$ is the integral part of $x \in \mathbb{R}$.) For $p = 2$, we consider the case where the $\bar{4}$-core is small, i.e., of size smaller than 4.

The elements of odd prime order p which we are then going to use for the vanishing property for spin characters of the double cover $\tilde{S}_n$ of the symmetric group S_n are those where the corresponding cycle type is of maximal p-bar weight, i.e., the cycle type has $\left\lfloor \frac{n}{p} \right\rfloor$ parts of size p. Indeed, the connection to the vanishing of spin character values is easily explained. The irreducible spin characters of $\tilde{S}_n$ are labelled by the bar partitions λ of n (and signs). The recursion formula given by Morris [**10**] for spin character values on elements of type $\mathcal{O}$ in $\tilde{S}_n$ shows that the irreducible spin character(s) labelled by λ vanishes on a p-element of maximal weight (where p is odd), if the $\bar{p}$-weight of λ is *not* maximal.

Our main result on bar partitions is the following:

Theorem 4.1. *Let λ be a bar partition of $n \in \mathbb{N}$. Then λ is of maximal $\bar{p}$-weight for all odd primes $p \leq n$, if and only if $\lambda = (n)$ or $\lambda = (n-1,1)$, where $n = 2^a + 2$ for some $a \in \mathbb{N}$, or one of the following occurs:*

$$\begin{array}{rrl} n = & 5: & \lambda = (3,2) \\ n = & 6: & \lambda = (3,2,1) \\ n = & 8: & \lambda = (5,2,1) \\ n = & 9: & \lambda = (4,3,2) \\ n = & 10: & \lambda = (4,3,2,1) \ \textit{or} \ (7,3) \end{array} .$$

If, in addition, also $\lambda_{(\bar{4})}$ is small, then $\lambda = (n)$ or λ is one of $(3,1)$*,* $(3,2,1)$*,* $(4,3,2,1)$*.*

The combinatorial classification result immediately has the desired consequence for the spin character zeros, as explained above; first we have to introduce some more notation (see [**7**], [**10**], [**12**]).

We denote by $\langle\mu\rangle$ the irreducible spin character of $\tilde{S}_n$ corresponding to $\mu \in D^+(n)$, and by $\langle\mu\rangle_+$, $\langle\mu\rangle_- = \mathrm{sgn}\cdot\langle\mu\rangle_+$, the irreducible spin characters of $\tilde{S}_n$ associated to $\mu \in D^-(n)$.
Furthermore, we let $\langle\langle\mu\rangle\rangle$ denote the irreducible spin character of $\tilde{A}_n$ corresponding to $\mu \in D^-(n)$ (which is the reduction of $\langle\mu\rangle_\pm$), and $\langle\langle\mu\rangle\rangle_\pm$ the irreducible spin characters of $\tilde{A}_n$ associated to $\mu \in D^+(n)$ (which are conjugate and sum to the reduction of $\langle\mu\rangle$, and which differ only on classes of cycle type $\lambda \in D^+$).

We refer to [**7**] for further details on the irreducible spin characters of $\tilde{A}_n$.

Let $n \in \mathbb{N}$, $n \geq 4$. First we observe that any irreducible spin character of a double cover $\tilde{S}_n$ of the symmetric group or a double cover $\tilde{A}_n$ of the alternating group has a zero of order 2. For this, note that the cycle types (2^a1^b), with $a > 0$, are neither of type $\mathcal{O}$ nor of type D and hence these classes do not split in the double cover groups. Thus all spin characters are zero on these classes. Hence in the following Theorem we are only interested in classes of odd prime order.

Theorem 4.2. *Let $n \in \mathbb{N}$, $n \geq 4$. Let χ be any irreducible spin character of a double cover of the symmetric group $\tilde{S}_n$ or the alternating group $\tilde{A}_n$. Then χ vanishes on some element of odd prime order, except if χ is a basic spin character, i.e., labelled by (n), or in the cases where χ is labelled by $(n-1,1)$ with $n = 2^a + 2$ for some $a \in \mathbb{N}$, or by one of the partitions $(3,2)$, $(3,2,1)$ or $(5,2,1)$.*

Remark 4.3. If an irreducible character χ of a finite group G has a zero at an element of prime order p, then p divides $\chi(1)$. Note that the irreducible spin characters of $\tilde{S}_n$ and $\tilde{A}_n$ of prime power degree have been classified in [**1**]; from Theorem 4.2 we can immediately recover the classification of irreducible spin characters of 2-power degree for these groups. In fact, here they are exactly those that do not have a zero at an element of odd prime order.
The converse of the statement above does not hold, even for $G = \tilde{S}_n$. The spin character $\langle 8,4\rangle$ is of degree $5280 = 2^4\cdot 3\cdot 5\cdot 11$, but the character does not vanish on any element of order 3.

Note that for $p > 2$ there is a simple relation between the $\bar{p}$-weight of a bar partition λ and the defect of the p-spin block containing the irreducible spin character(s) of S_n or A_n labelled by λ (see [**12**]). For $2 < p \leq n$, the basic spin character(s) of $\tilde{S}_n$ or $\tilde{A}_n$ are contained in one spin p-block which we call the *basic spin p-block* of $\tilde{S}_n$ or $\tilde{A}_n$, respectively. The following is then another direct consequence of Theorem 4.1 (note that for $a > 2$ the spin character to $(2^a+1,1)$ is not in the basic spin p-block for any odd prime p not dividing n and $n-1$).

Theorem 4.4. *Let $n \in \mathbb{N}$, $n \geq 4$.*
(i) *The basic spin characters $\langle n\rangle_{(\pm)}$, the spin characters $\langle n-1,1\rangle_{(\pm)}$ where $n = 2^a+2$ for some $a \in \mathbb{N}$, and the spin characters $\langle 3,2\rangle_\pm, \langle 3,2,1\rangle_\pm, \langle 5,2,1\rangle_\pm, \langle 4,3,2\rangle, \langle 4,3,2,1\rangle, \langle 7,3\rangle$ are the only irreducible spin characters of $\tilde{S}_n$ which are in spin p-blocks of maximal defect for all odd primes p.*
The spin characters $\langle 3,1\rangle$, $\langle 5,1\rangle$, $\langle 3,2\rangle_\pm$, $\langle 3,2,1\rangle_\pm$, $\langle 4,3,2\rangle$, $\langle 7,3\rangle$ are the only non-basic spin characters contained in the basic spin p-block for all odd primes $p \leq n$.
(ii) *The basic spin characters $\langle\langle n\rangle\rangle_{(\pm)}$, the spin characters $\langle\langle n-1,1\rangle\rangle_{(\pm)}$ where $n = 2^a+2$ for some $a \in \mathbb{N}$, and the spin characters $\langle\langle 3,2\rangle\rangle$, $\langle\langle 3,2,1\rangle\rangle$, $\langle\langle 5,2,1\rangle\rangle$, $\langle\langle 4,3,2\rangle\rangle_\pm$, $\langle\langle 4,3,2,1\rangle\rangle_\pm$, $\langle\langle 7,3\rangle\rangle_\pm$ are the only irreducible spin characters of $\tilde{A}_n$ which are in spin p-blocks of maximal defect for all odd primes p.*
The spin characters $\langle\langle 3,1\rangle\rangle_\pm, \langle\langle 5,1\rangle\rangle_\pm, \langle\langle 3,2\rangle\rangle, \langle\langle 3,2,1\rangle\rangle, \langle\langle 4,3,2\rangle\rangle_\pm, \langle\langle 7,3\rangle\rangle_\pm$ are the only non-basic spin characters contained in the basic spin p-block of $\tilde{A}_n$ for all odd primes $p \leq n$.

For $p = 2$, the blocks contain both ordinary and spin characters; in fact, the 2-block distribution of spin characters is more intricate and has been determined in [**3**]. Here the $\bar{4}$-combinatorics mentioned before fits with the distribution of spin characters into the 2-blocks of $\tilde{S}_n$ (see [**3**]). We note that when $n \equiv 3 \mod 4$, the basic spin character is not contained in the principal 2-block. Using also the 2-blocks, the non-basic spin characters may be even more finely separated from the basic spin characters; one easily checks that only the spin characters $\langle 3,1\rangle$ and $\langle 3,2,1\rangle_\pm$ are in the same p-block as the basic spin characters for all primes $p \leq n$ (analogously for $\tilde{A}_n$).

Now we want to indicate the strategy of the proof of the main classification result. We start by studying the bar lengths in bar partitions. We write $\bar{h}_\mu$ for the product of all the bar lengths of a bar partition μ.

From now on, $\lambda = (\lambda_1, \ldots, \lambda_l)$ is always a bar partition of n, of length l. The following easy result is very useful:

Proposition 4.5. *Assume that $\bar{w}_p(\lambda) = \left\lfloor \frac{n}{p} \right\rfloor$ for the odd prime $p \le n$.*

(i) *Let μ be obtained from λ by removing the first row. If p does not divide $\bar{h}_\mu$, then $p,\ 2p,\ \ldots,\ \left\lfloor \frac{n}{p} \right\rfloor p$ are first row bar lengths of λ.*
(ii) *If $n - \lambda_1 < p$, then $p,\ 2p,\ \ldots,\ \left\lfloor \frac{n}{p} \right\rfloor p$ are first row bar lengths of λ.*

Note that the first row bar lengths of λ, denoted $b_1, \ldots, b_{\lambda_1}$, can explicitly be given; the set of these numbers is

$$\{\lambda_1 + \lambda_2, \ldots, \lambda_1 + \lambda_l\} \cup \{1, \ldots, \lambda_1\} \setminus \{\lambda_1 - \lambda_2, \ldots \lambda_1 - \lambda_l\}\,.$$

In particular, the largest bar length in λ is $\lambda_1 + \lambda_2$.

As for the study of hook lengths of partitions, some number theoretic results about the distribution of primes are needed. In particular, a result due to Hanson is very useful; the exceptions occurring here are also a reason for exceptions occurring for small n in the classification theorem.

Theorem 4.6. [**6**] *The product of k consecutive numbers all greater than k contains a prime divisor greater than $\frac{3}{2}k$, with the only exceptions $3 \cdot 4$, $8 \cdot 9$ and $6 \cdot 7 \cdot 8 \cdot 9 \cdot 10$.*

In the case of partitions, we first dealt with the case of hooks in [**5**]. Here, one treats the "bar case" first, i.e., partitions of length at most 2.

Proposition 4.7. *Let $\lambda = (n-k, k)$ for some $k \in \mathbb{N}_0$, $k < n-k$. Then $\bar{w}_p(\lambda) = \left\lfloor \frac{n}{p} \right\rfloor$ for all odd primes $p \le n$ if and only if one of the following holds:*

(i) *$k = 0$, i.e., $\lambda = (n)$.*
(ii) *$k = 1$ and $n = 2^a + 2$ for some $a \in \mathbb{N}_0$, i.e., $\lambda = (2^a + 1, 1)$.*
(iii) *λ is one of $(3,2)$, $(7,3)$.*

If, in addition, also the $\bar{4}$-core is small, then $\lambda = (n)$ or λ is one of $(2,1), (3,1)$.

The following observation is crucial for getting a reduction procedure started in the general case.

Lemma 4.8. *Let $\lambda \in D(n)$. Let s be a bar length of λ with $\frac{n}{2} \le s$. Then s is a first row bar length of λ or $s = b_{23} = \lambda_2 + \lambda_3$. In the second case, b_1, b_2 are then the only first row bar lengths $\ge \frac{n}{2}$.*

Corollary 4.9. *Let $n = 13, 14$ or $n \ge 17$. Let $\lambda \in D(n)$ be of maximal $\bar{p}$-weight for all odd primes p with $\frac{n}{2} \le p \le n$. Then all bar lengths $\ge \frac{n}{2}$ are first row bar lengths of λ.*

Based on the following result we can then use the same algorithm as in [**4**]:

Proposition 4.10. *Let $\lambda \in D(n)$, $n \ge 17$, which is of maximal $\bar{p}$-weight for all odd primes $p \le n$. Let $s_1 < s_2 < \cdots < s_r \le n$ and $t_1 < t_2 < \cdots < t_r \le n$ be sequences of integers satisfying*

(i) *$s_i < t_i$ for all i;*
(ii) *s_1, t_1 are primes $> \frac{n}{2}$;*
(iii) *for $1 \le i \le r-1$, s_{i+1}, t_{i+1} have prime divisors exceeding $2n - s_i - t_i$.*

Then $s_1, \ldots, s_r$, $t_1, \ldots, t_r$ are first row bar lengths of λ.

It was already checked for the proof of the classification result in [**1**] that a suitable algorithm producing sequences as occurring in the proposition above ends close to n; also, the Theorem is easily checked for small n. We then obtain the following consequence:

Corollary 4.11. *Let $n \in \mathbb{N}$. Let λ be a bar partition of n of maximal $\bar{p}$-weight for all odd primes $p \le n$, $b_1 = \lambda_1 + \lambda_2$ its largest bar length.*

(i) *For $n \le 9.25 \cdot 10^8$, $n - b_1 \le 4$.*
(ii) *For $n > 9.25 \cdot 10^8$, $n - b_1 \le 225$.*

After this, we still have the tasks to reduce 225 to some manageable number, and to deal with the cases where $n - b_1$ is small. For this, we use a tailor-made number-theoretic Lemma for reducing $d = n - b_1$ and $k = \lambda_2 - \lambda_3 - 1$; it refines Hanson's Theorem in special situations.

Lemma 4.12. *Let $5 \leq m \leq 1000$. Then any product of m consecutive integers larger than $5.5 \cdot 10^8$ has a prime divisor $q > 2.15 \cdot m$, when $m \leq 10$, $q > 2.58 \cdot m$, when $11 \leq m \leq 21$, and $q > 3 \cdot m$, when $m \geq 22$.*

This Lemma also helps to deal with the cases of medium-sized d and k. The cases of small d and k are dealt with in a tedious case-by-case analysis; here the further exceptions for small n stated in the Theorem arise. This then finishes the proof.

References

[1] A. Balog, C. Bessenrodt, J. B. Olsson, K. Ono, Prime power degree representations of the symmetric and alternating groups, J. London Math. Soc. (2) 64 (2001) 344-356

[2] C. Bessenrodt, G. Malle, J. B. Olsson, Separating characters by blocks, to appear in: J. London Math. Soc.

[3] C. Bessenrodt, J. B. Olsson, The 2-blocks of the covering groups of the symmetric groups, Advances in Math. 129 (1997), 261 - 300

[4] C. Bessenrodt, J. B. Olsson, Prime power degree representations of the double covers of the symmetric and alternating groups, J. London Math. Soc. (2) 66 (2002) 313 - 324

[5] C. Bessenrodt, J. B. Olsson, Weights of partitions and character zeros, Electron. J. Comb. 11(2) (2004), R5

[6] D. Hanson, On a theorem of Sylvester and Schur, Canad. Math. Bull. 16 (1973) 195-199

[7] P. N. Hoffman, J. F. Humphreys, *Projective representations of the symmetric groups*, Clarendon Press, Oxford 1992

[8] G. James, A. Kerber, *The representation theory of the symmetric group*, Encyclopedia of Mathematics and its Applications, 16, Addison-Wesley 1981

[9] G. Malle, G. Navarro, J. B. Olsson, Zeros of characters of finite groups, J. Group Theory 3 (2000) 353-368

[10] A. O. Morris, The spin representation of the symmetric group, Canad. J. Math. **17** (1965), 543-549

[11] A. O. Morris, A. K. Yaseen, Some combinatorial results involving Young diagrams, *Math. Proc. Camb. Phil. Soc.* **99** (1986) 23-31

[12] J. B. Olsson, *Combinatorics and representations of finite groups*, Vorlesungen aus dem FB Mathematik der Univ. Essen, Heft 20, 1993

Fakultät für Mathematik und Physik, Universität Hannover, Welfengarten 1, D-30161 Hannover
E-mail address: bessen@math.uni-hannover.de

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A Rook Theory Model for the Generalized p,q-Stirling Numbers of the First and Second Kind

Karen Sue Briggs

Abstract. In (EJC 11 (2004), #R84), Remmel and Wachs presented two natural ways to define p,q-analogues of the generalized Stirling numbers of the first and second kind, $S^1(\alpha,\beta,r)$ and $S^2(\alpha,\beta,r)$ as introduced by Hsu and Shiue (Adv. App. Math 20 (1998), 366-384). In this paper, we present a rook theoretic model for each type of p,q-analogue based on a pair of boards parametrized by the nonnegative integers α, β, and r, so that rooks attack cells on its own board as well as on its companion board. For each model, we provide an analogue of Goldman, Joichi and White's product formula (Proc. Amer. Math. Soc. 52 (1975), 485-492) and demonstrate how each type of the generalized p,q-Stirling numbers of the first and second kind arises as a special case of these p,q-rook numbers.

Résumé. Remmel et Wachs, dans (EJC 11 (2004), #R84), ont présenté deux façons naturelles pour définir les p,q-analogues des nombres de Stirling généralisés, des première et deuxième sortes, $S^1(\alpha,\beta,r)$ et $S^2(\alpha,\beta,r)$, introduits par Hsu et Shiue (Adv. App. Math 20 (1998), 366-384). Dans cet article, nous présentons un model théorique des mouvements de la tour pour chaque type des p,q-analogues basé sur une paire de jeux paramétrisés par les entiers non-négatifs α, β, et r. Ainsi, la tour attaque les cases sur son propre jeu et celles de l'autre jeu. Pour chacun des modèles, nous donnons une formule analogue à celle du produit de Goldman, Joichi et White (Proc. Amer. Math. Soc. 52 (1975), 485-492) et démontrons comment chaque type de p,q-analogues des nombres de Stirling généralisés des première et deuxième sortes forment un cas spécial de nombres p,q-analogues pour les mouvements de la tour.

1. Introduction

In [**11**], Remmel and Wachs presented two natural ways to give p,q-analogues of Hsu and Shiue's generalized Stirling numbers of the first and second kind [**7**], respectively denoted $\overline{S}^1_{n,k}(\alpha,\beta,r)$ and $\overline{S}^2_{n,k}(\alpha,\beta,r)$ for $0 \le k \le n$, and defined by

$$x(x-\alpha)\cdots(x-(n-1)\alpha) = \sum_{k=0}^{n} \overline{S}^1_{n,k}(\alpha,\beta,r)(x-r)(x-r-\beta)\cdots(x-r-(k-1)\beta), \tag{1.1}$$

and

$$x(x-\beta)\cdots(x-(n-1)\beta) = \sum_{k=0}^{n} \overline{S}^2_{n,k}(\alpha,\beta,r)(x+r)(x+r-\alpha)\cdots(x+r-(k-1)\alpha). \tag{1.2}$$

From these definitions, one can clearly see that $\overline{S}^1_{n,k}(\alpha,\beta,r) = \overline{S}^2_{n,k}(\beta,\alpha,-r)$. Moreover, we find that $\overline{S}^1_{n,k}(1,0,0) = s_{n,k}$ and $\overline{S}^2_{n,k}(1,0,0) = S_{n,k}$ where $s_{n,k}$ and $S_{n,k}$ respectively denote the classical Stirling numbers of the first and second kind.

By setting

$$S^1_{n,k}(\alpha,\beta,r) = \overline{S}^1_{n,k}(\alpha,\beta,-r) \text{ and } S^2_{n,k}(\alpha,\beta,r) = \overline{S}^2_{n,k}(\alpha,\beta,-r)$$

2000 *Mathematics Subject Classification.* Primary 05A10, 05A15, 05A30; Secondary 05E15.

Key words and phrases. generalized stirling numbers, p,q-analogues, rook numbers, bipartite boards.

This research was completed as part of a larger project on rook theory with Jeffrey B. Remmel, Department of Mathematics, University of California, San Diego.

and replacing x by $t-r$ in equation (1.1) and x by t in equation (1.2), Remmel and Wachs obtained the following pair of equations:

$$(1.3)\qquad (t-r)(t-r-\alpha)\cdots(t-r-(n-1)\alpha)=\sum_{k=0}^{n} S^1_{n,k}(\alpha,\beta,r)t(t-\beta)\cdots(t-(k-1)\beta),$$

and

$$(1.4)\qquad t(t-\beta)\cdots(t-(n-1)\beta)=\sum_{k=0}^{n} S^2_{n,k}(\alpha,\beta,r)(t-r)(t-r-\alpha)\cdots(t-r-(k-1)\alpha).$$

Replacing $(t-\gamma)$ by two distinctly natural p,q-analogues, Remmel and Wachs then defined their two types of p,q-analogues of $S^1_{n,k}(\alpha,\beta,r)$ and $S^2_{n,k}(\alpha,\beta,r)$. The p,q-analogue of any real number γ is defined by

$$[\gamma]_{p,q}=\frac{p^\gamma-q^\gamma}{p-q},$$

so that when $\gamma=n$ is a nonnegative integer, $[n]_{p,q}=q^{n-1}+pq^{n-2}+\cdots p^{n-2}q+p^{n-1}$. Then, the p,q-analogues of $n!$ and $\binom{n}{k}$ are naturally defined by $[n]_{p,q}!=[n]_{p,q}[n-1]_{p,q}\cdots[1]_{p,q}$ and

$$\left[\begin{array}{c} n \\ k \end{array}\right]_{p,q}=\frac{[n]_{p,q}!}{[k]_{p,q}![n-k]_{p,q}!}.$$

For their type-I (p,q)-analogues of $S^1_{n,k}(\alpha,\beta,r)$ and $S^2_{n,k}(\alpha,\beta,r)$, Remmel and Wachs replaced $(t-\gamma)$ by $([t]_{p,q}-[\gamma]_{p,q})$ in (1.3) and (1.4). That is, they defined $S^{1,p,q}_{n,k}(\alpha,\beta,r)$ and $S^{2,p,q}_{n,k}(\alpha,\beta,r)$ for $0\le k\le n$ respectively by the following equations:

$$(1.5)\qquad \begin{aligned}&([t]_{p,q}-[r]_{p,q})([t]_{p,q}-[r+\alpha]_{p,q})\cdots([t]_{p,q}-[r+(n-1)\alpha]_{p,q})\\ &=\sum_{k=0}^{n} S^{1,p,q}_{n,k}(\alpha,\beta,r)([t]_{p,q})([t]_{p,q}-[\beta]_{p,q})\cdots([t]_{p,q}-[(k-1)\beta]_{p,q})\end{aligned}$$

and

$$(1.6)\qquad \begin{aligned}&([t]_{p,q}-[\beta]_{p,q})\cdots([t]_{p,q}-[(n-1)\beta]_{p,q})\\ &=\sum_{k=0}^{n} S^{2,p,q}_{n,k}(\alpha,\beta,r)([t]_{p,q}-[r]_{p,q})([t]_{p,q}-[r+\alpha]_{p,q})\cdots([t]_{p,q}-[r+(k-1)\alpha]_{p,q}).\end{aligned}$$

Moreover, they proved that when $0\le k\le n$, the $S^{1,p,q}_{n,k}(\alpha,\beta,r)$ and $S^{2,p,q}_{n,k}(\alpha,\beta,r)$ defined according to equations (1.5) and (1.6) satisfy the following recursions:

$$(1.7)\qquad S^{1,p,q}_{0,0}(\alpha,\beta,r)=1 \text{ and } S^{1,p,q}_{n,k}(\alpha,\beta,r)=0 \text{ if } k<0 \text{ or } k>n$$

and

$$(1.8)\qquad S^{1,p,q}_{n+1,k}(\alpha,\beta,r)=S^{1,p,q}_{n,k-1}(\alpha,\beta,r)+([k\beta]_{p,q}-[n\alpha+r]_{p,q})S^{1,p,q}_{n,k}(\alpha,\beta,r),$$

$$(1.9)\qquad S^{2,p,q}_{0,0}(\alpha,\beta,r)=1 \text{ and } S^{2,p,q}_{n,k}(\alpha,\beta,r)=0 \text{ if } k<0 \text{ or } k>n$$

and

$$(1.10)\qquad S^{2,p,q}_{n+1,k}(\alpha,\beta,r)=S^{2,p,q}_{n,k-1}(\alpha,\beta,r)+([k\alpha+r]_{p,q}-[n\beta]_{p,q})S^{2,p,q}_{n,k}(\alpha,\beta,r).$$

For their type-II (p,q)-analogues of $S^1_{n,k}(\alpha,\beta,r)$ and $S^2_{n,k}(\alpha,\beta,r)$, Remmel and Wachs replaced $(t-\gamma)$ by $[t-\gamma]_{p,q}$ in (1.3) and (1.4). That is, they defined $\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r)$ and $\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r)$ for $0\le k\le n$ by the following equations:

$$(1.11)\quad [t-r]_{p,q}[t-r-\alpha]_{p,q}\cdots[t-r-(n-1)\alpha]_{p,q}=\sum_{k=0}^{n}\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r)[t]_{p,q}[t-\beta]_{p,q}\cdots[t-(k-1)\beta]_{p,q}$$

and

$$(1.12)\quad [t]_{p,q}[t-\beta]_{p,q}\cdots[t-(k-1)\beta]_{p,q}=\sum_{k=0}^{n}\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r)[t-r]_{p,q}[t-r-\alpha]_{p,q}\cdots[t-r-(k-1)\alpha]_{p,q}.$$

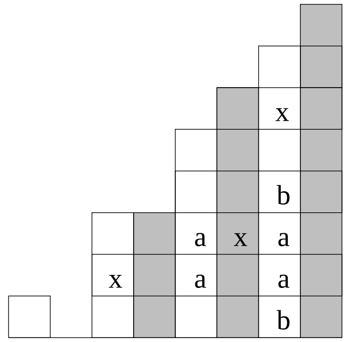


FIGURE 1. A placement in $(\mathcal{N}|\mathcal{F})^2_3(B_{\mathrm{BIP}}(1,0,3,3,5,6,7,8))$.

They further proved that when $0 \leq k \leq n$, the $\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r)$ and $\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r)$ defined according to equations (1.11) and (1.12) satisfy the following recursions:

$$\tilde{S}^{1,p,q}_{n+1,k}(\alpha,\beta,r) = q^{(k-1)\beta-n\alpha-r}\ \tilde{S}^{1,p,q}_{n,k-1}(\alpha,\beta,r) + p^{t-k\beta}[k\beta-n\alpha-r]_{p,q}\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r), \tag{1.13}$$

with initial conditions $\tilde{S}^{1,p,q}_{0,0}(\alpha,\beta,r) = 1$ and $\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r) = 0$ if $k < 0$ or $k > n$, and

$$\tilde{S}^{2,p,q}_{n+1,k}(\alpha,\beta,r) = q^{r+(k-1)\alpha-n\beta}\ \tilde{S}^{2,p,q}_{n,k-1}(\alpha,\beta,r) + p^{t-r-k\alpha}[k\alpha+r-n\beta]_{p,q}\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r), \tag{1.14}$$

with initial conditions $\tilde{S}^{2,p,q}_{0,0}(\alpha,\beta,r) = 1$ and $\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r) = 0$ if $k < 0$ or $k > n$.

Remmel and Wachs gave rook theory interpretations to $c^{i,j}_{n,k}(p,q) = (-1)^{n-k}S^{1,p,q}_{n,k}(j,0,i)$ and $S^{i,j}_{n,k}(p,q) = S^{2,p,q}_{n,k}(j,0,i)$ as well as $\tilde{c}^{i,j}_{n,k}(p,q) = (-1)^{n-k}\tilde{S}^{1,p,q}_{n,k}(j,0,i)$ and $\tilde{S}^{i,j}_{n,k}(p,q) = \tilde{S}^{1,p,q}_{n,k}(j,0,i)$ where i,j are nonnegative integers. Moreover, they were able to give combinatorial proofs of certain product formulas involving these polynomials. In this paper, we provide a generalization of their results by giving combinatorial interpretations to $S^{i,p,q}_{n,k}(\alpha,\beta,r)$ and $\tilde{S}^{i,p,q}_{n,k}(\alpha,\beta,r)$ when α, β and r are integers and $i \in \{1,2\}$, and we give combinatorial proofs to the product formulas that Remmel and Wachs did not provide.

2. A Rook Theoretic Model for $\mathbf{S^{1,p,q}_{n,k}}(\alpha,\beta,\mathbf{r})$ and $\mathbf{S^{2,p,q}_{n,k}}(\alpha,\beta,\mathbf{r})$

In this section, we give a rook theoretic model to interpret the type-I generalized p,q-Stirling numbers $S^{1,p,q}_{n,k}(\alpha,\beta,r)$ and $S^{2,p,q}_{n,k}(\alpha,\beta,r)$. The boards in our model are constructed as follows. Given any two finite sequences of nonnegative integers $\{a_1,a_2,\ldots,a_n\}$ and $\{b_1,b_2,\ldots,b_n\}$, we construct the *bipartite board* $B_{\mathrm{BIP}}(a_1,b_1,a_2,b_2,\ldots,\ a_n,b_n)$ whose column heights from left to right are $a_1,b_1,a_2,b_2,\ldots,a_n,b_n$. We will call the collection of columns whose heights are $a_1,a_2,\ldots,a_n$, the *Premier-columns* (P-columns), and the collection of columns whose heights are $b_1,b_2,\ldots,b_n$ the *Secondary-columns* (S-columns). For example, from the sequences $\{1,3,5,7\}$ and $\{0,3,6,8\}$, we obtain the board $B_{\mathrm{BIP}}(1,0,3,3,5,6,7,8)$ which is illustrated in Figure 1 with the P-columns given in white and the S-columns shaded in gray.

For any bipartite board $B = B_{\mathrm{BIP}}(a_1,b_1,a_2,b_2,\ldots,a_n,b_n)$, a rook r placed in a P-column (resp. S-column) of B is said to *j-attack* the cells in the P-columns of B that are strictly to the right of r in the first j rows that are weakly above r (resp. in the first j rows beginning with row 1) that are not j-attacked by any other rook that lies in a column to the left of r. Then, a placement $\mathbb{P}$ of rooks in $B_{\mathrm{BIP}}(a_1,b_1,a_2,b_2,\ldots,a_n,b_n)$ is called *j-nonattacking* if no rook in $\mathbb{P}$ is j-attacked by any rook in $\mathbb{P}$ to its left and there is at most one j-attacking rook per column pair $\{a_i,b_i\}$ for each $1 \leq i \leq n$. We let $(\mathcal{N}|\mathcal{F})^j_k(B)$ denote the set of all placements of k j-nonattacking rooks in B.

A placement in $(\mathcal{N}|\mathcal{F})^2_3(B_{\mathrm{BIP}}(1,0,3,3,5,6,7,8))$ is illustrated in Figure 1. As usual, rooks are denoted in the figure by an "x". In this example, the leftmost rook in B is placed in row 2 of column a_1 and 2-attacks the cells in rows 2 and 3 of columns a_3 and a_4. These cells 2-attacked by the leftmost rook contain an "a" in Figure 1. The second rook from the left in row 3 of column b_3 2-attacks the cells in rows 1 and 4 of column a_4. These cells 2-attacked by this second rook contain a "b" in Figure 1. The final rook of the placement is in row 6 of column a_4. Since there are no P-columns to the right of a_4, this rook does not j-attack any cells in the board.

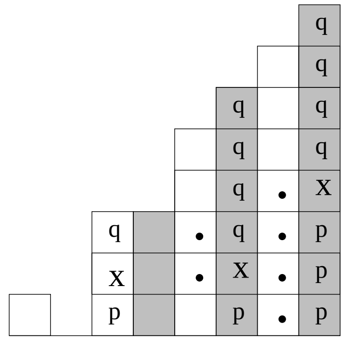


Figure 2. The p,q-weight of $\mathbb{P} \in (\mathcal{N}|\mathcal{F})_3^2(B_{\text{BIP}}(1,0,3,3,5,6,7,8))$ as contributed to $r_k^j(B,p,q)$.

For a nonnegative integer j, we say that a board $B = B_{\text{BIP}}(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$ is a *j-attacking bipartite board* if $0 \le a_1 \le a_2 \le \cdots \le a_n$, $0 \le b_1 \le b_2 \le \cdots \le b_n$, and for all placements of rooks in B, there are a sufficient number of cells in the P-columns of B for each rook to j-attack. By this definition, note that in the case when $b_1 = \cdots = b_n = 0$, the board $B_{\text{BIP}}(a_1, 0, a_2, 0, \cdots, a_n, 0)$ is a j-attacking bipartite board provided that for all $1 \le i < n$, $a_i \neq 0$ implies that $a_{i+1} \ge a_i + j - 1$. However, in the case when $b_i \neq 0$ for some $1 \le i < n$, B is a j-attacking bipartite board provided that $a_{j+1} \ge a_j + j$ for all $j > i \ge 1$. In Figure 1, the board $B_{\text{BIP}}(1,0,3,3,5,6,7,8)$ is a 2-attacking bipartite board.

Suppose that $\mathbb{P} \in (\mathcal{N}|\mathcal{F})_k^j(B)$ and set

$$\begin{aligned} n_S(\mathbb{P}) &= \text{the number of rooks of } \mathbb{P} \text{ placed in an } S\text{-column},\\ \mathcal{A}_B &= \text{the number of non-attacked cells in } B \text{ directly above some rook in } \mathbb{P},\\ \mathcal{B}_B &= \text{the number of non-attacked cells in } B \text{ directly below some rook in } \mathbb{P},\\ w_{p,q,B}^j(\mathbb{P}) &= (-1)^{n_S(\mathbb{P})} q^{\mathcal{A}_B} p^{\mathcal{B}_B}. \end{aligned}$$

The *type-I p,q-rook numbers*, denoted $r_k^j(B,p,q)$, are defined by

$$r_k^j(B,p,q) = \sum_{\mathbb{P} \in (\mathcal{N}|\mathcal{F})_k^j(B)} w_{p,q,B}^j(\mathbb{P}). \tag{2.1}$$

Here and in what follows, we will place a "$\bullet$" in the cells j-attacked by rooks in a given placement $\mathbb{P}$, a q in the cells that contribute a factor of q to $w_{p,q,B}^j(\mathbb{P})$, and a p in the cells that contribute a factor of p to $w_{p,q,B}^j(\mathbb{P})$. As illustrated in Figure 2 for $\mathbb{P} \in (\mathcal{N}|\mathcal{F})_3^2(B_{\text{BIP}}(1,0,3,3,5,6,7,8))$, $w_{p,q,B}^j(\mathbb{P}) = (-1)^2 q^9 p^5$.

Our first result is a p,q-analogue of Goldman, Joichi, and White's product formula [**5**].

Theorem 2.1. *Let $B = B_{BIP}(s, b_1, s+j, b_2, \ldots, s+(n-1)j, b_n)$. Then*

$$\begin{aligned} \sum_{k=0}^{n} r_{n-k}^j(B,p,q)([x]_{p,q} - [s]_{p,q})([x]_{p,q} - [s+j]_{p,q}) \cdots ([x]_{p,q} - [s+(k-1)j]_{p,q}) \\ = \quad ([x]_{p,q} - [b_1]_{p,q})([x]_{p,q} - [b_2]_{p,q}) \cdots ([x]_{p,q} - [b_n]_{p,q}). \end{aligned} \tag{2.2}$$

Proof. Given $B = B_{\text{BIP}}(s, b_1, s+j, b_2, \ldots, s+(n-1)j, b_n)$, we let $B_{(x,j)}$ be the board obtained from B by adjoining a single column of height $x + s + (i-1)j$ beneath the column pair $\{a_i, b_i\}$ for each $1 \le i \le n$. Here we call the line separating B from the adjoined rows the *bar*, the first x rows below the bar in $B_{(x,j)}$ the *x-adjoined rows* and the last $s + (n-1)j$ rows in $B_{(x,j)}$ below the bar the *j-adjoined rows*. Further, we will call the collection of cells in the column pair $\{a_i, b_i\}$ together with the $x + s + (i-1)j$ adjoined cells below it the ith *joined column*. The augmented board $B_{(x,j)}$ is illustrated in Figure 3.

For a given board B, placements of j-attacking rooks placed above the bar in $B_{(x,j)}$ will j-attack the same cells above the bar as described above. Additionally, any j-attacking rook r placed above the bar will attack all of the cells below it in its joined column as well as the first j rows in the j-adjoined rows strictly to the right of r not attacked by any rook to the left. A rook that is placed in one of the x-adjoined rows will attack all of the cells directly above it in the board B as well as the cells directly below it in the j-adjoined rows. A rook that is placed in a j-adjoined row will attack the cells directly above it in the x-adjoined

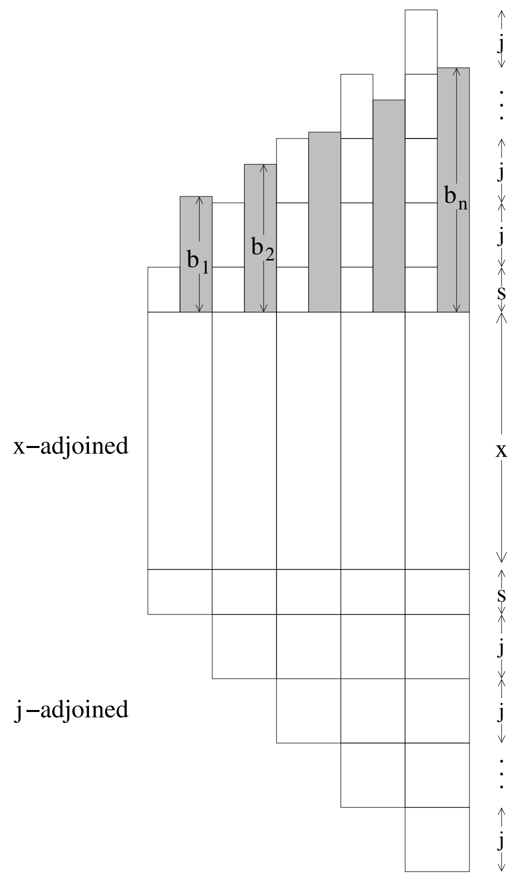


FIGURE 3. The board $B_{(x,j)}(s, b_1, s+j, b_2, \ldots, s+(n-1)j, b_n)$

rows and in the board B. The cells attacked by rooks of a placement in $B_{(x,j)}(1,0,3,3,5,6,7,8)$ have been illustrated in Figure 4.

Let $(\mathcal{N}|\mathcal{F})^j_n(B_{(x,j)})$ be the set of all placements of n rooks in $B_{(x,j)}$ such that no two rooks lie in the same joined column, no rook in $B_{(x,j)}$ is j-attacked by any rook in $B_{(x,j)}$ to its left. Thus, the placement in Figure 4 is in $(\mathcal{N}|\mathcal{F})^2_4(B_{\mathrm{BIP}}(1,0,3,3,5,6,7,8)_{(x,2)})$.

Then for positive integers x, the identity in (2.2) arises from two ways of counting

$$N = \sum_{\mathbb{P}\in(\mathcal{N}|\mathcal{F})^j_n(B_{(x,j)})} w^j_{p,q,B_{(x,j)}}(\mathbb{P}), \tag{2.3}$$

where $w^j_{p,q,B_{(x,j)}}(\mathbb{P})$ is defined as

$$w^j_{p,q,B_{(x,j)}}(\mathbb{P}) = (-1)^{n_S+n_j} q^{\mathcal{A}_{B_{(x,j)}}} p^{\mathcal{B}_{B_{(x,j)}}},$$

with

$$\begin{aligned}
n_S(\mathbb{P}) &= \text{the number of rooks of } \mathbb{P} \text{ placed in an } S\text{-column},\\
n_j(\mathbb{P}) &= \text{the number of rooks of } \mathbb{P} \text{ placed in a } j\text{-adjoined row},\\
\mathcal{A}_{B_{(x,j)}} &= \text{the number of non-attacked cells in } B_{(x,j)} \text{ directly above some rook in } \mathbb{P},\\
\mathcal{B}_{B_{(x,j)}} &= \text{the number of non-attacked cells in } B_{(x,j)} \text{ directly below some rook in } \mathbb{P}.
\end{aligned}$$

First we note that each placement $\mathbb{P} \in (\mathcal{N}|\mathcal{F})^j_n(B_{(x,j)})$ can be obtained by placing exactly one rook in each of the joined columns of $B_{(x,j)}$ proceeding from left to right. In the first joined column, the rook can be placed in either the first P-column, the first S-column, below the bar in the x-adjoined rows, or in the j-adjoined rows. It is easy to see that the contribution of the first joined column to N by placing the rook in the ith row from the top in the first P-column is $q^{i-1}p^{s-i}$ for a total contribution of $[s]_{p,q}$ to N. Likewise, the contribution of the first joined column to N by placing the rook in the ith row from the top in the first

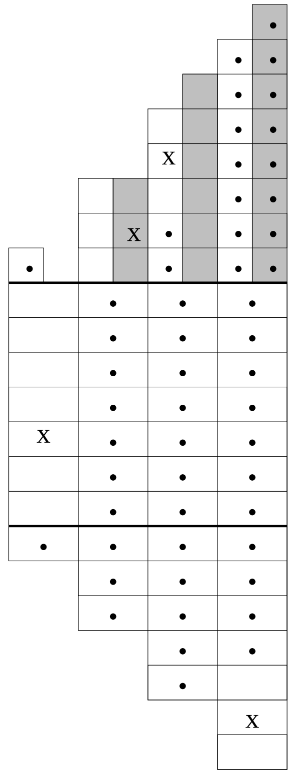


Figure 4. A placement in the board $B_{(x,2)}(B(1,0,3,3,5,6,7,8))$.

S-column is $-q^{i-1}p^{b_1-i}$ for a total contribution of $-[b_1]_{p,q}$ to N. Using the same analysis, we find that when the rook is placed below the bar in the x-adjoined rows, the contribution of the first joined column to N is $[x]_{p,q}$ while the total contribution is $-[s]_{p,q}$ from the placements in the j-adjoined rows. Therefore, the total contribution of the first joined column to N is $[x]_{p,q} - [b_1]_{p,q}$.

We now argue that regardless of the placement of the rook in the first joined column, the contributions from the P-column and the j-adjoined of the second adjoined column will cancel. To see this, first consider the case when a rook in the first joined column had been placed in B. Such a rook would attack exactly j cells in the P-columns as well as j cells in each row of the j-adjoined columns weakly to the right of the rook. In this case, the contribution from the second P-column is $[s]_{p,q}$ while the contribution from the second j-adjoined column is $-[s]_{p,q}$. On the other hand, if the rook in the first joined column had been placed below the bar, then the contribution from the second P-column is $[s+j]_{p,q}$ while the contribution from the second j-adjoined column is $-[s+j]_{p,q}$. To this end, we can argue as above, that the contribution of the second adjoined column to N is $[x]_{p,q} - [b_2]_{p,q}$.

Continuing in this way, we find that the total contribution of all n adjoined columns to N is

$$([x]_{p,q} - [b_1]_{p,q})([x]_{p,q} - [b_2]_{p,q}) \cdots ([x]_{p,q} - [b_n]_{p,q}).$$

Now suppose that a placement $\mathbb{Q}$ of $n-k$ rooks is fixed in B. Then a placement $\mathbb{P} \in (\mathcal{N}|\mathcal{F})_n^j(B_{(x,j)})$ can be obtained from $\mathbb{Q}$ by placing the remaining k rooks below the bar. As prescribed, each of the $n-k$ rooks in B will attack j cells in the j-adjoined rows in the columns weakly to the right of each rook. As such, there will be x places in the x-adjoined rows and $s+(i-1)j$ places in the j-adjoined rows in which to place the ith rook below the bar from the left, for $1 \le i \le k$. Therefore, the placement of the k rooks below the bar will contribute a factor of $([x]_{p,q} - [s]_{p,q})([x]_{p,q} - [s+j]_{p,q}) \cdots ([x]_{p,q} - [s+(k-1)j]_{p,q})$ to N. Furthermore, each rook placed below the bar will attack the cells above it in B implying that $w^j_{p,q,B}(\mathbb{Q}) = w^j_{p,q,B}(\mathbb{P} \cap B)$.

Thus,

$$\begin{aligned}
N &= \sum_{k=0}^{n} \sum_{\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^j_{n-k}(B)} \sum_{\substack{\mathbb{P}\in(\mathcal{N}|\mathcal{F})^j_n(B_{(x,j)})\\ \mathbb{P}\cap B=\mathbb{Q}}} w^j_{p,q,B_{(x,j)}}(\mathbb{P})\\
&= \sum_{k=0}^{n} \sum_{\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^j_{n-k}(B)} w^j_{p,q,B}(\mathbb{P}\cap B)([x]_{p,q}-[s]_{p,q})([x]_{p,q}-[s+j]_{p,q})\cdots([x]_{p,q}-[s+(k-1)j]_{p,q})\\
&= \sum_{k=0}^{n} r^j_{n-k}(B,p,q)([x]_{p,q}-[s]_{p,q})([x]_{p,q}-[s+j]_{p,q})\cdots([x]_{p,q}-[s+(k-1)j]_{p,q}).
\end{aligned}$$

□

We are now in a position to give combinatorial interpretations to $S^{1,p,q}_{n,k}(\alpha,\beta,r)$ and $S^{2,p,q}_{n,k}(\alpha,\beta,r)$ defined by (1.5) and (1.6). To begin, let x, y, μ and ν be nonnegative integers and let $B^{x,y}_{\mu,\nu,n}$ denote the bipartite board $B_{\text{BIP}}(x,y,x+\mu,y+\nu,x+2\mu,y+2\nu,\ldots,x+(n-1)\mu,y+(n-1)\nu)$. Then,

THEOREM 2.2. *If n and k are nonnegative integers for which $0<k<n$, then*

$$S^{1,p,q}_{n,k}(\alpha,\beta,r) = r^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n},p,q) \quad \text{and} \tag{2.4}$$

$$S^{2,p,q}_{n,k}(\alpha,\beta,r) = r^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n},p,q). \tag{2.5}$$

PROOF. We begin by noting that the identities in (2.4) and (2.5) can be proved by showing that the p,q-rook numbers $r^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n},p,q)$ satisfy the same recursion as $S^{1,p,q}_{n,k}(\alpha,\beta,r)$ given in (1.7) and (1.8) and that $r^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n},p,q)$ satisfy the same recursion as $S^{2,p,q}_{n,k}(\alpha,\beta,r)$ given in (1.9) and (1.10).

For $n=0$, $B^{x,y}_{\mu,\nu,0}=\emptyset$. So, it immediately follows from our definition that

$$r^{\beta}_0(B^{0,r}_{\beta,\alpha,0},p,q)=1 \quad \text{and} \quad r^{\alpha}_0(B^{r,0}_{\alpha,\beta,0},p,q)=1.$$

Clearly, $r^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n},p,q)=0$ and $r^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n},p,q)=0$ if $k>n$ or $k<0$ since both $(\mathcal{N}|\mathcal{F})^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n})$ and $(\mathcal{N}|\mathcal{F})^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n})$ are empty if $k>n$ or $k<0$. Therefore, to verify the equalities in (2.4) and (2.5), it remains to show that for all $n\geq 1$ and $0\leq k\leq n$,

$$r^{\beta}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1},p,q) = r^{\beta}_{n-(k-1)}(B^{0,r}_{\beta,\alpha,n},p,q) + ([k\beta]_{p,q}-[n\alpha+r]_{p,q})\, r^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n},p,q) \tag{2.6}$$

and

$$r^{\alpha}_{n+1-k}(B^{r,0}_{\alpha,\beta,n+1},p,q) = r^{\alpha}_{n-(k-1)}(B^{r,0}_{\alpha,\beta,n},p,q) + ([k\alpha+r]_{p,q}-[n\beta]_{p,q})\, r^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n},p,q). \tag{2.7}$$

To prove (2.6), we note that the set of elements in $(\mathcal{N}|\mathcal{F})^{\beta}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1})$ can be partitioned into the sets No, $P-Last$, and $S-Last$ where No consists of the placements of $(\mathcal{N}|\mathcal{F})^{\beta}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1})$ with no rook in the column pair $\{a_{n+1},b_{n+1}\}$, $P-Last$ consists of the placements of $(\mathcal{N}|\mathcal{F})^{\beta}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1})$ with a rook in the P-column a_{n+1}, and $S-Last$ consists of the placements of $(\mathcal{N}|\mathcal{F})^{\beta}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1})$ with a rook in the S-column b_{n+1}. Then,

$$\begin{aligned}
r^{\beta}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1},p,q) &= \sum_{\mathbb{P}\in(\mathcal{N}|\mathcal{F})^{\beta}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1})} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P})\\
&= \sum_{\mathbb{P}\in No} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}) + \sum_{\mathbb{P}\in P-Last} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}) + \sum_{\mathbb{P}\in S-Last} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}).
\end{aligned}$$

It is easy to see that a placement in $\mathbb{P}\in No$ has $n-(k-1)$ rooks to the left of the column pair $\{a_{n+1},b_{n+1}\}$. Thus, $w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}) = w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n}}(\mathbb{Q})$ where $\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^{\beta}_{n-(k-1)}(B^{0,r}_{\beta,\alpha,n})$ is the placement that would result in eliminating the last pair of columns $\{a_{n+1},b_{n+1}\}$ from $B^{0,r}_{\beta,\alpha,n+1}$. Therefore,

$$\sum_{\mathbb{P}\in No} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}) = \sum_{\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^{\beta}_{n-(k-1)}(B^{0,r}_{\beta,\alpha,n})} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n}}(\mathbb{Q}) = r^{\beta}_{n-(k-1)}(B^{0,r}_{\beta,\alpha,n},p,q).$$

To compute $\sum_{\mathbb{P}\in P-Last} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P})$, we first observe that each $\mathbb{Q} \in (\mathcal{N}|\mathcal{F})^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n})$ can be extended to $k\beta$ placements in $P - Last$ by placing an additional rook in a non-attacked cell of column a_{n+1}. This follows since each of the $n-k$ rooks of a fixed $\mathbb{Q} \in (\mathcal{N}|\mathcal{F})^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n})$ attacks β cells in column a_{n+1} leaving $n\beta - (n-k)\beta = k\beta$ non-attacked cells in column a_{n+1} in which to place the additional rook. Next, we note that $n_S(\mathbb{P}) = n_S(\mathbb{Q})$. So, if the additional rook is placed in the ith non-attacked cell from the top, then the weight of the corresponding placement $\mathbb{P}^i$ is $q^{i-1}p^{k\beta-i}w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n}}(\mathbb{Q})$. Therefore,

$$\begin{aligned}\sum_{\mathbb{P}\in P-Last} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}) &= \sum_{\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n})} (p^{k\beta-1} + qp^{k\beta-2} + \cdots + q^{k\beta-2}p + q^{k\beta-1})\; w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n}}(\mathbb{Q}) \\ &= [k\beta]_{p,q} r^{\beta}_{n,k}(B^{0,r}_{\beta,\alpha,n}, p, q).\end{aligned}$$

Finally, we observe that each rook of $\mathbb{P} \in S - Last$ attacks β cells of column a_{n+1} but no cells of column b_{n+1}. Accordingly, $\sum_{\mathbb{P}\in S-Last} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P})$ could be computed by extending each $\mathbb{Q} \in (\mathcal{N}|\mathcal{F})^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n})$ to $n\alpha + r$ distinct placements in $(\mathcal{N}|\mathcal{F})^{\beta}_{n-k+1}(B^{0,r}_{\beta,\alpha,n+1})$ by placing an additional rook in any of the $n\alpha + r$ non-attacked cells of column b_{n+1}. For such a placement $\mathbb{P}^i$ obtained by placing the additional rook in the ith non-attacked cell from the top, we note that $n_S(\mathbb{P}^i) = 1 + n_S(\mathbb{Q})$ and consequently $w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}^i) = -q^{i-1}p^{n\alpha+r+i}w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n}}(\mathbb{Q})$. Therefore, it follows that

$$\begin{aligned}&\sum_{\mathbb{P}\in S-Last} w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}) \\ &= \sum_{\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n})} -(p^{n\alpha+r-1} + qp^{n\alpha+r-2} + \cdots + q^{n\alpha+r-2}p + q^{n\alpha+r-1})\; w^{\beta}_{p,q,B^{0,r}_{\beta,\alpha,n}}(\mathbb{Q}) \\ &= -[n\alpha+r]_{p,q}\, r^{\beta}_{n,k}(B^{0,r}_{\beta,\alpha,n}, p, q).\end{aligned}$$

In the same way, we prove (2.7) by partitioning $(\mathcal{N}|\mathcal{F})^{\alpha}_{n+1-k}(B^{r,0}_{\alpha,\beta,n+1})$ into the sets No, $P - Last$, and $S - Last$ where No consists of the placements of $(\mathcal{N}|\mathcal{F})^{\alpha}_{n+1-k}(B^{r,0}_{\alpha,\beta,n+1})$ with no rook in the column pair $\{a_{n+1}, b_{n+1}\}$, $P - Last$ consists of the placements of $(\mathcal{N}|\mathcal{F})^{\alpha}_{n+1-k}(B^{r,0}_{\alpha,\beta,n+1})$ with a rook in the P-column a_{n+1}, and $S - Last$ consists of the placements of $(\mathcal{N}|\mathcal{F})^{\alpha}_{n+1-k}(B^{r,0}_{\alpha,\beta,n+1})$ with a rook in the S-column b_{n+1}. The recursion in (2.7) will follow by showing that

$$\begin{aligned}r^{\alpha}_{n+1-k}(B^{0,r}_{\beta,\alpha,n+1}, p, q) &= \sum_{\mathbb{P}\in(\mathcal{N}|\mathcal{F})^{\alpha}_{n+1-k}(B^{r,0}_{\alpha,\beta,n+1})} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P}) \\ &= \sum_{\mathbb{P}\in No} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P}) + \sum_{\mathbb{P}\in P-Last} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P}) + \sum_{\mathbb{P}\in S-Last} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P}).\end{aligned}$$

Again, it is easy to see that

$$\sum_{\mathbb{P}\in No} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P}) = r^{\alpha}_{n-(k-1)}(B^{r,0}_{\alpha,\beta,n}, p, q).$$

To compute $\sum_{\mathbb{P}\in P-Last} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P})$, we observe that each fixed placement $\mathbb{Q} \in (\mathcal{N}|\mathcal{F})^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n})$ can be extended to $k\alpha + r$ distinct placements in $(\mathcal{N}|\mathcal{F})^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n})$ by placing an additional rook in one of the $n\alpha + r - \alpha(n-k) = k\alpha + r$ non-attacked cells of column a_{n+1}. If the additional rook is placed in the ith non-attacked cells from the top of column a_{n+1}, then the weight of the corresponding placement $\mathbb{P}^i$ is $q^{i-1}p^{k\alpha+r-i}w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n}}(\mathbb{Q})$. It follows that

$$\begin{aligned}\sum_{\mathbb{P}\in P-Last} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P}) &= \sum_{\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n})} (p^{k\alpha+r-1} + qp^{k\alpha+r-2} + \cdots + q^{k\alpha+r-2}p + q^{k\alpha+r-1})\; w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n}}(\mathbb{Q}) \\ &= [k\alpha+r]_{p,q} r^{\alpha}_{n,k}(B^{r,0}_{\alpha,\beta,n}, p, q).\end{aligned}$$

As above, we observe that each rook of $\mathbb{P} \in S-Last$ attacks α cells of column a_{n+1} but no cells of column b_{n+1}. Therefore, $\sum_{\mathbb{P}\in S-Last} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P})$ could be computed by extending each $\mathbb{Q} \in (\mathcal{N}|\mathcal{F})^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n})$ to $n\beta$ distinct placements in $(\mathcal{N}|\mathcal{F})^{\alpha}_{n-k+1}(B^{r,0}_{\alpha,\beta,n+1})$ by placing an additional rook in any of the $n\beta$ non-attacked cells of column b_{n+1}. For such a placement $\mathbb{P}^i$ obtained by placing the additional rook in the ith non-attacked cell from the top, we note that $n_S(\mathbb{P}^i) = 1 + n_S(\mathbb{Q})$ and thus $w^{\alpha}_{p,q,B^{0,r}_{\beta,\alpha,n+1}}(\mathbb{P}^s) = -q^{i-1}p^{n\beta+i}w^{\alpha}_{p,q,B^{0,r}_{\beta,\alpha,n}}(\mathbb{Q})$. To this end,

$$\begin{aligned}\sum_{\mathbb{P}\in S-Last} w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n+1}}(\mathbb{P}) &= \sum_{\mathbb{Q}\in(\mathcal{N}|\mathcal{F})^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n})} -(p^{n\beta-1}+qp^{n\beta-2}+\cdots+q^{n\beta-2}p+q^{n\beta-1})\, w^{\alpha}_{p,q,B^{r,0}_{\alpha,\beta,n}}(\mathbb{Q}) \\ &= -[n\beta]_{p,q} r^{\alpha}_{n,k}(B^{r,0}_{\alpha,\beta,n},p,q).\end{aligned}$$

□

We end this section by noting that as a consequence of Theorems 2.1 and 2.2, our single model described above yields a combinatorial interpretation to both (1.5) and (1.6). In particular, (1.5) is obtained from (2.2) by setting $s = 0$, $j = \beta$, and $b_i = r + (i-1)\alpha$. Likewise, setting $s = r$, $j = \alpha$, and $b_i = (i-1)\beta$ in (2.2) produces (1.5).

3. A Rook Theoretic Model for $\tilde{\mathbf{S}}^{\mathbf{1},\mathbf{p},\mathbf{q}}_{\mathbf{n},\mathbf{k}}(\alpha,\beta,\mathbf{r})$ and $\tilde{\mathbf{S}}^{\mathbf{2},\mathbf{p},\mathbf{q}}_{\mathbf{n},\mathbf{k}}(\alpha,\beta,\mathbf{r})$

To define the second type of p,q-rook numbers, let B be a j-attacking bipartite board and suppose $\mathbb{P} \in (\mathcal{N}|\mathcal{F})^j_k(B)$. Assume additionally that the k rooks are in the column pairs $\{a_i, b_i\}$ with labels $1 \le c_1 < \cdots < c_k \le n$ and that there are j_i non-attacked cells in the column containing the rook among the pair $\{a_{c_i}, b_{c_i}\}$ for $1 \le i \le k$. Setting

$$\begin{aligned} a_B &= \text{the number of non-attacked cells in } B \text{ directly above some rook in } \mathbb{P}, \\ b_B &= \text{the number of non-attacked cells in } B \text{ directly below some rook in } \mathbb{P}, \\ \varepsilon_B &= \text{the number of non-attacked cells in a } P\text{-column of an } \{a_i,b_i\} \text{ pair containing no rook,}\end{aligned}$$

we define the *type-II p,q-rook numbers*, denoted $\tilde{r}^j_k(B,p,q)$, by

$$\tilde{r}^j_k(B,p,q) = q^{-(b_1+\cdots+b_n)} \sum_{\mathbb{P}\in(\mathcal{N}|\mathcal{F})^j_k(B)} (-1)^{n_S} q^{\varepsilon_B(\mathbb{P})+a_B(\mathbb{P})} p^{b_B(\mathbb{P})+kt-(j_1+j_2+\cdots+j_k)}. \tag{3.1}$$

The following result gives the generalized product formula for the type-II p,q-rook numbers.

Theorem 3.1. *Let $B = B_{BIP}(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$ be a j-attacking bipartite board. Then for each nonnegative integer n,*

$$\begin{aligned}\sum_{k=0}^{n} \tilde{r}^j_{n-k}(B,p,q)[t]_{p,q}[t-j]_{p,q}\cdots[t-(k-1)j]_{p,q} \\ = \prod_{i=1}^{n} q^{-b_i}\big([t+a_i-(i-1)j]_{p,q} - p^{t+a_i-(i-1)j-b_i}[b_i]_{p,q}\big).\end{aligned} \tag{3.2}$$

The proof of Theorem 3.1 is similar to that of Theorem 2.1. The idea is to consider all placements of n j-attacking rooks in the board $B^j_t(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$ which is obtained from B by adjoining t rows below the n P-columns, labeled from bottom to top by $1, 2, \ldots, t$. The board $B^j_t(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$ is illustrated in Figure 5. Here, rooks placed in B will j-attack in the P-columns as usual, while a rook that is placed in row i below a P-column will j-attack the cells to the right in the first j rows weakly above it in the list of rows $i, i+1, \ldots, t, 1, \ldots, i-1$ that have not been j-attacked by a rook from the left. Such a placement of n rooks is illustrated in Figure 5 with $j = 2$.

It can also be shown that the type-II p,q-rook numbers on specific j-attacking bipartite boards satisfy the same recursions as the type-II Stirling numbers of the first and second kind. To see this, we first note that

$$[k\beta - n\alpha - r]_{p,q} = q^{-n\alpha-r}\left([k\beta]_{p,q} - p^{k\beta-n\alpha-r}[n\alpha+r]\right) \tag{3.3}$$

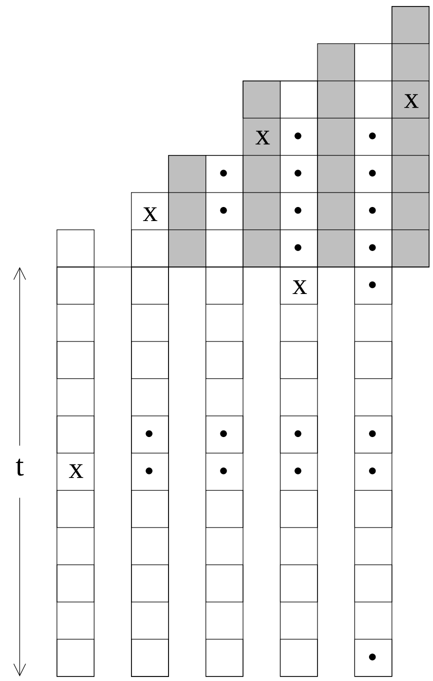


Figure 5. The board $B_t^j(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$

and

$$[k\alpha + r - n\beta]_{p,q} = q^{-n\beta}\left([k\alpha + r]_{p,q} - p^{k\alpha+r-n\beta}[n\beta]\right). \tag{3.4}$$

Then, substituting the identity (3.3) into the recursion (1.13) and (3.4) into (1.14) yields the following three term recursions:

$$\begin{aligned} \tilde{S}^{1,p,q}_{n+1,k}(\alpha,\beta,r) &= q^{(k-1)\beta-n\alpha-r}\ \tilde{S}^{1,p,q}_{n,k-1}(\alpha,\beta,r) + p^{t-k\beta}q^{-n\alpha-r}[k\beta]_{p,q}\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r) \\ &\quad -p^{t-n\alpha-r}q^{-n\alpha-r}[n\alpha+r]_{p,q}\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r). \end{aligned} \tag{3.5}$$

$$\begin{aligned} \tilde{S}^{2,p,q}_{n+1,k}(\alpha,\beta,r) &= q^{r+(k-1)\alpha-n\beta}\ \tilde{S}^{2,p,q}_{n,k-1}(\alpha,\beta,r) + p^{t-r-k\alpha}q^{-n\beta}[k\alpha+r]_{p,q}\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r) \\ &\quad -p^{t-n\beta}q^{-n\beta}[n\beta]_{p,q}\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r). \end{aligned} \tag{3.6}$$

As in the proof of Theorem 2.2, we can show that the rook numbers $\tilde{r}^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n})$ and $\tilde{r}^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n})$ satisfy the respective recursions in (3.5) and (3.5) by again partitioning the set of placements $(\mathcal{N}|\mathcal{F})^{\beta}_{n+1}(B^{0,r}_{\beta,\alpha,n+1})$ and $(\mathcal{N}|\mathcal{F})^{\alpha}_{n+1}(B^{r,0}_{\alpha,\beta,n+1})$ into *No*, $P - Last$, and $S - Last$. We summarize these results in the following:

Theorem 3.2. *If n and k are nonnegative integers for which $0 < k < n$, then*

$$\tilde{S}^{1,p,q}_{n,k}(\alpha,\beta,r) = \tilde{r}^{\beta}_{n-k}(B^{0,r}_{\beta,\alpha,n}) \quad \textit{and} \tag{3.7}$$

$$\tilde{S}^{2,p,q}_{n,k}(\alpha,\beta,r) = \tilde{r}^{\alpha}_{n-k}(B^{r,0}_{\alpha,\beta,n}). \tag{3.8}$$

As a consequence of Theorems 3.1 and 3.2, this single rook theoretic model yields a combinatorial interpretation for the identities given in (1.11) and (1.12). To see this, we observe that

$$q^{-b_i}\left([t + a_i - (i-1)j]_{p,q} - p^{t+a_i-(i-1)j-b_i}[b_i]_{p,q}\right) = [t + a_i - (i-1)j - b_i]_{p,q}.$$

Then from (3.2), (1.11) is obtained by setting $j = \beta$, $a_i = (i-1)\beta$, and $b_i = r + (i-1)\alpha$ as is (1.12) by setting $j = \alpha$, $a_i = r + (i-1)\alpha$, and $b_i = (i-1)\beta$, and replacing t with $t - r$.

4. Directions

While our models have provided a rook theoretic interpretation for both types of p,q-analogues of the generalized Stirling numbers of the first and second kind, their product formulas and recursions, we have yet to produce analogues of the orthogonality relations given by Hsu and Shiue [**7**] for arbitrary parameters α, β, and r:

$$\sum_{k=i}^{n} \overline{S}^{1}_{n,k}(\alpha,\beta,r)\overline{S}^{2}_{k,i}(\alpha,\beta,r) = \sum_{k=i}^{n} \overline{S}^{2}_{n,k}(\alpha,\beta,r)\overline{S}^{1}_{k,i}(\alpha,\beta,r) = \chi(i=n).$$

Although, Remmel and Wachs gave direct combinatorial interpretations of the following p,q-analogues of the orthogonality relations

$$\sum_{k=r}^{n} S^{2,p,q}_{n,k}(j,0,i)S^{1,p,q}_{k,r}(j,0,i) = \chi(r=n)$$

and

$$\sum_{k=r}^{n} p^{\binom{n-k+1}{2}}\tilde{S}^{2,p,q}_{n,k}(j,0,i)(pq)^{\binom{k}{2}j}p^{-ir}q^{-ik}\tilde{S}^{1,p,q}_{k,r}(j,0,i) = \chi(r=n),$$

they did not provide the p,q-orthogonality relations for arbitrary parameters α, β, and r. We will pursue this problem in a subsequent paper.

References

[1] K. S. Briggs, A Combinatorial Interpretation for p,q-Hit Numbers, *Discrete Mathematics*, submitted.

[2] K. S. Briggs, Q-analogues and p,q-analogues of rook numbers and hit numbers and their extensions, Ph.D. thesis, University of California, San Diego (2003).

[3] K. S. Briggs and J. B. Remmel, A p,q-analogue of a Formula of Frobenius, *Electron. J. Comb.* **10** (2003), #R9.

[4] A. M. Garsia and J. B. Remmel, Q-Counting Rook Configurations and a Formula of Frobenius, *J. Combin. Theory Ser. A* **41**, 246-275.

[5] J. R. Goldman, J. T. Joichi, and D. E. White, Rook Theory I. Rook equivalence of Ferrers boards, *Proc. Amer. Math Soc.* **52** (1975), 485-492.

[6] H. G. Gould, The q-Stirling numbers of the first and second kinds, *Duke Math. J.* **28** (1961), 281-289.

[7] L. C. Hsu and P. J. S. Shiue, A Unified Approach to Generalized Stirling Numbers, *Advances in Applied Mathematics* **20** (1998), 366-384.

[8] A. de Médicis and P. Leroux, A unified combinatorial approach for q-(and p,q-)Stirling numbers, *J. Statist. Plann. Inference* **34** (1993), 89-105.

[9] A. de Médicis and P. Leroux, Generalized Stirling Numbers, Convolution Formulae and p,q-analogues, *Can. J. Math.* **47** (1995), 474-499.

[10] S. C. Milne, Restricted growth functions, rank row matchings of partition lattices, and q-Stirling numbers, *Adv. in Math.* **43** (1982), 173-196.

[11] J. B. Remmel and M. Wachs, Rook Theory, Generalized Stirling Numbers and (p,q)-analogues, *Electron. J. Comb.* **11** (2004), #R84.

[12] I. Kaplansky and J. Riordan, The problem of rooks and its applications, *Duke Math. J.* **13** (1946), 259-268.

[13] M. Wachs and D. White, p,q-Stirling numbers and set partition statistics, *J. Combin. Theory Ser. A* **56** (1991), 27-46.

[14] M. Wachs, σ-Resticted growth functions and p,q-Stirling numbers, *J. Combin. Theory Ser. A* **68** (1994), 470-480.

Department of Mathematics, Clayton State University, USA, 30260

E-mail address: `karenbriggs@clayton.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Restricted Dumont permutations, Dyck paths, and noncrossing partitions

Alexander Burstein, Sergi Elizalde, and Toufik Mansour

Abstract. We complete the enumeration of Dumont permutations of the second kind avoiding a pattern of length 4 which is in turn a Dumont permutation of the second kind. We also consider some combinatorial statistics on Dumont permutations avoiding certain patterns of length 3 and 4 and give a natural bijection between 3142-avoiding Dumont permutations of the second kind and noncrossing partitions that uses cycle decomposition, as well as bijections between 132-, 231- and 321-avoiding Dumont permutations and Dyck paths.

Résumé. Nous complétons l'énumeration des permutations Dumont de deuxième espèce évitant un motif de longueur 4 étant elle-même une permutation Dumont de deuxième espèce. Nous considérons aussi quelques statistiques combinatoires sur les permutations Dumont évitant certains motifs de longueur 3 et 4 et nous démontrons une bijection naturelle entre les permutations Dumont de deuxième espèce évitant le motif 3142 et les partitions non-croisées via le biais de décompositions cycliques, aussi bien qu'une bijection entre les permutations Dumont de deuxième espèce évitant les motifs 132, 231, 321 et les chemins de Dyck.

1. Preliminaries

The main goal of this paper is to give analogues of known enumerative results on certain classes of permutations characterized by pattern-avoidance. Instead of taking the symmetric group S_n, we consider the subset of Dumont permutations (see definition below), and we identify classes of restricted permutations with enumerative properties that are analogous to the case of general permutations. More precisely, we study the number of Dumont permutations of length $2n$ avoiding either a 3-letter pattern or a 4-letter pattern. We also give direct bijections between equinumerous sets of restricted Dumont permutations of length $2n$ and other objects such as restricted permutations of length n, Dyck paths of semilength n, or noncrossing partitions of $[n] = \{1, 2 \dots, n\}$.

1.1. Patterns. Let $\sigma \in S_n$ and $\tau \in S_k$ be two permutations. We say that τ *occurs* in σ, or that σ *contains* τ, if σ has a subsequence $(\sigma(i_1), \dots, \sigma(i_k))$, $1 \le i_1 < \dots < i_k \le n$, that is order-isomorphic to τ (in other words, for any j_1 and j_2, $\sigma(i_{j_1}) \le \sigma(i_{j_2})$ if and only if $\tau(j_1) \le \tau(j_2)$). Such a subsequence is called an *occurrence* (or an *instance*) of τ in σ. In this context, the permutation τ is called a *pattern*. If τ does not occur in σ, we say that σ *avoids* τ, or is τ*-avoiding*. We denote by $S_n(\tau)$ the set of permutations in S_n avoiding a pattern τ. If T is a set of patterns, then $S_n(T) = \bigcap_{\tau \in T} S_n(\tau)$, i.e. $S_n(T)$ is the set of permutations in S_n avoiding all patterns in T.

The first results in the extensive body of research on permutations avoiding a 3-letter pattern are due to Knuth [**9**], but the intensive study of patterns in permutations began with Simion and Schmidt [**16**], who considered permutations and involutions avoiding each set T of 3-letter patterns. One of the most frequently considered problems is the enumeration of $S_n(\tau)$ and $S_n(T)$ for various patterns τ and sets of patterns T. The inventory of cardinalities of $|S_n(T)|$ for $T \subseteq S_3$ is given in [**16**], and a similar inventory for $|S_n(\tau_1, \tau_2)|$, where $\tau_1 \in S_3$ and $\tau_2 \in S_4$ is given in [**23**]. Some results on $|S_n(\tau_1, \tau_2)|$ for $\tau_1, \tau_2 \in S_4$ are obtained in [**22**]. The exact formula for $|S_n(1234)|$ and the generating function for $|S_n(12 \dots k)|$ are found

2000 *Mathematics Subject Classification.* 05A05, 05A15.
Key words and phrases. Dumont permutations, Dyck paths, forbidden subsequences, noncrossing partitions.

in [**7**]. Bóna [**2**] has found the exact value of $|S_n(1342)| = |S_n(1423)|$, and Stankova [**18, 19**] showed that $|S_n(3142)| = |S_n(1342)|$. For a survey of results on pattern avoidance, see [**1, 8**].

Another problem is finding equinumerously avoided (sets of) patterns, i.e. sets T_1 and T_2 such that $|S_n(T_1)| = |S_n(T_2)|$ for any $n \geq 0$. Such (sets of) patterns are called *Wilf-equivalent* and said to belong to the same *Wilf class*. The following symmetry operations on S_n map every pattern onto a Wilf-equivalent pattern:

- *reversal* r: $r(\tau)(i) = \tau(n+1-i)$, i.e. $r(\tau)$ is τ read right-to-left.
- *complement* c: $c(\tau)(i) = n+1-\tau(i)$, i.e. $c(\tau)$ is τ read upside down.
- $r \circ c = c \circ r$: $r \circ c(\tau)(i) = n+1-\tau(n+1-i)$, i.e. $r \circ c(\tau)$ is τ read right-to-left upside down.

The set of patterns $\langle r, c\rangle(\tau) = \{\tau, r(\tau), c(\tau), r(c(\tau)) = c(r(\tau))\}$ is called the *symmetry class* of τ.

Sometimes we will represent a permutation $\pi \in S_n$ by placing dots on an $n \times n$ board. For each $i = 1, \ldots, n$ we will place a dot with abscissa i and ordinate $\pi(i)$ (the origin of the board is at the bottom-left corner).

1.2. Dumont permutations. In this paper we give a complete answer for the above problems when we restrict our attention to the set of Dumont permutations. A *Dumont permutation of the first kind* is a permutation $\pi \in S_{2n}$ where each even entry is followed by a descent and each odd entry is followed by an ascent or ends the string. In other words, for every $i = 1, 2, \ldots, 2n$,

$$\pi(i) \text{ is even} \implies i < 2n \text{ and } \pi(i) > \pi(i+1),$$
$$\pi(i) \text{ is odd} \implies \pi(i) < \pi(i+1) \text{ or } i = 2n.$$

A *Dumont permutation of the second kind* is a permutation $\pi \in S_{2n}$ where all entries at even positions are deficiencies and all entries at odd positions are fixed points or excedances. In other words, for every $i = 1, 2, \ldots, n$,

$$\pi(2i) < 2i,$$
$$\pi(2i-1) \geq 2i-1.$$

We denote the set of Dumont permutations of the first (resp. second) kind of length $2n$ by $\mathfrak{D}^1_{2n}$ (resp. $\mathfrak{D}^2_{2n}$). For example, $\mathfrak{D}^1_2 = \mathfrak{D}^2_2 = \{21\}$, $\mathfrak{D}^1_4 = \{2143, 3421, 4213\}$, $\mathfrak{D}^2_4 = \{2143, 3142, 4132\}$. We also define $\mathfrak{D}^1$-*Wilf-equivalence* and $\mathfrak{D}^2$-*Wilf-equivalence* similarly to the Wilf-equivalence on S_n. Dumont [**4**] showed that

$$|\mathfrak{D}^1_{2n}| = |\mathfrak{D}^2_{2n}| = G_{2n+2} = 2(1-2^{2n+2})B_{2n+2},$$

where G_n is the nth Genocchi number, a multiple of the Bernoulli number B_n. Lists of Dumont permutations $\mathfrak{D}^1_{2n}$ and $\mathfrak{D}^2_{2n}$ for $n \leq 4$ as well as some basic information and references for Genocchi numbers and Dumont permutations may be obtained in [**15**] and [**17**, A001469]. The exponential generating functions for the unsigned and signed Genocchi numbers are as follows:

$$\sum_{n=1}^{\infty} G_{2n}\frac{x^{2n}}{(2n)!} = x\tan\frac{x}{2}, \qquad \sum_{n=1}^{\infty}(-1)^n G_{2n}\frac{x^{2n}}{(2n)!} = \frac{2x}{e^x+1} - x = -x\tanh\frac{x}{2}.$$

Some cardinalities of sets of restricted Dumont permutations of length $2n$ parallel those of restricted permutations of length n. For example, the following results were obtained in [**3, 11**]:

- $|\mathfrak{D}^1_{2n}(\tau)| = C_n$ for $\tau \in \{132, 231, 312\}$, where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the n-th Catalan number.
- $|\mathfrak{D}^2_{2n}(321)| = C_n$.
- $|\mathfrak{D}^1_{2n}(213)| = C_{n-1}$, so r, c and $r \circ c$ do not necessarily produce $\mathfrak{D}^1$-Wilf-equivalent patterns.
- $|\mathfrak{D}^2_{2n}(231)| = 2^{n-1}$, while $|\mathfrak{D}^2_{2n}(312)| = 1$ and $|\mathfrak{D}^2_{2n}(132)| = |\mathfrak{D}^2_{2n}(213)| = 0$ for $n \geq 3$, so r, c and $r \circ c$ do not necessarily produce $\mathfrak{D}^2$-Wilf-equivalent patterns either.
- $|\mathfrak{D}^2_{2n}(3142)| = C_n$.
- $|\mathfrak{D}^1_{2n}(1342, 1423)| = |\mathfrak{D}^1_{2n}(2341, 2413)| = |\mathfrak{D}^1_{2n}(1342, 2413)| = s_{n+1}$, the $(n+1)$-st little Schröder number [**17**, A001003], given by $s_1 = 1$, $s_{n+1} = -s_n + 2\sum_{k=1}^{n} s_k s_{n-k}$ $(n \geq 2)$.
- $|\mathfrak{D}^1_{2n}(2413, 3142)| = C(2; n)$, the generalized Catalan number (see [**17**, A064062]).

Note that the these results parallel some enumerative avoidance results in S_n, where the same or similar cardinalities are obtained:

- $|S_n(\tau)| = C_n = \frac{1}{n+1}\binom{2n}{n}$, the nth Catalan number, for any $\tau \in S_3$.

- $|S_n(123,213)| = |S_n(132,231)| = 2^{n-1}$.
- $|S_n(3142,2413)| = |S_n(4132,4231)| = |S_n(2431,4231)| = r_{n-1}$, the $(n-1)$-st large Schröder number [**17**, A006318], given by $r_0 = 1$, $r_n = r_{n-1} + \sum_{j=0}^{n-1} r_k r_{n-k}$, or by $r_n = 2s_n$ for $n \geq 1$.

In this paper, we establish several enumerative and bijective results on restricted Dumont permutations.

In Section 2 we give direct bijections between $\mathfrak{D}^1_{2n}(132)$, $\mathfrak{D}^1_{2n}(231)$, $\mathfrak{D}^2_{2n}(321)$ and the class of Dyck paths of semilength n (paths from $(0,0)$ to $(2n,0)$ with steps $\mathbf{u} = (1,1)$ and $\mathbf{d} = (1,-1)$ that never go below the x-axis). This allows us to consider some permutation statistics, such as length of the longest increasing (or decreasing) subsequence, and study their distribution on the sets $\mathfrak{D}^1_{2n}(132)$, $\mathfrak{D}^1_{2n}(231)$ and $\mathfrak{D}^2_{2n}(321)$.

In Section 3, we consider Dumont permutations of the second kind avoiding patterns in $\mathfrak{D}^2_4$. Note that [**3**] showed that $|\mathfrak{D}^2_{2n}(3142)| = C_n$ using block decomposition (see [**12**]), which is very surprising given that it is by far a more difficult task to count all permutations avoiding a single 4-letter pattern (e.g., see [**2, 7, 18, 19, 21**]).

Furthermore, we prove that $\mathfrak{D}^2_{2n}(4132) = \mathfrak{D}^2_{2n}(321)$ and, thus, $|\mathfrak{D}^2_{2n}(4132)| = C_n$. The fact that permutations of different lengths are equinumerously avoided is another striking difference between restricted Dumont permutations and restricted permutations.

Refining the result $|\mathfrak{D}^2_{2n}(3142)| = C_n$ in [**3**], we consider some combinatorial statistics on $\mathfrak{D}^2_{2n}(3142)$ such as the number of fixed points and 2-cycles, and give a natural bijection between permutations in $\mathfrak{D}^2_{2n}(3142)$ with k fixed points and the set $NC(n, n-k)$ of noncrossing partitions of $[n]$ into $n-k$ parts that uses cycle decomposition. This is yet another surprising difference since pattern avoidance on permutations so far has not been shown to be related to their cycle decomposition in any natural way.

Finally, we prove that $|\mathfrak{D}^2_{2n}(2143)| = a_n a_{n+1}$, where $a_{2m} = \frac{1}{2m+1}\binom{3m}{m}$ and $a_{2m+1} = \frac{1}{2m+1}\binom{3m+1}{m+1}$. This allows us to relate 2143-avoiding Dumont permutations of the second kind with pairs of northeast lattice paths from $(0,0)$ to $(2n,n)$ and $(2n+1,n)$ that do not get above the line $y = x/2$.

Thus, we complete the enumeration problem of $\mathfrak{D}^2_{2n}(\tau)$ for all $\tau \in \mathfrak{D}^2_4$.

2. Dumont permutations avoiding a single 3-letter pattern

In this section we consider some permutation statistics and study their distribution on certain classes of restricted Dumont permutations. We focus on the sets $\mathfrak{D}^1_{2n}(132)$, $\mathfrak{D}^1_{2n}(231)$ and $\mathfrak{D}^2_{2n}(321)$, whose cardinality is given by the Catalan numbers, as shown in [**3, 11**]. We construct direct bijections between these sets and the class of Dyck paths of semilength n, which we denote $\mathcal{D}_n$.

2.1. 132**-avoiding Dumont permutations of the first kind.** Here we present a bijection f_1 between $\mathfrak{D}^1_{2n}(132)$ and $S_n(132)$, which will allow us to enumerate 132-avoiding Dumont permutations of the first kind with respect to the length of the longest increasing subsequences. The bijection is defined as follows. Let $\pi = \pi_1\pi_2\cdots\pi_{2n} \in \mathfrak{D}^1_{2n}(132)$. First delete all the even entries of π. Next, replace each of the remaining entries π_i by $(\pi_i+1)/2$. Note that we only obtain integer numbers since the π_i that were not erased are odd. Clearly, since π was 132-avoiding, the sequence $f_1(\pi)$ that we obtain is a 132-avoiding permutation, that is, $f_1(\pi) \in S_n(132)$. For example, if $\pi = 64357821$, then deleting the even entries we get 3571, so $f_1(\pi) = 2341$.

To see that f_1 is indeed a bijection, we now describe the inverse map. Let $\sigma \in S_n(132)$. First replace each entry σ_i with $\sigma'_i := 2\sigma_i - 1$. Now, for every i from 1 to n, proceed according to one of the two following cases. If $\sigma'_i > \sigma'_{i+1}$, insert $\sigma'_i + 1$ immediately to the right of σ'_i. Otherwise (that is, $\sigma'_i < \sigma'_{i+1}$ or σ'_{i+1} is not defined), insert $\sigma'_i + 1$ immediately to the right of the rightmost element to the left of σ'_i that is bigger than σ'_i, or to the beginning of the sequence if such element does not exist. For example, if $\sigma = 546231$, after the first step we get $(9,7,11,3,5,1)$, so $f_1^{-1}(\sigma) = (9,10,8,7,11,12,4,3,5,6,2,1)$.

Recall Krattenthaler's bijection between 132-avoiding permutations and Dyck paths [**10**]. We denote it by $\varphi : S_n(132) \to \mathcal{D}_n$, and it can be defined as follows. Given a permutation $\pi \in S_n(132)$ represented as an $n \times n$ board, where for each entry $\pi(i)$ there is a dot in the i-th column from the left and row $\pi(i)$ from the bottom, consider a lattice path from $(n,0)$ to $(0,n)$ not above the antidiagonal $y = n - x$ that leaves all dots to the right and stays as close to the antidiagonal as possible. Then $\varphi(\pi)$ is the Dyck path obtained from this path by reading an $\mathbf{u}$ every time the path goes west and a $\mathbf{d}$ every time it goes north. Composing f_1 with the bijection φ we obtain a bijection $\varphi \circ f_1 : \mathfrak{D}^1_{2n}(132) \to \mathcal{D}_n$.

Again through φ, the set $S_{2n}(132)$ is in bijection with $\mathcal{D}_{2n}$. Considering $\mathfrak{D}^1_{2n}(132)$ as a subset of $S_{2n}(132)$, we observe that $g_1 := \varphi \circ f_1^{-1} \circ \varphi^{-1}$ is an injective map from $\mathcal{D}_n$ to $\mathcal{D}_{2n}$. Here is a way to describe it directly only in terms of Dyck paths. Recall that a valley in a Dyck path is an occurrence of $\mathbf{du}$, and that a tunnel is a horizontal segment whose interior is below the path and whose endpoints are lattice points belonging to the path (see [**5, 6**] for more precise definitions). Let $D \in \mathcal{D}_n$. For each valley in D, consider the tunnel whose left endpoint is at the bottom of the valley. Mark the up-step and the down-step that delimit this tunnel. Now, replace each unmarked down-step **d** with **dud**. Replace each marked up-step **u** with **uu**, and each marked **d** with **dd**. The path that we obtain after these operations is precisely $g_1(D) \in \mathcal{D}_{2n}$. The reason is that through φ, each entry of the permutation has an associated tunnel in the path (as described in [**5**]), and these operations on the steps of the path create tunnels that correspond to the even elements of $f_1^{-1}(\varphi^{-1}(D))$.

For example, if $D =$ **uduudud**, then underlining the marked steps we get $\mathbf{ud\underline{u}ud\underline{udd}}$, so $g_1(D) =$ **ududuuududuudddd**.

Denote by $\mathrm{lis}(\pi)$ (resp. $\mathrm{lds}(\pi)$) the length of the longest increasing (resp. decreasing) subsequence of π. Using the above bijections we obtain the following result.

THEOREM 2.1. *Let $L_k(z) := \sum_{n\geq 0} |\{\pi \in \mathfrak{D}^1_{2n}(132) : \mathrm{lis}(\pi) \leq k)\}| \, z^n$ be the generating function for $\{132, 12\cdots(k+1)\}$-avoiding Dumont permutations of the first kind. Then we have the recurrence*

$$L_k(z) = 1 + \frac{zL_{k-1}(z)}{1 - zL_{k-2}(z)},$$

with $L_{-1}(z) = 0$ and $L_0(z) = 1$.

PROOF. As shown in [**10**], the length of the longest increasing subsequence of a permutation $\pi \in S_{2n}(132)$ corresponds to the height of the path $\varphi(\pi) \in \mathcal{D}_{2n}$. Next we describe the statistic, which we denote λ, on the set of Dyck paths $\mathcal{D}_n$ that, under the injection $g_1 : \mathcal{D}_n \hookrightarrow \mathcal{D}_{2n}$, corresponds to the height in $\mathcal{D}_{2n}$. Let $D \in \mathcal{D}_n$. For each peak p of D, define $\lambda(p)$ to be the height of p plus the number of tunnels below p whose left endpoint is at a valley of D. Now let $\lambda(D) := \max_p\{\lambda(p)\}$ where p ranges over all the peaks of D. From the description of g_1 it follows that for any $D \in \mathcal{D}_n$, $\mathrm{height}(g_1(D)) = \lambda(D)$. Thus, enumerating permutations in $\mathfrak{D}^1_{2n}(132)$ according to the parameter lis is equivalent to enumerating paths in $\mathcal{D}_n$ according to the parameter λ. More precisely, $L_k(z) = \sum_{D\in\mathcal{D}:\lambda(D)\leq k} z^{|D|}$. To find an equation for L_k, we use that every nonempty Dyck path D can be uniquely decomposed as $D = A\mathbf{u}B\mathbf{d}$, where $A, B \in \mathcal{D}$. We obtain that

$$L_k(z) = 1 + zL_{k-1}(z) + z(L_k(z) - 1)L_{k-2}(z),$$

where the term $zL_{k-1}(z)$ corresponds to the case where A is empty (for then $\lambda(\mathbf{u}B\mathbf{d}) = \lambda(B) + 1$, and $z(L_k(z) - 1)L_{k-2}(z)$ to the case there A is not empty. From this we obtain the recurrence

$$L_k(z) = 1 + \frac{zL_{k-1}(z)}{1 - zL_{k-2}(z)},$$

where $L_{-1}(z) = 0$ and $L_0(z) = 1$ by definition. □

It also follows from the definition of φ that the length of the longest decreasing subsequence of $\pi \in S_{2n}(132)$ corresponds to the number of peaks of the path $\varphi(\pi) \in \mathcal{D}_{2n}$. Looking at the description of g_1, we see that a peak is created in $g_1(D)$ for each unmarked down-step of d. The number of marked down-steps is the number of valleys of D. Therefore, if $D \in \mathcal{D}_n$, we have that the number of peaks of $g_1(D)$ is $\mathrm{peaks}(g_1(D)) = \mathrm{peaks}(D) + n - \mathrm{valleys}(D) = n + 1$. Hence, we have that for every $\pi \in \mathfrak{D}^1_{2n}(132)$, $\mathrm{lds}(\pi) = n + 1$.

2.2. 231-avoiding Dumont permutations of the first kind. As we did in the case of 132-avoiding Dumont permutations, we can give the following bijection f_2 between $\mathfrak{D}^1_{2n}(231)$ and $S_n(231)$. Let $\pi \in \mathfrak{D}^1_{2n}(231)$. First delete all the odd entries of π. Next, replace each of the remaining entries π_i by $\pi_i/2$. Note that we only obtain integer entries since the remaining π_i were even. Compare this to the analogous transformation described in Section 3.1 for Dumont permutations of the second kind. Clearly the sequence $f_2(\pi)$ that we obtain is a 231-avoiding permutation (since so was π), that is, $f_2(\pi) \in S_n(231)$. For example, if $\pi = (2, 1, 10, 8, 4, 3, 6, 5, 7, 9)$, then deleting the odd entries we get $(2, 10, 8, 4, 6)$, so $f_2(\pi) = 15423$.

To see that f_2 is indeed a bijection, we define the inverse map as follows. Let $\sigma \in S_n(231)$. First replace each entry k with $2k$. Now, for every i from 1 to $n - 1$, insert $2i - 1$ immediately to the left of the first entry to the right of $2i$ that is bigger than $2i$ (if such an entry does not exist, insert $2i - 1$ at the end of the sequence). For example, if $\sigma = 7215346$, after the first step we get $(14, 4, 2, 10, 6, 8, 12)$, so $f_2^{-1}(\sigma) = (14, 4, 2, 1, 3, 10, 6, 5, 8, 7, 9, 12, 11, 13)$.

Consider now the bijection $\varphi^R : S_n(231) \longrightarrow \mathcal{D}_n$ that is obtained by composing φ defined above with the reversal operation that sends $\pi = \pi_1\pi_2 \cdots \pi_n \in S_n(231)$ to $\pi^R = \pi_n \cdots \pi_2\pi_1 \in S_n(132)$.

Through φ^R, the set $S_{2n}(231)$ is in bijection with $\mathcal{D}_{2n}$, so we can identify $\mathfrak{D}^1_{2n}(231)$ with a subset of $\mathcal{D}_{2n}$. The map $g_2 := \varphi^R \circ f_2^{-1} \circ (\varphi^R)^{-1}$ is an injection from $\mathcal{D}_n$ to $\mathcal{D}_{2n}$. Here is a way to describe it directly only in terms of Dyck paths. Given $D \in \mathcal{D}_n$, all we have to do is replace each down-step **d** of D with **udd**. The path that we obtain is precisely $g_2(D) \in \mathcal{D}_{2n}$. For example, if $D =$ **uduuududdd** (this example corresponds to the same σ given above), then $g_2(D) =$ **uudduuuudduudduddudd**. Given $g_2(D)$, one can easily recover D by replacing every **udd** by **d**.

Some properties of φ trivially translate to properties of φ^R. In particular, the length of the longest increasing subsequence of a 231-avoiding permutation π equals the number of peaks of $\varphi^R(\pi)$, and the length of the longest decreasing subsequence of π is precisely the height of $\varphi^R(\pi)$.

It follows from the description of g_2 in terms of Dyck paths that for any $D \in \mathcal{D}_n$, $g_2(D)$ has exactly n peaks (one for each down-step of D). Therefore, for any $\pi \in \mathfrak{D}^1_{2n}(231)$, the number of right-to-left minima of π is $\mathrm{rlm}(\pi) = n$. In fact it is not hard to see directly from the definition of 231-avoiding Dumont permutations that the right-to-left minima of $\pi \in \mathfrak{D}^1_{2n}(231)$ are precisely its odd entries, which necessarily form an increasing subsequence.

Also from the description of g_2 we see that $\mathrm{height}(g_2(D)) = \mathrm{height}(D) + 1$. In terms of permutations, this translates to the fact that if $\pi \in S_n(231)$, then $\mathrm{lds}(f_2(\pi)) = \mathrm{lds}(\pi) + 1$. This allows us to enumerate 231-avoiding Dumont permutations with respect to the statistic lds. Indeed, $|\{\pi \in \mathfrak{D}^1_{2n}(231) : \mathrm{lds}(\pi) = k\}| = |\{D \in \mathcal{D}_n : \mathrm{height}(D) = k - 1\}|$.

2.3. 321**-avoiding Dumont permutations of the second kind.** Let us first notice that a permutation $\pi \in \mathfrak{D}^2_{2n}(321)$ cannot have any fixed points. Indeed, assume that $\pi_i = i$. Then, if we write $\pi = \sigma i \tau$, the fact that π is 321-avoiding implies that σ is a permutation of $\{1, 2, \ldots, i-1\}$ and τ is a permutation of $\{i+1, i+2, \ldots, n\}$. Since $\pi \in \mathfrak{D}^2_{2n}$, i must be odd, but then the first element of τ is in an even position, and it is either a fixed point or an excedance, which contradicts the definition of Dumont permutations of the second kind.

It is known (see e.g. [**14**]) that a permutation is 321-avoiding if and only if both the subsequence determined by its excedances and the one determined by the remaining elements are increasing. It follows that a permutation in $\mathfrak{D}^2_{2n}(321)$ is uniquely determined by the values of its excedances. Another consequence is that if $\pi \in \mathfrak{D}^2_{2n}(321)$, then $\mathrm{lis}(\pi) = n$.

We can give a bijection between $\mathfrak{D}^2_{2n}(321)$ and $\mathcal{D}_n$. We define it in two parts. For the first part, we use the bijection ψ between $S_n(321)$ and $\mathcal{D}_n$ that was defined in [**5**], and which is closely related to the bijection between $S_n(123)$ and $\mathcal{D}_n$ given in [**10**]. Given $\pi \in S_n(321)$, consider again the $n \times n$ board with a dot in the i-th column from the left and row $\pi(i)$ from the bottom, for each i. Take the path with *north* and *east* steps that goes from $(0, 0)$ to the (n, n), leaving all the dots to the right, and staying always as close to the diagonal as possible. Then $\psi(\pi)$ is the Dyck path obtained from this path by reading an up-step every time the path goes north and a down-step every time it goes east.

If we apply ψ to a permutation $\pi \in \mathfrak{D}^2_{2n}(321)$ we get a Dyck path $\psi(\pi) \in \mathcal{D}_{2n}$. The second part of our bijection is just the map g_2^{-1} defined above, which consists in replacing every occurrence of **udd** with a **d**. It is not hard to check that $\pi \mapsto g_2^{-1}(\psi(\pi))$ is a bijection from $\mathfrak{D}^2_{2n}(321)$ to $\mathcal{D}_n$. For example, for $\pi = (3, 1, 5, 2, 6, 4, 9, 7, 10, 8)$, we have that $\psi(\pi) =$ **uuudduudduddduuuddudd**, and $g_2^{-1}(\psi(\pi)) =$ **uududduudd**.

3. Dumont permutations avoiding a single 4-letter pattern

In this section we will determine the structure of permutations in $\mathfrak{D}^2_{2n}(\tau)$ and find the cardinality $|\mathfrak{D}^2_{2n}(\tau)|$ for each $\tau \in \mathfrak{D}^2_4 = \{2143, 3142, 4132\}$.

It was shown in [**3**] that $|\mathfrak{D}^2_{2n}(3142)| = C_n$. In Section 3.1, we refine this result with respect to the number of fixed points and 2-cycles in permutations in $\mathfrak{D}^2_{2n}(3142)$ and use cycle decomposition to give a natural bijection between permutations in $\mathfrak{D}^2_{2n}(3142)$ with k fixed points and the set $NC(n, n-k)$ of noncrossing partitions of $[n]$ into $n-k$ parts. In Section 3.2, we prove that $\mathfrak{D}^2_{2n}(4132) = \mathfrak{D}^2_{2n}(321)$ and, thus,

$|\mathfrak{D}^2_{2n}(4132)| = C_n$. Finally, in Section 3.3 we prove that $|\mathfrak{D}^2_{2n}(2143)| = a_n a_{n+1}$, where $a_{2m} = \frac{1}{2m+1}\binom{3m}{m}$ and $a_{2m+1} = \frac{1}{2m+1}\binom{3m+1}{m+1}$. Thus, we can relate permutations in $\mathfrak{D}^2_{2n}(2143)$ and pairs of northeast lattice paths from $(0,0)$ to $(2n,n)$ and $(2n+1,n)$ that stay on or below $y = x/2$. This completes the enumeration problem of $\mathfrak{D}^2_{2n}(\tau)$ for $\tau \in \mathfrak{D}^2_4$.

3.1. Avoiding 3142. It was shown in [**3**] that $|\mathfrak{D}^2_{2n}(3142)| = C_n$; moreover, the permutations $\pi \in \mathfrak{D}^2_{2n}(3142)$ can be recursively described as follows:

$$\pi = (2k, 1, r \circ c(\pi') + 1, \pi'' + 2k), \tag{3.1}$$

where $\pi' \in \mathfrak{D}^2_{2k-2}(3142)$ and $\pi'' \in \mathfrak{D}^2_{2n-2k}(3142)$ (see Figure 1). From this block decomposition, it is easy to see that the subsequence of odd integers in π is increasing. Moreover, the odd entries are exactly those on the main diagonal and the first subdiagonal (i.e. those i for which $\pi(i) = i$ or $\pi(i) = i - 1$).

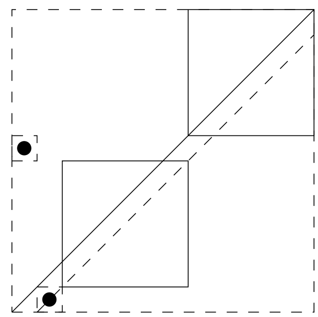


FIGURE 1. The block decomposition of a permutation in $D^2_{2n}(3142)$.

In subsections 3.1.1 and 3.1.2 we use the above decomposition to derive two bijections from $\mathfrak{D}^2_{2n}(3142)$ to sets of cardinality C_n.

3.1.1. *Subsequence of even entries.* The first bijection is $\phi : \mathfrak{D}^2_{2n}(3142) \to E_n \subset S_n$, where

$$E_n = \left\{(1/2)\pi_{ev} \mid \pi \in \mathfrak{D}^2_{2n}(3142)\right\},$$

and π_{ev} (resp. π_{ov}) is the subsequence of even (resp. odd) values in π. (Here $\frac{1}{2}\pi_{ev}$ is the permutation obtained by dividing all entries in π_{ev} by 2; in other words, if $\sigma = \frac{1}{2}\pi_{ev}$, then $\sigma(i) = \pi_{ev}(i)/2$ for all $i \in [n]$.) Define $\phi(\pi) = \frac{1}{2}\pi_{ev}$ for each $\pi \in \mathfrak{D}^2_{2n}(3142)$.

Permutations in E_n have a block decomposition similar to those in $\mathfrak{D}^2_{2n}(3142)$, namely,

$$\sigma \in E_n \iff \sigma = (k, r \circ c(\sigma'), k + \sigma'') \text{ for some } \sigma' \in E_{k-1} \text{ and } \sigma'' \in E_{n-k}.$$

The inverse $\phi^{-1} : E_n \to \mathfrak{D}^2_{2n}(3142)$ is easy to describe. Let $\sigma \in E_n$. Then $\pi = \phi^{-1}(\sigma)$ is obtained as follows: let $\pi_{ev} = 2\sigma$ (i.e. $\pi_{ev}(i) = 2\sigma(i)$ for all $i \in [n]$), then for each $i \in [n]$ insert $2i-1$ immediately before $2\sigma(i)$ if $\sigma(i) < i$ or immediately after $2\sigma(i)$ if $\sigma(i) \geq i$. For instance, if $\sigma = 3124 \in E_4$, then $\pi_{ev} = 6248$ and $\pi = 61\,32\,54\,87 \in \mathfrak{D}^2_8(3142)$.

It is not difficult to show that E_n consists of exactly those permutations that, written in cyclic form, correspond to noncrossing partitions of $[n]$ by replacing pairs of parentheses with slashes. We remark that E_n is also the set of permutations whose tableaux (see [**20**]) have a single 1 in each column.

THEOREM 3.1. *For a permutation ρ, define*

$$\mathrm{fix}(\rho) = |\{i \mid \rho(i) = i\}|, \qquad \mathrm{exc}(\rho) = |\{i \mid \rho(i) > i\}|,$$
$$\mathrm{fix}_{-1}(\rho) = |\{i \mid \rho(i) = i-1\}|, \qquad \mathrm{def}(\rho) = |\{i \mid \rho(i) < i\}|.$$

Then for any $\pi \in \mathfrak{D}^2_{2n}(3142)$ and $\sigma = \phi(\pi) \in E_n$, we have

$$\mathrm{fix}(\pi) + \mathrm{fix}_{-1}(\pi) = n, \tag{3.2}$$

$$\mathrm{fix}(\pi) = \mathrm{def}(\sigma), \tag{3.3}$$

$$\mathrm{fix}_{-1}(\pi) = \mathrm{exc}(\sigma) + \mathrm{fix}(\sigma), \tag{3.4}$$

$$\mathrm{fix}(\sigma) = \#\text{ 2-cycles in } \pi. \tag{3.5}$$

Proof. Equation (3.2) follows from the fact that odd integers in π are exactly those on the main diagonal and first subdiagonal.

Let π and σ be as above and let $i \in [n]$. Then there are two cases: either $2i-1 = \pi(2i)$ or $2i-1 = \pi(2i-1)$.

Case 1: $\pi(2i) = 2i-1$. Then $\pi(2i-1) \geq 2i$, and hence $\pi(2i-1)$ must be even.

Case 2: $\pi(2i-1) = 2i-1$. Then $\pi(2i) \leq 2i-2$, and hence $\pi(2i)$ must be even.

In either case, for each $i \in [n]$, we have $\{\pi(2i-1), \pi(2i)\} = \{2i-1, 2s_i\}$ for some $s_i \in [n]$. Define $\sigma(i) = s_i$. Then $\sigma(i) \geq i$ if $2i-1 \in \mathrm{fix}_{-1}(\pi)$, and $\sigma(i) \leq i-1$ if $2i-1 \in \mathrm{fix}(\pi)$. This proves (3.3) and (3.4).

Finally, let $i \in [n]$ be such that $\sigma(i) = i$. Since $2\sigma(i) \in \{\pi(2i-1), \pi(2i)\}$ and $\pi(2i) < 2i$, it follows that $2i = 2\sigma(i) = \pi(2i-1)$, so $2i-1 = \pi(2i)$, and thus π contains a 2-cycle $(2i-1, 2i)$.

Conversely, let (ab) be a 2-cycle of π, and assume that $b > a$. Then $\pi(a) > a$, so a must be odd, say $a = 2i-1$ for some $i \in [n]$. Then $b = \pi^{-1}(a) \in \{2i-1, 2i\}$, so $b = 2i$, and thus $(ab) = (2i-1, 2i)$. This proves (3.5). $\square$

Theorem 3.2. *Let $A(q,t,x) = \sum_{n\geq 0} \sum_{\pi \in \mathfrak{D}^2_{2n}(3142)} q^{\mathrm{fix}(\pi)} t^{\#\text{ 2-cycles in }\pi} x^n$ be the generating function for 3142-avoiding Dumont permutations of the second kind with respect to the number of fixed points and the number of 2-cycles. Then*

$$A(q,t,x) = \frac{1 + x(q-t) - \sqrt{1 - 2x(q+t) + x^2((q+t)^2 - 4q)}}{2xq(1 + x(1-t))}. \tag{3.6}$$

Proof. By the correspondences in Theorem 3.1, it follows that

$$A(q,t,x) = \sum_{n\geq 0} \sum_{\sigma \in E_n} q^{\mathrm{def}(\sigma)} t^{\mathrm{fix}(\sigma)} x^n.$$

For convenience, let us define a related generating function $B(q,t,x) = \sum_{n\geq 0} \sum_{\sigma \in E_n} q^{\mathrm{def}(\sigma)} t^{\mathrm{fix}_{-1}(\sigma)} x^n$. From the block decomposition of permutations $\sigma \in E_n$ as $\sigma = (k, r \circ c(\sigma'), k + \sigma'')$ for some $\sigma' \in E_{k-1}$, $\sigma'' \in E_{n-k}$, it follows that

$$A(q,t,x) = 1 + xtA(q,t,x) + x(B(1/q,t,xq) - 1)A(q,t,x). \tag{3.7}$$

The term $xtA(q,t,x)$ corresponds to the case $k=1$, in which σ' is empty and k is a fixed point. When $k > 1$, σ'' still contributes as $A(q,t,x)$, and the contribution of σ' is $B(1/q,t,xq) - 1$, since elements with $\sigma'(i) = i-1$ become fixed points of σ, and all elements of σ' other than its deficiencies become deficiencies of σ.

A similar reasoning gives the following equation for $B(q,t,x)$:

$$B(q,t,x) = 1 + xA(1/q,t,xq)B(q,t,x).$$

Solving for B we have $B(q,t,x) = \frac{1}{1 - xA(1/q,t,xq)}$, and plugging $B(1/q,t,xq) = \frac{1}{1 - xqA(q,t,x)}$ into (3.7) gives

$$A(q,t,x) = 1 + x\left(\frac{1}{1 - xqA(q,t,x)} + t - 1\right)A(q,t,x).$$

Solving this quadratic equation gives the desired formula for $A(q,t,x)$. $\square$

3.1.2. *Cycle decomposition.* Letting $t = 1$ in (3.6), we obtain

Corollary 3.3. We have

$$\sum_{n\geq 0} \sum_{\pi \in \mathfrak{D}^2_{2n}(3142)} q^{\mathrm{fix}(\pi)} x^n = A(q,1,x) = \frac{1 + x(q-1) - \sqrt{1 - 2x(q+1) + x^2(q-1)^2}}{2xq},$$

i.e. the number of permutations in $\pi \in \mathfrak{D}^2_{2n}(3142)$ with k fixed points is the Narayana number $N(n,k) = \frac{1}{n}\binom{n}{k}\binom{n}{k+1}$, which is also the number of noncrossing partitions of $[n]$ into $n-k$ parts.

Proof. Even though the generating function is an immediate consequence of Theorem 3.2, we will give a combinatorial proof of the corollary, by exhibiting a natural bijection $\psi : \mathfrak{D}^2_{2n}(3142) \to NC(n)$, where $NC(n)$ is the set of noncrossing partitions of $[n]$. We start by considering a permutation $\pi \in \mathfrak{D}^2_{2k}(3142)$. Iterating the block decomposition (3.1), we obtain

$$\begin{aligned}\pi &= (2k_1, 1, c \circ r(\pi_1) + 1, 2k_2, 2k_1 + 1, c \circ r(\pi_2) + 2k_1 + 1, \cdots, 2k_r, 2k_{r-1} + 1, c \circ r(\pi_r) + 2k_{r-1} + 1) \\ &= (2k_1, 1, 2k_1 - r(\pi_1), 2k_2, 2k_1 + 1, 2k_2 - r(\pi_2), \cdots, 2k_r, 2k_{r-1} + 1, 2k_r - r(\pi_r)),\end{aligned}$$

where $1 \le k_1 < k_2 < \cdots < k_r = k$, $\pi_i \in \mathfrak{D}^2_{2(k_i-k_{i-1}-1)}(3142)$ $(1 \le i \le r)$, and we define $k_0 = 0$. Note that each permutation $c \circ r(\pi_i) + 2k_{i-1} + 1 = 2k_i - r(\pi_i)$ of $[2k_{i-1}+2, 2k_i - 1]$ occurs at positions $[2k_{i-1}+3, 2k_i]$ in π.

Now consider

$$\pi' = (2k+2, 1, c \circ r(\pi) + 1) = (2k_r + 2, 1, 2k_r + 2 - r(\pi)).$$

Let $k'_i = k - k_i = k_r - k_i$. By (3.1), we have $\pi \in \mathfrak{D}^2_{2k+2}(3142)$, $\pi_i \in \mathfrak{D}^2_{2(k'_{i-1}-k'_i-1)}(3142)$ $(1 \le i \le r)$, $k'_r = 0$, $k'_0 = k$, and

$$\pi' = (2k+2, 1, \pi_r + 2, 2k'_{r-1}+1, 2, \pi_{r-1} + 2k'_{r-1} + 2, 2k'_{r-2}+1, 2k'_{r-1}+2, \ldots, \pi_1 + 2k'_1 + 2, 2k+1, 2k'_1 + 2).$$

Note that, for each $i = 1, 2, \ldots, r$, the permutation $\pi_i + 2k'_i + 2$ of $[2k'_i + 3, 2k'_{i-1}]$ occurs at positions $[2k'_i + 3, 2k'_{i-1}]$ in π'. Moreover, the entries $2k'_i + 1$ $(0 \le i \le r-1)$ occur at positions $2k'_i + 1$ in π', and thus are fixed points of π'. Finally, each entry $2k'_i + 2$ $(1 \le i \le r)$ occurs at position $2k'_{i-1} + 2$, 1 occurs at position $2 = 2k'_r + 2$, and $2k+2 = 2k'_0 + 2$ occurs at position 1.

Thus, $\gamma = (2k'_0+2, 2k'_1+2, 2k'_2+2, \ldots, 2k'_{r-1}+2, 2k'_r+2, 1) = (2k+2, 2k'_1+2, 2k'_2+2, \ldots, 2k'_{r-1}+2, 2, 1)$ is a cycle of π', and each remaining nontrivial cycle of π' is completely contained in some $\pi_i + 2k'_i + 2$, which is a 3142-avoiding Dumont permutation of the second kind of $[2k'_i + 3, 2k'_{i-1}]$. Note that

$$2k'_i + 2 < 2k'_i + 3 < 2k'_{i-1} < 2k'_{i-1} + 2,$$

so all entries of each remaining cycle of π' are contained between two consecutive entries of γ.

Now let G be the subset of $[2k+2]$ consisting of the entries of γ. Then, clearly,

$$G/\{2k'_{r-1}+1\}/\ldots/\{2k'_1+1\}/\{2k'_0+1\}/[2k'_r+3, 2k'_{r-1}]/\ldots/[2k'_1+3, 2k'_0]$$

is a noncrossing partition of $[2k+2]$. Now it is easy to see by induction on the size of π' that the subsets of π' formed by entries of the cycles in cycle decomposition of π' form a noncrossing partition of π'. Moreover, all the entries of G except the smallest entry are even, so likewise the cycle decomposition of π' determines a unique noncrossing partition of π'_{ev}, hence a unique noncrossing partition of $[n]$.

Finally, any permutation $\hat{\pi} \in \mathfrak{D}^2_{2n}(3142)$ can be written as $\hat{\pi} = (\pi', \pi'' + 2k + 2)$, where π' is as above and $\pi'' \in \mathfrak{D}^2_{2n-2k-2}(3142)$, so the cycles of any permutation in $\mathfrak{D}^2_{2n}(3142)$ determine a unique noncrossing partition of $[n]$.

Notice also that each cycle in the decomposition of $\hat{\pi}$ contains exactly one odd entry, the least entry in each cycle, so the number of odd entries of $\hat{\pi}$ which are not fixed points, $\mathrm{fix}_{-1}(\hat{\pi}) = n - \mathrm{fix}(\hat{\pi})$, is the number of parts in $\psi(\hat{\pi})$. This finishes the proof. ☐

For example, if

$$\begin{aligned}\hat{\pi} &= 12, 1, 6, 3, 5, 4, 7, 2, 10, 9, 11, 8, 16, 13, 15, 14\\ &= (12, 8, 2, 1)(6, 4, 3)(10, 9)(16, 14, 13)(15)(11)(7)(5) \in \mathfrak{D}^2_{16}(3142),\end{aligned}$$

then $\psi(\hat{\pi}) = 641/32/5/87 \in NC(8)$. Note also that $\hat{\pi}_{ev} = 63215487 = (641)(32)(5)(87)$.

3.2. Avoiding 4132. For Dumont permutations of the second kind avoiding the pattern 4132 we have the following result.

Theorem 3.4. *For any $n \ge 0$, $\mathfrak{D}^2_{2n}(4132) = \mathfrak{D}^2_{2n}(321)$. Moreover, $|\mathfrak{D}^2_{2n}(4132)| = C_n$, where C_n is the nth Catalan number. Thus,* 4132 *and* 3142 *are $\mathfrak{D}^2$-Wilf-equivalent.*

Proof. The pattern 321 is contained in 4132. Therefore, if π avoids 321, then π avoids 4132, so $\mathfrak{D}^2_{2n}(321) \subseteq \mathfrak{D}^2_{2n}(4132)$. Now let us prove that $\mathfrak{D}^2_{2n}(4132) \subseteq \mathfrak{D}^2_{2n}(321)$. Let $n \ge 4$ and let $\pi \in \mathfrak{D}^2_{2n}(4132)$ contain an occurrence of 321. Choose the leftmost occurrence of 321 in π, namely, $\pi(i_1) > \pi(i_2) > \pi(i_3)$ with $1 \le i_1 < i_2 < i_3 \le 2n$ such that $i_1+i_2+i_3$ is minimal. If i_1 is an even number, then $\pi(i_1-1) \ge i_1 - 1 \ge \pi(i_1)$, so the occurrence $\pi(i_1-1)\pi(i_1)\pi(i_2)$ of pattern 321 contradicts minimality of our choice. Therefore, i_1 is odd. If $i_2 \ne i_1 + 1$, then from the minimality of the occurrence we get that $\pi(i_1+1) < \pi(i_3)$. Hence, π contains 4132 a contradiction. So $i_2 = i_1 + 1$. If i_3 is odd, then $\pi(i_3) \ge i_3 > i_1 + 1 \ge \pi(i_1+1)$, which contradicts the fact that $\pi(i_1) > \pi(i_1+1) > \pi(i_3)$. So i_3 is even.

Therefore, the leftmost occurrence of 321 is given by $\pi(2i+1)\pi(2i+2)\pi(j)$ where $4 \le 2i+2 \le j \le 2n$ (since $\pi(2) = 1$, we must have $i \ge 1$). By minimality of the occurrence, we have $\pi(m) \le 2i$ for all $m \le 2i$.

On the other hand, $\pi(i_3) < \pi(2i+2) \leq 2i+1$ which means that $\pi(i_3) \leq 2i$. Hence, π must contain at least $2i+1$ letters smaller than $2i$, a contradiction.

Therefore, if $\pi \in \mathfrak{D}^2_{2n}(4132)$ then $\pi \in \mathfrak{D}^2_{2n}(321)$. The rest is a consequence of [**11**, Theorem 4.3]. $\square$

3.3. Avoiding 2143. Dumont permutations of the second kind that avoid 2143 are enumerated by the following theorem, which we prove in this section.

THEOREM 3.5. *For any $n \geq 0$, $|\mathfrak{D}^2_{2n}(2143)| = a_n a_{n+1}$, where*

$$a_{2m} = \frac{1}{2m+1}\binom{3m}{m},$$
$$a_{2m+1} = \frac{1}{2m+1}\binom{3m+1}{m+1} = \frac{1}{m+1}\binom{3m+1}{m}.$$

REMARK 3.6. Note that the sequence $\{a_n\}$ enumerates, among other objects, pairs of northeast lattice paths from $(0,0)$ to $(n, \lfloor n/2 \rfloor)$ that do not get above the line $y = x/2$ (see [**17**, A047749] and references therein). Also note that $\{a_{2m+1}\}$ is the convolution of $\{a_{2m}\}$ with itself, while the convolution of $\{a_{2m}\}$ with $\{a_{2m+1}\}$ is $\{a_{2m+2}\}$. Alternatively, if $f(x)$ and $g(x)$ are the ordinary generating functions for $\{a_{2m}\}$ and $\{a_{2m+1}\}$, then $f(x) = 1 + xf(x)g(x)$ and $g(x) = f(x)^2$, so $f(x) = 1 + xf(x)^3$. Now the Lagrange inversion applied to the last two equations yields the formulas for a_n.

Note that Theorem 3.5 implies that $\lim_{n\to\infty} |\mathfrak{D}^2_{2n}(2143)|^{\frac{1}{2n}} = \frac{3^3}{2^2} = \frac{27}{4}$. In comparison, [**13**] and [**21**] imply that $|S_n(2143)| = |S_n(1234)|$ and hence $\lim_{n\to\infty} |S_n(2143)|^{\frac{1}{n}} = \lim_{n\to\infty} |S_n(1234)|^{\frac{1}{n}} = (4-1)^2 = 9$.

LEMMA 3.7. *Let $\pi \in \mathfrak{D}^2_{2n}(2143)$. Then the subsequence $(\pi(1), \pi(3), \ldots, \pi(2n-1))$ is a permutation of $\{n+1, n+2, \ldots, 2n\}$ and the subsequence $(\pi(2), \pi(4), \ldots, \pi(2n))$ is a permutation of $\{1, 2, \ldots, n\}$.*

PROOF. Assume the lemma is false. Let i be the smallest integer such that $\pi(2i) \geq n+1$. Then $\pi(2i-1) \geq 2i-1 \geq \pi(2i) \geq n+1$. Therefore, if $j \geq i$, then $\pi(2j-1) \geq 2j-1 \geq 2i-1 \geq n+1$. In fact, note that for any $1 \leq j \leq n$, $\pi(2j-1) \geq 2j-1 \geq \pi(2j)$.

By minimality of i, we have $\pi(2j) \leq n$ for $j < i$, so if $\pi(2j-1) \leq n$ for some $j < i$, then $(\pi(2j-1), \pi(2j), \pi(2i-1), \pi(2i))$ is an occurrence of pattern 2143 in π. Hence, $\pi(2j-1) \geq n+1$ for all $j < i$.

Thus, we have $\pi(2j-1) \geq n+1$ for any $1 \leq j \leq n$, and $\pi(2i) \geq n+1$, so π must have at least $n+1$ entries between $n+1$ and $2n$, which is impossible. The lemma follows. $\square$

For $\pi \in \mathfrak{D}^2_{2n}(2143)$, we denote $\pi_o = (\pi(1), \pi(3), \ldots, \pi(2n-1)) - n$ and $\pi_e = (\pi(2), \pi(4), \ldots, \pi(2n))$. By Lemma 3.7, $\pi_o, \pi_e \in S_n(2143)$. For example, given $\pi = 71635482 \in \mathfrak{D}^2_8(2143)$, we have $\pi_o = 3214$ and $\pi_e = 1342$. Note that $\pi(2i-1) = \pi_o(i) + n$ and $\pi(2i) = \pi_e(i)$.

LEMMA 3.8. *For any permutation $\pi \in \mathfrak{D}^2_{2n}(2143)$, and π_o and π_e defined as above, the following is true:*

(1) *$\pi_o \in S_n(132)$ and the entries of π_o are on a board with n columns aligned at the top of sizes $2, 4, 6, \ldots, 2\lfloor \frac{n}{2} \rfloor, n, \ldots, n$ from right to left (see the first and third boards in Figure 2).*

(2) *$\pi_e \in S_n(213)$ and the entries of π_e are on a board with n columns aligned at the bottom of sizes $1, 3, 5, \ldots, 2\lfloor \frac{n}{2} \rfloor - 1, n, \ldots, n$ from left to right (see the second and fourth boards in Figure 2).*

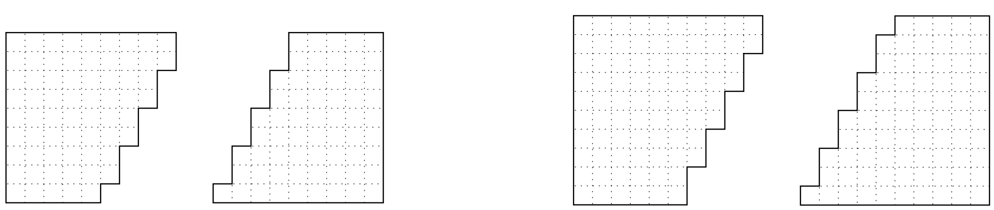


FIGURE 2. The boards of Lemma 3.8 for $n = 9$ (left) and $n = 10$ (right).

PROOF. If 132 occurs in π_o at positions $i_1 < i_2 < i_3$, then 2143 occurs in π at positions $2i_1 - 1 < 2i_1 < 2i_2 - 1 < 2i_3 - 1$ since $\pi(2i_1) < \pi(2i_1 - 1)$. Similarly, if 213 occurs in π_e at positions $i_1 < i_2 < i_3$, then 2143 occurs in π at positions $2i_1 < 2i_2 < 2i_3 - 1 < 2i_3$ since $\pi(2i_3 - 1) > \pi(2i_3)$. The rest simply follows from the definition of $\mathfrak{D}^2_{2n}$ and Lemma 3.7. $\square$

Let us call a permutation as in part (1) of Lemma 3.8 an *upper board*, and a permutation as in part (2) of Lemma 3.8 a *lower board*. Note that $\pi_e(1) = 1$ and $213 = r \circ c(132)$. Hence it is easy to see that $\pi_e = (1, r \circ c(\pi') + 1)$ with $\pi' \in S_{n-1}(132)$ of upper type. Let b_n be the number of lower boards in $S_n(213)$. Then the number of upper boards in $S_n(132)$ is b_{n+1}.

LEMMA 3.9. *Let $\pi_1 \in S_n(132)$ be an upper board and $\pi_2 \in S_n(213)$ be a lower board. Let $\pi \in S_{2n}$ be defined by $\pi = (\pi_1(1)+n, \pi_2(1), \pi_1(2)+n, \pi_2(2), \dots, \pi_1(n)+n, \pi_2(n))$ (i.e. such that $\pi_o = \pi_1$ and $\pi_e = \pi_2$). Then $\pi \in \mathfrak{D}^2_{2n}(2143)$.*

PROOF. Clearly $\pi \in \mathfrak{D}^2_{2n}$. It is not difficult to see that if π contains 2143, then "2" and "1" are deficiencies (i.e., they are at even positions and come from π_2) and "4" and "3" are excedances or fixed points (i.e. they are at odd positions and come from π_1). Such an occurrence is represented in Figure 3, where an entry $\pi(i)$ is plotted by a dot with abscissa i and ordinate $\pi(i)$, and the two diagonal lines indicate the positions of the fixed points and elements with $\pi(i) = i - 1$.

Say the pattern 2143 occurs at positions $2i_1 < 2i_2 < 2i_3 - 1 < 2i_4 - 1$. We have $\pi(2j) \le 2j - 1 < 2i_2 - 1$ for any $j < i_2$. On the other hand, the subdiagonal part of π avoids 213, so $\pi(2j) < \pi(2i_1) \le 2i_1 - 1 < 2i_2 - 1$ for any $j \ge i_2$. Thus, $\pi(2j) < 2i_2 - 1$ for any $1 \le j \le n$. Similarly, $\pi(2j-1) \ge 2j - 1 > 2i_3 - 1$ for any $j > i_3$, and $\pi(2j-1) > \pi(2i_4) \ge 2i_4 - 1 > 2i_3 - 1$ for any $j \le i_3$ since the superdiagonal part of π avoids 132. Thus, $\pi(2j-1) > 2i_3 - 1$ for any $1 \le j \le n$.

Therefore, no entry of π lies in the interval $[2i_2 - 1, 2i_3 - 1]$, which is nonempty since $2i_2 < 2i_3 - 1$. This is, of course, impossible, so the lemma follows. $\square$

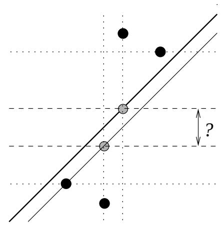


FIGURE 3. This situation is impossible in Lemma 3.9: no value between the grey points (inclusive) can occur in π.

Hence, there is a bijection between $\pi \in \mathfrak{D}^2_{2n}(2143)$ and pairs (π_1, π_2), where $\pi_1 \in S_n(132)$ is an upper board and $\pi_2 \in S_n(213)$ is a lower board. Thus, $|\mathfrak{D}^2_{2n}(2143)| = b_n b_{n+1}$, where b_n is the number of lower boards $\pi \in S_n(213)$ and b_{n+1} is the number of upper boards $\pi \in S_n(132)$ (see the remark after Lemma 3.8).

LEMMA 3.10. *Let $F(x) = \sum_{m=0}^{\infty} b_{2m} x^m$ and $G(x) = \sum_{m=0}^{\infty} b_{2m+1} x^m$. Then we have $b_0 = 1$ and*

$$b_{2m} = \sum_{i=0}^{m-1} b_{2i} b_{2m-2i-1}, \quad b_{2m+1} = \sum_{i=0}^{m} b_{2i} b_{2m-2i},$$
$$F(x) = 1 + xF(x)G(x), \quad G(x) = F(x)^2.$$

PROOF. Let $\pi \in S_n(213)$ be a lower board, and let $i \ge 0$ be maximal such that $\pi(i+1) = 2i+1$. Such an i always exists since $\pi(1) = 1$. Then $\pi(j) \le 2j - 2$ for $j \ge i + 2$. Furthermore, π avoids 213, so if $j_1, j_2 > i + 1$, and $\pi(j_1) > \pi(i+1) > \pi(j_2)$, then $j_1 < j_2$. In other words, all entries of π greater than and to the right of $2i+1$ must come before all entries less than and to the right of $2i+1$ (see Figure 4, the areas that cannot contain entries of π are shaded). In addition, $\pi(j) \le 2i + 1$ for $j \le i + 1$, so $\pi(j) > 2i + 1$ only if $j > i + 1$. There are $n - 2i - 1$ values greater than $2i + 1$ in π, hence they must occupy the $n - 2i - 1$ positions immediately to the right of $\pi(i+1)$, i.e. positions $i + 2$ through $n - i$. It is not difficult now to see from the above argument that all entries of π greater than $2i + 1$ must lie on a board of lower type in $S_{n-2i-1}(213)$, while the entries less than $2i + 1$ in π must lie on two boards whose concatenation is a lower board in $S_{2i}(213)$ (unshaded areas in Figure 4). $\square$

Thus, we get the same generating function equations as in Remark 3.6, so $F(x) = f(x)$, $G(x) = g(x)$, and hence $b_n = a_n$ for all $n \ge 0$. This proves Theorem 3.5.

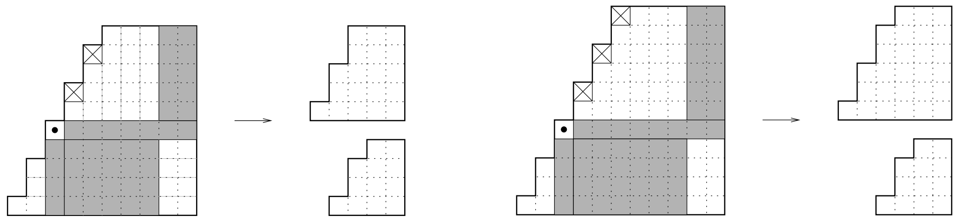


FIGURE 4. A lower board $\pi \in S_n(213)$ ($n = 10$ (even), left, and $n = 11$ (odd), right) decomposed into two lower boards according to the largest i such that $\pi(i+1) = 2i+1$ (here $i = 2$).

We can give a direct bijection showing that $b_n = a_n$. It is well-known that a_{2n} (resp. a_{2n+1}) is the number of northeast lattice paths from $(0,0)$ to $(2n, n)$ (resp. from $(0,0)$ to $(2n+1, n)$) that do not get above the line $y = x/2$. The following bijection uses the same idea as a bijection of Krattenthaler [**10**] from the set of 132-avoiding permutations in S_n to Dyck paths of semilength n, which is described in Section 2.1.

We introduce a bijection between the set of lower boards in $S_n(213)$ and northwest paths from $(n,0)$ to $(\lceil n/2 \rceil, n)$ that stay on or above the line $y = 2n - 2x$ (see Figure 5). Given a lower board in $S_n(213)$ represented as an $n \times n$ binary array, consider a lattice path from $(n,0)$ to $(\lceil n/2 \rceil, n)$ that leaves all dots to the left and stays as close to the $y = 2n - 2x$ as possible. We claim that such a path must stay on or above the line $y = 2n - 2x$. Indeed, considering rows of a lower board from top to bottom, we see that at most one extra column appears on the left for every two consecutive rows. Therefore, our path must shift at least r columns to the right for every $2r$ consecutive rows starting from the top. The rest is easy to see.

Conversely, given a northwest path from $(n,0)$ to $(\lceil n/2 \rceil, n)$ not below the line $y = 2n - 2x$, fill the corresponding board from top to bottom (i.e. from row n to row 1) so that the dots are in the rightmost column to the left of the path that still contains no dots.

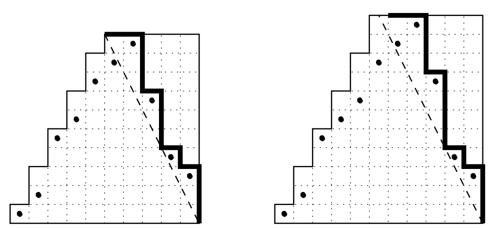


FIGURE 5. A bijection between lower boards in $S_n(213)$, for $n = 10$ (left) and $n = 11$ (right), and northwest paths from $(n,0)$ to $(\lceil n/2 \rceil, n)$ not below $y = 2n - 2x$.

The *median Genocchi number* (or *Genocchi number of the second kind*) H_n [**17**, A005439] counts the number of derangements in $\mathfrak{D}^2_{2n}$ (also, the number of permutations in $\mathfrak{D}^1_{2n}$ which begin with n or $n+1$). Using the preceding argument, we can also count the number of derangements in $\mathfrak{D}^2_{2n}(2143)$.

THEOREM 3.11. *The number of derangements in $\mathfrak{D}^2_{2n}(2143)$ is a_n^2, where a_n is as in Theorem 3.5.*

PROOF. Notice that the fixed points of a permutation $\pi \in \mathfrak{D}^2_{2n}(2143)$ correspond to the dots in the lower right (southeast) corner cells on its upper board (except the lowest right corner when n is odd) (see Figure 2). It is easy to see that deletion of those cells on an upper board produces a rotation of a lower board by 180°. This, together with the preceding lemmas, implies the theorem. $\square$

The following theorem gives the generating function for the distribution of the number of fixed points among permutations in $\mathfrak{D}^2_{2n}(2143)$.

THEOREM 3.12. *We have*

$$\sum_{\pi \in \mathfrak{D}^2_{2n}(2143)} q^{\mathrm{fix}(\pi)} = a_n [x^{n+1}] \left(\frac{1}{1 - xf(x^2)} \cdot \frac{1}{1 - qx^2 f(x^2)^2} \right). \tag{3.8}$$

where $f(x) = \sum_{n \geq 0} a_{2n} x^n$ is a solution of $f(x) = 1 + xf(x)^3$.

Note that $\sum_{n\geq 0} a_{2n}x^{2n} = f(x^2)$ and $g(x) = \sum_{n\geq 0} a_{2n+1}x^n = f(x)^2$ implies $\sum_{n\geq 0} a_{2n+1}x^{2n+1} = xf(x^2)^2$. Hence,

$$\sum_{n\geq 0} a_n x^n = f(x^2) + xf(x^2)^2 = \frac{1}{1-xf(x^2)}.$$

Proof. Let $\pi \in \mathfrak{D}^2_{2n}(2143)$. Note that all fixed points must be on the upper board of π. Therefore, the lower board of π may be any 213-avoiding lower board. This accounts for the factor a_n. Now consider the product of two rational functions on the right. This products corresponds to the fact that the upper board B of π is a concatenation of two objects: the upper board B' of rows below the lowest (smallest) fixed point, and the upper board B'' of rows not below the lowest fixed point. It is easy to see that B' may be any 132-avoiding upper board. Note that B'' must necessarily have an even number of rows and that B'' is a concatenation of a sequence of "slices" between consecutive fixed points, where the ith slice consists of an even number of rows below the $(i+1)$-th smallest fixed point but not below the ith smallest fixed point.

Thus, we obtain a block decomposition of the upper board B (similar to the one in the Figure 4 for lower boards) into an possibly empty upper board B' and a sequence B'' of nonempty upper boards $B''_1, B''_2, \ldots$, where each B''_i contains an even number of rows and exactly 1 fixed point of π. Taking generating functions yields the product of functions on the right-hand side of (3.8). $\square$

In conclusion, we note that not all results of the full paper fit in the length of this extended abstract. Using the same methods as in [**12**], we may similarly obtain the generating function for the number of Dumont permutations of the first kind simultaneously avoiding certain pairs of 4-letter patterns and another pattern of arbitrary length in terms of Chebyshev polynomials.

References

[1] M. Bóna, *Combinatorics of Permutations*, Chapman & Hall/CRC Press, 2004.

[2] M. Bóna, Exact enumeration of 1342-avoiding permutations: A close link with labeled trees and planar maps. *J. Combin. Theory, Ser. A* **80** (1997), 257–272.

[3] A. Burstein, Restricted Dumont permutations, *Ann. Combin.* **9** (2005), no. 3, 269–280.

[4] D. Dumont, Interpretations combinatoires des nombres de Genocchi, *Duke J. Math.* **41** (1974), 305–318.

[5] S. Elizalde, Fixed points and excedances in restricted permutations, *Proceedings of FPSAC'03*, June 2003, University of Linköping, Sweden, `arxiv:math.CO/0212221`.

[6] S. Elizalde, I. Pak, Bijections for refined restricted permutations, *J. Combin. Theory Ser. A* **105** (2004), 207–219.

[7] I. Gessel, Symmetric functions and p-recursiveness, *J. Combin. Theory, Ser. A* **53** (1990), 257–285.

[8] S. Kitaev, T. Mansour, A survey of certain pattern problems, preprint.

[9] D.E. Knuth, The Art of Computer Programming, vols. 1, 3, Addison-Wesley, NY, 1968, 1973.

[10] C. Krattenthaler, Permutations with restricted patterns and Dyck paths, *Adv. Appl. Math.* **27** (2001), 510–530.

[11] T. Mansour, Restricted 132-Dumont permutations, *Australasian J. Combin.* **29** (2004), 103–118.

[12] T. Mansour, A. Vainshtein, Restricted permutations and Chevyshev polynomials, *Sem. Loth. Comb.* (2002), B47c.

[13] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, *Adv. Math.* **41** (1981), 115–136.

[14] A. Reifegerste, On the diagram of 132-avoiding permutations, *European J. Combin.* **24** (2003), 759–776.

[15] F. Ruskey, Combinatorial Object Server, `http://www.theory.csc.uvic.ca/~cos/inf/perm/GenocchiInfo.html`.

[16] R. Simion, F.W. Schmidt, Restricted permutations, *Europ. J. Combin.* **6** (1985), 383–406.

[17] N.J.A. Sloane, S. Plouffe, The Encyclopedia of Integer Sequences, Academic Press, New York, 1995. Online at `http://www.research.att.com/~njas/sequences`.

[18] Z. Stankova, Forbidden subsequences, *Discrete Math.* **132** (1994), 291–316.

[19] Z. Stankova, Classification of forbidden subsequences of length 4, *Europ. J. Combin.* **17** (1996), 501–517.

[20] E. Steíngrimsson, L.K. Williams, Permutation tableaux and permutation patterns, preprint, `arXiv:math.CO/0507149`.

[21] J. West, Permutations with forbidden subsequences and stack-sortable permutations, Ph.D. thesis, M.I.T., 1990.

[22] J. West, Permutation trees and the Catalan and Schröder numbers *Discrete Math.* **146** (1995), 247–262.

[23] J. West, Generating trees and forbidden subsequences, *Discrete Math.* **157** (1996), 363–374.

Department of Mathematics, Iowa State University, Ames, IA 50011-2064 USA
E-mail address: `burstein@math.iastate.edu`

Department of Mathematics, Dartmouth College, 6188 Bradley Hall, Hanover, NH 03755 USA
E-mail address: `sergi.elizalde@dartmouth.edu`

Department of Mathematics, University of Haifa, 31905 Haifa, Israel
E-mail address: `toufik@math.haifa.ac.il`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Restricted Patience Sorting and Barred Pattern Avoidance

Alexander Burstein and Isaiah Lankham

Abstract. Patience Sorting is a combinatorial algorithm that can be viewed as an iterated, non-recursive form of the Schensted Insertion Algorithm. In recent work the authors have shown that Patience Sorting provides an algorithmic description for permutations avoiding the barred (generalized) permutation pattern $3-\bar{1}-42$. Motivated by this and a recently formulated geometric form for Patience Sorting in terms of certain intersecting lattice paths, we study the related themes of restricted input and avoidance of similar barred permutation patterns. One such result is to characterize those permutations for which Patience Sorting is an invertible algorithm as the set of permutations simultaneously avoiding the barred patterns $3-\bar{1}-42$ and $3-\bar{1}-24$. We then enumerate this avoidance set, which involves convolved Fibonacci numbers.

Résumé. *Patience Sorting* est un algorithme combinatoire que l'on peut comprendre comme étant une version itérée, non-récursive de la correspondence de Schensted. Dans leur travail récent les auteurs ont démontré que *Patience Sorting* donne une description algorithmique des permutations évitant le motif barré (généralisé) $3-\bar{1}-42$. Motivés par ceci et par une forme récemment formulée de *Patience Sorting* en termes de certaine parcours du treillis intersectants, nous étudions les thèmes connexe d'input restreinte et permutations qui évitent de similaire motifs barrés. Un de nos résultats est de caractériser les permutations pour lesquelles *Patience Sorting* est un algorithme inversible comme étant l'ensemble des permutations évitant simultanément les motifs barrés $3-\bar{1}-42$ and $3-\bar{1}-24$. Nous énumérons ensuite cet ensemble, qui utilise des convolutions des nombres de Fibonacci.

1. Introduction

The term *Patience Sorting* was introduced in 1962 by C. L. Mallows [**12, 13**] while studying a card sorting algorithm invented by A. S. C. Ross. Given a shuffled deck of cards $\sigma = c_1 c_2 \cdots c_n$ (which we take to be a permutation $\sigma \in \mathfrak{S}_n$), Ross proposed the following algorithm:

Step 1 Use what Mallows called a "patience sorting procedure" to form the subsequences $r_1, r_2, \ldots, r_m$ of σ (called *piles*) as follows:

- Place the first card c_1 from the deck into a pile r_1 by itself.
- For each remaining card c_i ($i = 2, \ldots, n$), consider the cards $d_1, d_2, \ldots, d_k$ atop the piles $r_1, r_2, \ldots, r_k$ that have already been formed.
  - If $c_i > \max\{d_1, d_2, \ldots, d_k\}$, then put c_i into a new right-most pile r_{k+1} by itself.
  - Otherwise, find the left-most card d_j that is larger than c_i and put the card c_i atop pile r_j.

2000 *Mathematics Subject Classification.* Primary: 05A05, 05A15, 05A18; Secondary: 05E10.

Key words and phrases. patience sorting, barred and generalized permutation patterns, shadow diagrams, intersecting lattice paths, Skew Fibonacci-Pascal triangle, convolved Fibonacci numbers.

The work of the first author was supported in part by the U.S. National Security Agency Young Investigator Grant MSPR-05Y-113.

The work of the second author was supported in part by the U.S. National Science Foundation under Grants DMS-0135345 and DMS-0304414.

Step 2 Gather the cards up one at a time from these piles in ascending order.

We call **Step 1** of the above algorithm *Patience Sorting* and denote by $R(\sigma) = \{r_1, r_2, \ldots, r_m\}$ the *pile configuration* associated to the permutation $\sigma \in \mathfrak{S}_n$. Moreover, given any pile configuration R, one forms its *reverse patience word* $RPW(R)$ by listing the piles in R "from bottom to top, left to right" (i.e., by reversing the so-called "far-eastern reading"). In [**5**] these words are characterized as being exactly the elements of the avoidance set $S_n(3\text{-}\bar{1}\text{-}42)$. That is, they are permutation avoiding the generalized pattern 2-31 unless every occurrence of 2-31 is contained within an occurrence of the generalized pattern 3-1-42. (A review of generalized permutation patterns can be found in Section 1.2 below).

We illustrate the formation of $R(\sigma)$ and $RPW(R)$ in the following example.

Example 1.1. Let $\sigma = 64518723 \in \mathfrak{S}_8$. Then we form the pile configuration $R(\sigma)$ as follows:

Form a new pile with **6**:	$\mathbf{6}$	Then place **4** atop 6:	$\begin{array}{l}\mathbf{4}\\6\end{array}$	Form a new pile with **5**:	$\begin{array}{ll}4&\\6&\mathbf{5}\end{array}$
Add **1** to the left-most pile:	$\begin{array}{ll}\mathbf{1}&\\4&\\6&5\end{array}$	Form a new pile with **8**:	$\begin{array}{lll}1&&\\4&&\\6&5&\mathbf{8}\end{array}$	Then place **7** atop 8:	$\begin{array}{lll}1&&\\4&&\mathbf{7}\\6&5&8\end{array}$
Add **2** to the middle pile:	$\begin{array}{lll}1&&\\4&\mathbf{2}&7\\6&5&8\end{array}$	Finally, place **3** atop 7:	$\begin{array}{lll}1&&\mathbf{3}\\4&2&7\\6&5&8\end{array}$		

Then, by reading up the columns of $R(\sigma)$ from left to right, $RPW(R(\sigma)) = 64152873 \in S_8(3\text{-}\bar{1}\text{-}42)$.

Given $\sigma \in \mathfrak{S}_n$, the formation of $R(\sigma)$ can be viewed as an iterated, non-recursive form of the Schensted Insertion Algorithm for interposing values into the rows of a standard Young tableau (see [**2**]). In [**5**] the authors augment the formation of $R(\sigma)$ so that the resulting extension of Patience Sorting becomes a full non-recursive analogue of the celebrated Robinson-Schensted-Knuth (or RSK) Correspondence. As with RSK, this Extended Patience Sorting Algorithm (given as Algorithm 1.2 in Section 1.1 below) takes a simple idea — that of placing cards into piles — and uses it to build a bijection between elements of the symmetric group $\mathfrak{S}_n$ and certain pairs of combinatorial objects. In the case of RSK, one uses the Schensted Insertion Algorithm to build a bijection with (unrestricted) pairs of standard Young tableau having the same shape (see [**16**]). However, in the case of Patience Sorting, one achieves a bijection between permutations and (somewhat more restricted) pairs of pile configurations having the same shape. We denote this latter bijection by $\sigma \xleftrightarrow{PS} (R(\sigma), S(\sigma))$ and call $R(\sigma)$ (resp. $S(\sigma)$) the *insertion piles* (resp. *recording piles*) corresponding to σ. Collectively, we also call $(R(\sigma), S(\sigma))$ the *stable pair* of pile configurations corresponding to σ and characterize such pairs in [**5**] using a somewhat involved pattern avoidance condition on their reverse patience words.

Barred (generalized) permutation patterns like $3\text{-}\bar{1}\text{-}42$ arise quite naturally when studying Patience Sorting. We discuss and enumerate the avoidance classes for several related patterns in Section 2. Then, in Section 3, we examine properties of Patience Sorting under restricted input that can be characterized using such patterns. One such characterization, discussed in Section 3.1, is for the crossings in the initial iteration of the Geometric Patience Sorting Algorithm given by the authors in [**6**]. This geometric form for the Extended Patience Sorting Algorithm is naturally dual to Viennot's Geometric RSK (originally defined in [**19**]) and gives, among other things, a geometric interpretation for the stable pairs of $3\text{-}\bar{1}\text{-}42$-avoiding permutations corresponding to a permutation under Extended Patience Sorting. However, unlike Viennot's geometric form for RSK, the shadow lines in Geometric Patience Sorting are allowed to cross. While a complete characterization for these crossings is given in [**6**] in terms of the pile configurations formed, this new result is the first step in providing a characterization for the permutations involved in terms of barred pattern avoidance.

We close this introduction by describing both the Extended and Geometric Patience Sorting Algorithms. We also briefly review the notation of generalized permutation patterns.

1.1. Extended and Geometric Patience Sorting. Mallows' original "patience sorting procedure" can be extended to a full bijection between the symmetric group $\mathfrak{S}_n$ and certain restricted pairs of pile configurations using the following algorithm (which was first introduced in [**5**]):

ALGORITHM 1.2 (Extended Patience Sorting Algorithm). Given $\sigma = c_1c_2\cdots c_n \in \mathfrak{S}_n$, inductively build *insertion piles* $R(\sigma) = \{r_1, r_2, \ldots, r_m\}$ and *recording piles* $S(\sigma) = \{s_1, s_2, \ldots, s_m\}$ as follows:

- Place the first card c_1 from the deck into a pile r_1 by itself, and set $s_1 = \{1\}$.
- For each remaining card c_i $(i = 2, \ldots, n)$, consider the cards $d_1, d_2, \ldots, d_k$ atop the piles $r_1, r_2, \ldots, r_k$ that have already been formed.
  - If $c_i > \max\{d_1, d_2, \ldots, d_k\}$, then put c_i into a new pile r_{k+1} by itself and set $s_{k+1} = \{i\}$.
  - Otherwise, find the left-most card d_j that is larger than c_i and put the card c_i atop pile r_j while simultaneously putting i at the bottom of pile s_j.

Note that the pile configurations that comprise a resulting stable pair must have the same "shape", which we define as follows:

DEFINITION 1.3. Given a pile configuration $R = \{r_1, r_2, \ldots, r_m\}$ on n cards, we call the composition $\gamma = (|r_1|, |r_2|, \ldots, |r_m|)$ of n the *shape* of R and denote this by $\mathrm{sh}(R) = \gamma \models n$.

The idea behind Algorithm 1.2 is that we are using the auxiliary pile configuration $S(\sigma)$ to implicitly label the order in which the elements of the permutation $\sigma \in \mathfrak{S}_n$ are added to the usual Patience Sorting pile configuration $R(\sigma)$ (which we now call the "insertion piles" of σ in this context by analogy to RSK). It is clear that this information then allows us to uniquely reconstruct σ by reversing the order in which the cards were played. As with normal Patience Sorting, we visualize the pile configurations $R(\sigma)$ and $S(\sigma)$ by listing their constituent piles vertically as illustrated in the following example.

EXAMPLE 1.4. Given $\sigma = 64518723 \in \mathfrak{S}_8$ from Example 1.1 above, we simultaneously form the following pile configurations with shape $\mathrm{sh}(R(\sigma)) = \mathrm{sh}(S(\sigma)) = (3, 2, 3)$ under Extended Patience Sorting (Algorithm 1.2):

$$R(\sigma) = \begin{matrix} 1 & & 3 \\ 4 & 2 & 7 \\ 6 & 5 & 8 \end{matrix} \quad \text{and} \quad S(\sigma) = \begin{matrix} 1 & & 5 \\ 2 & 3 & 6 \\ 4 & 7 & 8 \end{matrix}$$

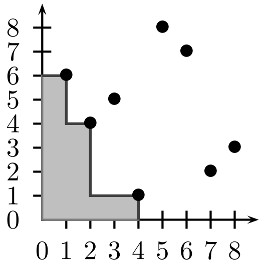
Note that the insertion piles $R(\sigma)$ are the same as the pile configuration formed in Example 1.1 and that $RPW(S(64518723)) = 42173865 \in S_8(3\text{-}\bar{1}\text{-}42)$.

In order to now describe a natural geometric form for this Extended Patience Sorting Algorithm, we begin with the following fundamental definition.

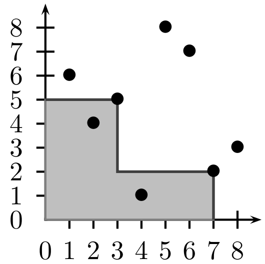
DEFINITION 1.5. Given a lattice point $(m, n) \in \mathbb{Z}^2$, we define the *(southwest) shadow* of (m, n) to be the quarter space $U(m, n) = \{(x, y) \in \mathbb{R}^2 \mid x \leq m,\ y \leq n\}$.

As with the northeasterly-oriented shadows that Viennot used when building his geometric form for RSK (see [**19**]), the most important use of these southwesterly-oriented shadows is in building shadowlines (which is illustrated in Figure 1(a)):

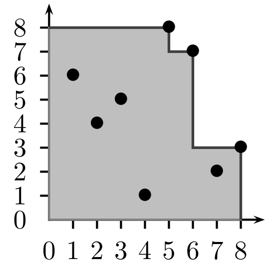
DEFINITION 1.6. The *(southwest) shadowline* of $(m_1, n_1), (m_2, n_2), \ldots, (m_k, n_k) \in \mathbb{Z}^2$ is defined to be the boundary of the union of the shadows $U(m_1, n_1), U(m_2, n_2), \ldots, U(m_k, n_k)$.



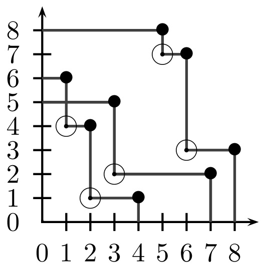
(a) Shadowline $L_1^{(0)}(64518723)$.



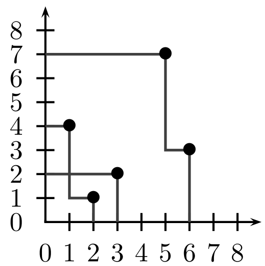
(b) Shadowline $L_2^{(0)}(64518723)$.



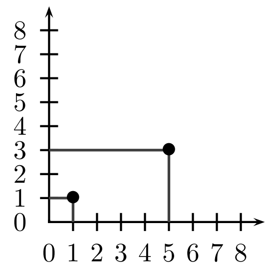
(c) Shadowline $L_3^{(0)}(64518723)$.



(d) Salient points ($\odot$) for $D^{(0)}(64518723)$.



(e) Shadow Diagram $D^{(1)}(64518723)$.



(f) Shadow Diagram $D^{(2)}(64518723)$.

FIGURE 1. Examples of Shadowline and Shadow Diagram Construction.

In particular, we wish to associate to each permutation a certain collection of (southwest) shadowlines called its *shadow diagram*. However, unlike the northeasterly-oriented shadowlines used to define the northeast shadow diagrams of Geometric RSK [**19**], these southwest shadowlines are allowed to intersect as illustrated in Figure 1(d)–(e). (We characterize those permutations having intersecting shadowlines under Definition 1.7 in Theorem 3.6 below.)

DEFINITION 1.7. The *(southwest) shadow diagram* $D^{(0)}(\sigma)$ of $\sigma = \sigma_1\sigma_2\cdots\sigma_n \in \mathfrak{S}_n$ consists of the (southwest) shadowlines $D^{(0)}(\sigma) = \{L_1^{(0)}(\sigma), L_2^{(0)}(\sigma), \ldots, L_k^{(0)}(\sigma)\}$ formed as follows:

- $L_1^{(0)}(\sigma)$ is the shadowline for those lattice points $(x, y) \in \{(1, \sigma_1), (2, \sigma_2), \ldots, (n, \sigma_n)\}$ such that the shadow $U(x, y)$ does not contain any other lattice points.
- While at least one of the points $(1, \sigma_1), (2, \sigma_2), \ldots, (n, \sigma_n)$ is not contained in the shadowlines $L_1^{(0)}(\sigma), L_2^{(0)}(\sigma), \ldots, L_j^{(0)}(\sigma)$, define $L_{j+1}^{(0)}(\sigma)$ to be the shadowline for the points
$$(x, y) \in A := \{(i, \sigma_i) \mid (i, \sigma_i) \notin \bigcup_{k=1}^{j} L_k^{(0)}(\sigma)\}$$
such that the shadow $U(x, y)$ does not contain any other lattice points from the set A.

In other words, we define a shadow diagram by inductively eliminating points in the permutation diagram until every point has been used to define a shadowline (as illustrated in Figure 1(a)–(c)).

One can prove (see [**5**]) that the ordinates (i.e., y-coordinates) of the points used to define each shadowline in the shadow diagram $D^{(0)}(\sigma)$ are exactly the *left-to-right minima subsequences* (a.k.a. *basic subsequences*) in the permutation $\sigma \in \mathfrak{S}_n$. These are defined as follows:

DEFINITION 1.8. Let $\pi = \pi_1\pi_2\cdots\pi_l$ be a partial permutation on the set $[n] = \{1, 2, \ldots, n\}$. Then the *left-to-right minima* (resp. *maxima*) *subsequence* of π consists of those components π_j of π such that $\pi_j = \min\{\pi_i \mid 1 \leq i \leq j\}$ (resp. $\pi_j = \max\{\pi_i \mid 1 \leq i \leq j\}$).

We then inductively define the left-to-right minima (resp. maxima) subsequences $s_1, s_2, \ldots, s_k$ of the permutation σ by taking s_1 to be the left-to-right minima (resp. maxima) subsequence for σ itself and then each remaining subsequence s_i to be the left-to-right minima (resp. maxima) subsequence for the partial permutation obtained by removing the elements of $s_1, s_2, \ldots, s_{i-1}$ from σ.

Finally, one can produce a sequence $D(\sigma) = (D^{(0)}(\sigma),\ D^{(1)}(\sigma),\ D^{(2)}(\sigma),\ \ldots)$ of shadow diagrams for a given permutation $\sigma \in \mathfrak{S}_n$ by recursively applying Definition 1.7 to the southwest corners (called *salient points*) of a given set of shadowlines (as illustrated in Figure 1(d)–(f)). The only difference is that, with each iteration, newly formed shadowlines can only connect salient points along the same pre-existing shadowline. One can then uniquely reconstruct the pile configurations $R(\sigma)$ and $S(\sigma)$ from these shadowlines by taking their intersections with the x- and y-axes in a certain canonical order (as detailed in [**6**]).

DEFINITION 1.9. We call $D^{(k)}(\sigma)$ the k^{th} *iterate* of the *exhaustive shadow diagram* $D(\sigma)$ for the permutation $\sigma \in \mathfrak{S}_n$.

1.2. Generalized Pattern Avoidance. We first recall the following definition:

DEFINITION 1.10. Let $\sigma = \sigma_1\sigma_2\cdots\sigma_n \in \mathfrak{S}_n$ and $\pi \in \mathfrak{S}_m$ with $m \leq n$. Then we say that σ *contains* the *(classical) permutation pattern* π if there exists a subsequence $(\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_m})$ of σ (meaning $i_1 < i_2 < \cdots < i_m$) such that the word $\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_m}$ is order-isomorphic to π. I.e., each $\sigma_{i_j} < \sigma_{i_{j+1}}$ if and only if $\pi_j < \pi_{j+1}$.

Note, though, that the elements in the subsequence $(\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_m})$ are not required to be contiguous in σ. This motivates the

DEFINITION 1.11. A *generalized permutation pattern* is a classical permutation pattern π in which one assumes that every element in the subsequence $(\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_m})$ of σ must be taken contiguously unless a dash is inserted between the corresponding order-isomorphic elements of the pattern π.

Finally, if σ does not contain a subsequence that is order-isomorphic to π, then we say that σ *avoids* the pattern π. This motivated the

DEFINITION 1.12. Given any collection $\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)}$ of permutation patterns (classical or generalized), we denote by

$$S_n(\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)}) = \bigcap_{i=1}^{k} S_n(\pi^{(i)}) = \bigcap_{i=1}^{k} \{\sigma \in \mathfrak{S}_n \mid \sigma \text{ avoids } \pi^{(i)}\}$$

the *avoidance set* of permutations $\sigma \in \mathfrak{S}_n$ such that σ simultaneously avoids each of the patterns $\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)}$. Furthermore, the set

$$\bigcup_{n \geq 1} S_n(\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)})$$

is called the *(pattern) avoidance class* with *basis* $\{\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(k)}\}$.

More information about permutation patterns in general can be found in [**4**].

2. Barred and Unbarred Generalized Pattern Avoidance

An important further generalization of the notion of generalized permutation pattern requires that the context in which the occurrence of a generalized pattern occurs be taken into account. The resulting concept of *barred permutation patterns*, along with the accompanying notation, first arose within the study of stack-sortability of permutations by J. West [**20**]. Given how naturally these barred patterns now arise in the study of Patience Sorting (as illustrated in both [**5**] and Section 3 below), we initiate their systematic study in this section.

Definition 2.1. A *barred (generalized) permutation pattern* β is a generalized permutation pattern in which overbars are used to indicate that barred values cannot occur at the barred positions. As before, we denote by $S_n(\beta^{(1)}, \ldots, \beta^{(k)})$ the set of all permutations $\sigma \in \mathfrak{S}_n$ that simultaneously avoid $\beta^{(1)}, \ldots, \beta^{(k)}$ (i.e., permutations that contain no subsequence that is order-isomorphic to any of the $\beta^{(1)}, \ldots, \beta^{(k)}$).

Example 2.2. A permutation $\sigma = \sigma_1\sigma_2\cdots\sigma_n \in \mathfrak{S}_n \notin S_n(3\text{-}\bar{5}\text{-}2\text{-}4\text{-}1)$ contains an occurrence of the barred permutation pattern 3-$\bar{5}$-2-4-1 if it contains an occurrence of the generalized pattern 3-2-4-1 (i.e., contains a subsequence $(\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_m})$ that is order-isomorphic to the classical pattern 3241) in which no value larger than the element playing the role of "4" is allowed to occur between the elements playing the roles of "3" and "2". This is one of the two basis elements for the pattern avoidance class used to characterize the set of 2-stack-sortable permutations [**10, 11, 20**]. (The other pattern is 2-3-4-1, i.e., the classical pattern 2341.)

Despite the added complexity involved in avoiding barred permutation patterns, it is still sometimes possible to characterize the avoidance class for a barred permutation pattern in terms of an unbarred generalized permutation pattern. The following theorem gives such a characterization for the pattern 3-$\bar{1}$-42. (Note, though, that there is no equivalent characterization for such barred permutation patterns as 1$\bar{3}$-42 and 3-$\bar{5}$-2-4-1.)

Theorem 2.3. *Let $B_n = \frac{1}{e}\sum_{k\geq 0}\frac{k^n}{k!}$ denote the n^{th} Bell number. Then*
(1) $S_n(3\text{-}\bar{1}\text{-}42) = S_n(3\text{-}\bar{1}\text{-}4\text{-}2) = S_n(23\text{-}1)$
(2) $|S_n(3\text{-}\bar{1}\text{-}42)| = B_n$

Proof. (Sketch)
As in [**7**], we see that each of these sets consists of permutations having the form

$$\sigma = \sigma_1 a_1 \sigma_2 a_2 \ldots \sigma_k a_k,$$

where $a_k > a_{k-1} > \cdots > a_2 > a_1$ are the successive right-to-left minima of σ (reversing the order of the elements in Definition 1.8) and where each segment $\sigma_i a_i$ is a decreasing subsequence. $\square$

Remark 2.4. We emphasize the following important consequences of Theorem 2.3.

(1) Even though $S_n(3\text{-}\bar{1}\text{-}42) = S_n(23\text{-}1)$ by Theorem 2.3(1), it is more natural to use avoidance of the barred pattern 3-$\bar{1}$-42 in studying Patience Sorting. As shown in [**5**] and elaborated upon in Section 3 below, $S_n(3\text{-}\bar{1}\text{-}42)$ is the set of equivalence classes of $\mathfrak{S}_n$ modulo the transitive closure of the relation 3-$\bar{1}$-42 $\sim$ 3-$\bar{1}$-24. (I.e., two permutations $\sigma, \tau \in \mathfrak{S}_n$ are equivalent if the elements creating an occurrence of one of these patterns in σ form an occurrence of the other pattern in τ.) Moreover, each permutation $\sigma \in \mathfrak{S}_n$ in a given equivalence class has the same pile configuration $R(\sigma)$ under Patience Sorting, a description of which is significantly more difficult to describe for occurrences of the unbarred generalized permutation pattern 23-1.

(2) Marcus and Tardos proved in [**14**] that the avoidance set $S_n(\pi)$ for any classical pattern π grows at most exponentially fast as $n \to \infty$. (This was previously known as the Stanley-Wilf Conjecture.) The Bell numbers, though, satisfy $\log B_n = n(\log n - \log\log n + O(1))$ and so exhibit superexponential growth. (See [**18**] for more information about Bell numbers.) While it was previously known that the Stanley-Wilf Conjecture does not extend to generalized permutation patterns (see, e.g., [**7**]), it took Theorem 2.3(2) (originally proven in [**5**] using Patience Sorting) to provide the first verification that one also cannot extend the Stanley-Wilf Conjecture to barred generalized permutation patterns.

A further abstraction of barred permutation pattern avoidance (called *Bruhat-restricted avoidance*) was recently given by A. Woo and A. Yong in [**21**]. The result in Theorem 2.3(2) has led A. Woo to conjecture to the second author that the Stanley-Wilf ex-Conjecture also does not extend to this new notion of pattern avoidance.

We conclude this section with a simple corollary to Theorem 2.3 that gives similar equivalences and enumerations for some barred permutation patterns that also arise naturally in the study of Patience Sorting (see Proposition 3.2 and Theorem 3.6 in Section 3 below).

Corollary 2.5. *Using the notation in Theorem 2.3,*

(1) $S_n(31\text{-}\bar{4}\text{-}2) = S_n(3\text{-}1\text{-}\bar{4}\text{-}2) = S_n(3\text{-}12)$
(2) $S_n(\bar{2}\text{-}41\text{-}3) = S_n(\bar{2}\text{-}4\text{-}1\text{-}3) = S_n(2\text{-}4\text{-}1\text{-}\bar{3}) = S_n(2\text{-}41\text{-}\bar{3})$
(3) $|S_n(\bar{2}\text{-}41\text{-}3)| = |S_n(31\text{-}\bar{4}\text{-}2)| = |S_n(3\text{-}\bar{1}\text{-}42)| = B_n$.

Proof. (Sketches)

(1) Take reverse complements in $S_n(3\text{-}\bar{1}\text{-}42)$ and apply Theorem 2.3.
(2) Similar to (1). (Note that (2) is also proven in [**1**].)
(3) This follows from the fact that the patterns 3-1-$\bar{4}$-2 and $\bar{2}$-4-1-3 are inverses of each other. □

3. Patience Sorting under Restricted Input

3.1. Patience Sorting on Restricted Permutations. The similarities between the Extended Patience Sorting Algorithm (Algorithm 1.2) and RSK applied to permutations is perhaps most observable in the following simple proposition:

Proposition 3.1. *Let* $\imath_k = 1\text{-}2\text{-}\cdots\text{-}k = 12\cdots k$ *and* $\jmath_k = k\text{-}\cdots\text{-}2\text{-}1 = k\cdots 21$ *be the classical monotone permutation patterns. Then there is*

(1) *a bijection between* $S_n(\imath_{k+1})$ *and pairs of pile configurations having the same composition shape* $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_m) \models n$ *but with at most* k *piles (i.e.,* $m \le k$*).*
(2) *a bijection between* $S_n(\jmath_{k+1})$ *and pairs of pile configurations having the same composition shape* $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_m) \models n$ *but with no pile having more than* k *cards in it (i.e.,* $\gamma_i \le k$ *for each* $i = 1, 2, \dots, m$*).*

Proof. (Sketches)

(1) Given $\sigma \in \mathfrak{S}_n$, a bijection is formed in [**2**] between the set of piles $R(\sigma) = \{r_1, r_2, \dots, r_k\}$ formed under Patience Sorting and the components of a particular distinguished longest increasing subsequence in σ. Since avoiding the monotone pattern $\imath_{k+1}$ is equivalent to restricting the length of the longest increasing subsequence in a permutation, the result then follows.
(2) Follows from (1) by reversing each of the permutations in $S_n(\imath_{k+1})$ in order to form $S_n(\jmath_{k+1})$. □

Proposition 3.1 states that Patience Sorting can be used to efficiently compute the length of both the longest increasing and longest decreasing subsequences in a given permutation. In particular, one can compute these lengths without needing to examine every subsequence of a permutation, just as with RSK. However, while both RSK and Patience Sorting can be used to implement this computation in $O(n \log(n))$ time, an extension of this technique is given in [**3**] that also simultaneously tabulates all of the longest increasing or decreasing subsequences without incurring any additional asymptotic computational cost.

As mentioned in Section 2 above, Patience Sorting also has immediate connections to certain barred permutation patterns:

Proposition 3.2.

(1) $S_n(3\text{-}\bar{1}\text{-}42) = \{RPW(R(\sigma)) \mid \sigma \in \mathfrak{S}_n\}$. *In particular, given* $\sigma \in S_n(3\text{-}\bar{1}\text{-}42)$*, the entries in each column of the insertion piles* $R(\sigma)$ *(when read from bottom to top) occupy successive positions in the permutation* σ*.*
(2) $S_n(\bar{2}\text{-}41\text{-}3) = \{RPW(R(\sigma^{-1})) \mid \sigma \in \mathfrak{S}_n\}$. *In particular, given* $\sigma \in S_n(\bar{2}\text{-}41\text{-}3)$*, the columns of the insertion piles* $R(\sigma)$ *(when read from top to bottom) contain successive values.*

Proof. Part (1) is proven in [**5**], and part (2) follows immediate by taking inverses in (1). □

As an immediate corollary, we can characterize an important category of classical permutation patterns in terms of barred permutation patterns.

Definition 3.3. Given a composition $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_m) \models n$, the *(classical) layered permutation pattern* $\pi_\gamma \in \mathfrak{S}_n$ is the permutation

$$\gamma_1 \cdot (\gamma_1 - 1) \cdots 1 \cdot (\gamma_1 + \gamma_2) \cdot (\gamma_1 + \gamma_2 - 1) \cdots (\gamma_1 + 1) \cdots n \cdot (n-1) \cdots (\gamma_1 + \gamma_2 + \cdots + \gamma_{m-1} + 1).$$

Example 3.4. Given $\gamma = (3, 2, 3) \models 8$, the corresponding layered pattern is $\pi_{(3,2,3)} = \widehat{321}\widehat{54}\widehat{876} \in \mathfrak{S}_8$ (following the notation in [**15**]). Moreover, applying Extended Patience Sorting (Algorithm 1.2) to $\pi_{(3,2,3)}$:

$$R(\pi_{(3,2,3)}) = \begin{array}{ccc} 1 & & 6 \\ 2 & 4 & 7 \\ 3 & 5 & 8 \end{array} \quad \text{and} \quad S(\pi_{(3,2,3)}) = \begin{array}{ccc} 1 & & 6 \\ 2 & 4 & 7 \\ 3 & 5 & 8 \end{array}$$

Note in particular that $\pi_{(3,2,3)}$ satisfies both of the conditions in Proposition 3.2, which illustrates the following characterization of layered patterns:

Corollary 3.5. *$S_n(3\text{-}\bar{1}\text{-}42, \bar{2}\text{-}41\text{-}3)$ is the set of layered patterns in $\mathfrak{S}_n$.*

Proof. Apply Proposition 3.2 noting that $S_n(3\text{-}\bar{1}\text{-}42, \bar{2}\text{-}41\text{-}3) = S_n(23\text{-}1, 31\text{-}2)$ (as considered in [**8**]). $\square$

As a consequence of this interaction between Patience Sorting and barred permutation patterns, we can now explicitly characterize those permutations for which the initial iteration of Geometric Patience Sorting (as defined in Section 1.1 above) yields non-crossing lattice paths.

Theorem 3.6. *The set $S_n(3\text{-}\bar{1}\text{-}42, 31\text{-}\bar{4}\text{-}2)$ consists of all reverse patience words having non-intersecting shadow diagrams. (I.e., no shadowlines cross in the 0^{th} iterate shadow diagram.) Moreover, given a permutation $\sigma \in S_n(3\text{-}\bar{1}\text{-}42, 31\text{-}\bar{4}\text{-}2)$, the values in the bottom rows of $R(\sigma)$ and $S(\sigma)$ increase from left to right.*

Proof. From Theorem 2.3 and Corollary 2.5, $R(S_n(3\text{-}\bar{1}\text{-}42, 31\text{-}\bar{4}\text{-}2)) = R(S_n(23\text{-}1, 3\text{-}12))$ consists exactly of set partitions of $[n] = \{1, 2, \ldots, n\}$ whose components can be ordered so that both the minimal and maximal elements of the components simultaneously increase. (These are called *strongly monotone partitions* in [**9**]).

Let $\sigma \in S_n(3\text{-}\bar{1}\text{-}42, 31\text{-}\bar{4}\text{-}2)$. Since σ avoids $3\text{-}\bar{1}\text{-}42$, we have that $\sigma = RPW(R(\sigma))$ by Proposition 3.2. Thus, the i^{th} shadowline $L_i^{(0)}(\sigma)$ of σ is the boundary of the union of shadows with generating points in decreasing segments $\sigma_i a_i$, $i \in [k]$, where $\sigma_i a_i$ are as in the proof of Theorem 2.3. Let b_i be the i^{th} left-to-right maximum of σ. Then b_i is the left-most (i.e. maximal) entry of $\sigma_i a_i$, so $\sigma_i a_i = b_i \sigma_i' a_i$ for some decreasing subsequence σ_i'. Note that σ_i' may be empty so that $b_i = a_i$.

Since b_i is the i^{th} left-to-right maximum of σ, it must be at the bottom of the i^{th} column of $R(\sigma)$ (similarly, a_i is at the top of the i^{th} column). So the bottom rows of both $R(\sigma)$ and $S(\sigma)$ must be in increasing order.

Now consider the i^{th} and j^{th} shadowlines $L_i^{(0)}(\sigma)$ and $L_j^{(0)}(\sigma)$ of σ, respectively, where $i < j$. We have that $b_i < b_j$ from which the initial horizontal segment of the i^{th} shadowline is lower than that of the j^{th} shadowline. Moreover, a_i is to the left of b_j, so the remaining segment of the i^{th} shadowline is completely to the left of the remaining segment of the j^{th} shadowline. Thus, $L_i^{(0)}(\sigma)$ and $L_j^{(0)}(\sigma)$ do not intersect. $\square$

In [**6**] the authors actually give the following stronger result:

Theorem 3.7. *Each iterate $D_{SW}^{(m)}(\sigma)$ ($m \geq 0$) of $\sigma \in \mathfrak{S}_n$ is free from crossings if and only if every row in both $R(\sigma)$ and $S(\sigma)$ is monotone increasing from left to right.*

However, this only characterizes the output of the Extended Patience Sorting Algorithm involved. As such, Theorem 3.6 provides the first step toward characterizing those permutations that result in non-crossing lattice paths under Geometric Patience Sorting.

We conclude this section by noting that, while the strongly monotone condition implied by simultaneously avoiding 3-$\bar{1}$-42 and 31-$\bar{4}$-2 is necessary to alleviate such crossings, it is clearly not sufficient. (The problem lies with what we call "polygonal crossings" in the shadow diagrams in [**6**], which occur in permutations like $\sigma = 45312$.) Thus, to avoid crossings at all iterations of Geometric Patience Sorting, we need to impose further "ordinally increasing" conditions on the set partition associated to a given permutation under Patience Sorting. In particular, in addition to requiring just the minima and maxima elements in the set partition to increase as in the strongly monotone partitions encountered in the proof of Theorem 3.6, it is necessary to require that every record value simultaneously increase under an appropriate ordering of the blocks. That is, under a single ordering of these blocks, we must simultaneously have that the largest elements in each block increase, then the next largest elements, then the next-next largest elements, and so on. E.g., the partition $\{\{5,3,1\},\{6,4,2\}\}$ of the set $[6] = \{1,2,\ldots,6\}$ satisfies this condition.

3.2. Invertibility of Patience Sorting. It is clear that the pile configurations corresponding to two permutations under the Patience Sorting Algorithm need not be distinct in general (e.g., $R(3142) = R(3412)$). As proven in [**5**], two permutations give rise to the same pile configuration under Patience Sorting if and only if they have the same left-to-right minima subsequences (e.g., 3142 and 3412 both have the left-to-right minima subsequences 31 and 42). In this section we characterize permutations having distinct pile configurations under Patience Sorting in terms of certain barred permutation patterns. We then establish a non-trivial enumeration for the resulting avoidance sets.

THEOREM 3.8. *A pile configuration pile R has a unique preimage $\sigma \in \mathfrak{S}_n$ under Patience Sorting if and only if $\sigma \in S_n(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)$.*

PROOF. (Sketch)
It is clear that every pile configuration R has at least one preimage, namely its reverse patience word $\sigma = RPW(R)$. By Proposition 3.2, reverse patience words are exactly those permutations that avoid the barred pattern 3-$\bar{1}$-42. Furthermore, as shown in [**5**], two permutations have the same insertion piles under Extended Patience Sorting (Algorithm 1.2) if and only if one can be obtained from the other by a sequence of order-isomorphic exchanges 3-$\bar{1}$-24 $\rightsquigarrow$ 3-$\bar{1}$-42 or 3-$\bar{1}$-42 $\rightsquigarrow$ 3-$\bar{1}$-24. (I.e., the occurrence of one pattern is reordered to form an occurrence of the other pattern.) Thus, it is easy to see that R has a unique preimage σ if and only if σ has no occurrence of 3-$\bar{1}$-42 or 3-$\bar{1}$-24. $\square$

Given this pattern avoidance characterization of invertibility, we have the following recurrence relation for the number of permutations having distinct pile configurations under Patience Sorting:

LEMMA 3.9. *Set $f(n) = |S_n(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)|$ and, for $k \le n$,*

$$f(n,k) = |\{\sigma \in S_n(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24) \,:\, \sigma(1) = k\}|.$$

Then $f(n) = \sum_{k=1}^{n} f(n,k)$, and we have the following recurrence relation for $f(n,k)$:

$$f(n,0) = 0 \quad \textit{for} \quad n \ge 1 \tag{3.1}$$

$$f(n,1) = f(n,n) = f(n-1) \quad \textit{for} \quad n \ge 1 \tag{3.2}$$

$$f(n,2) = 0 \quad \textit{for} \quad n \ge 3 \tag{3.3}$$

$$f(n,k) = f(n,k-1) + f(n-1,k-1) + f(n-2,k-2) \quad \textit{for} \quad n \ge 3 \tag{3.4}$$

subject to the initial conditions $f(0) = f(0,0) = 1$.

PROOF. First note that Equation (3.1) is the obvious boundary condition for $k = 0$.

Now suppose that the first letter of $\sigma \in S_n(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)$ is $\sigma(1) = 1$ or n. Then $\sigma(1)$ cannot be part of any occurrence of 3-$\bar{1}$-42 or 3-$\bar{1}$-24 in σ. Thus, deletion of $\sigma(1)$, and subtraction of 1 from each component if $\sigma(1) = 1$, yields a bijection with $S_{n-1}(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)$ so that Equation (3.2) follows.

Similarly, suppose that the first letter of $\sigma \in S_n(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)$ is $\sigma(1) = 2$. Then the first column of $R(\sigma)$ must be $\begin{smallmatrix}1\\2\end{smallmatrix}$ regardless of where 1 occurs in σ. Therefore, $R(\sigma)$ has a unique preimage σ if and only if $\sigma = 21 \in \mathfrak{S}_2$ so that Equation (3.3) follows.

Finally, suppose that $\sigma \in S_n(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)$ with $3 \le k \le n$. Since σ avoids 3-$\bar{1}$-42, σ is a RPW by Proposition 3.2, and hence the left prefix of σ from k to 1 is a decreasing subsequence. Let σ' be the permutation obtained by interchanging the values k and $k-1$ in σ. Then the only instances of the patterns 3-$\bar{1}$-42 and 3-$\bar{1}$-24 in σ' must involve both k and $k-1$. Note that the number of σ for which no instances of these patterns are created by interchanging k and $k-1$ is exactly $f(n, k-1)$.

There are then two cases in which an instance of the barred pattern 3-$\bar{1}$-42 or 3-$\bar{1}$-24 will be created in σ' by this interchange:

Case 1. If $k-1$ occurs between $\sigma(1) = k$ and 1 in σ, then $\sigma(2) = k-1$, so interchanging k and $k-1$ creates an instance of the pattern 23-1 via the subsequence $(k-1, k, 1)$ in σ'. Thus, by Theorem 2.3, σ' contains 3-$\bar{1}$-42 from which $\sigma' \in S_n(3\text{-}\bar{1}\text{-}42)$ if and only if $k-1$ occurs after 1 in σ. Note also that if $\sigma(2) = k-1$, then deleting k yields a bijection with permutations in $S_{n-1}(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)$ that start with $k-1$. So the number of permutations counted in Case 1 is exactly $f(n-1, k-1)$.

Case 2. If $k-1$ occurs to the right of 1 in σ, then σ' both contains the subsequence $(k-1, 1, k)$ and avoids the pattern 3-$\bar{1}$-42, so it must also contain the pattern 3-$\bar{1}$-24. If an instance of 3-$\bar{1}$-24 in σ' involves both $k-1$ and k, then $k-1$ and k must play the roles of "3" and "4", respectively. If the value ℓ preceding k is not 1, then the subsequence $(k-1, 1, \ell, k)$ is an instance of 3-1-24, so $(k-1, \ell, k)$ is not an instance of 3-$\bar{1}$-24. Therefore, for σ' to contain 3-$\bar{1}$-24, k must follow 1 in σ', and so $k-1$ follows 1 in σ. If the letter preceding 1 is some $m < k$, then the subsequence $(m, 1, k-1)$ is an instance of 3-$\bar{1}$-24 in σ, which is impossible. Therefore, k must precede 1 in σ, from which σ must start with the initial segment $(k, 1, k-1)$. But then deleting the values k and 1 and then subtracting 1 from each component yields a bijection with permutations in $S_{n-2}(3\text{-}\bar{1}\text{-}42, 3\text{-}\bar{1}\text{-}24)$ that start with $k-2$. It follows that the number of permutations counted in Case 2 is then exactly $f(n-2, k-2)$, which yields Equation (3.4). $\square$

If we denote by

$$\Phi(x, y) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} f(n, k) x^n y^k$$

the bivariate generating function for the sequence $\{f(n, k)\}_{n \ge k \ge 0}$, then Equation (3.4) implies that

$$(1 - y - xy - x^2y^2)\Phi(x, y) = 1 - y - xy + xy^2 - xy^2\Phi(xy, 1) + xy(1 - y - xy)\Phi(x, 1).$$

Moreover, using the kernel method, one can show that

$$x + 1 + \frac{\sqrt{1 + 2x + 5x^2} - x - 1}{2} \cdot F(x) - F\left(\frac{\sqrt{1 + 2x + 5x^2} - x - 1}{2x}\right) = 0$$

where $F(x) = \sum_{n \ge 0} f(n) x^n = \Phi(x, 1)$ is the generating function for the sequence $\{f(n)\}_{n \ge 0}$.

We conclude with the following main enumerative result about invertibility of Patience Sorting.

Theorem 3.10. *Denote by F_n the n^{th} Fibonacci number (with $F_0 = F_1 = 1$) and by*

$$a(n, k) = \sum_{\substack{n_1, \ldots, n_k \ge 0 \\ n_1 + \cdots + n_k = n-k-2}} F_{n_1} F_{n_2} \ldots F_{n_k}$$

convolved Fibonacci numbers for $n \ge k + 2$ (where $a(n, k) := 0$ otherwise). Then, defining

$$X = \begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \\ f(4) \\ \vdots \end{bmatrix}, \quad F = \begin{bmatrix} 1 \\ F_0 \\ F_1 \\ F_2 \\ F_3 \\ \vdots \end{bmatrix}, \quad \text{and} \quad \mathbf{A} = (a(n,k))_{n,k\geq 0} = \begin{bmatrix} 0 & & & & & \\ 0 & 0 & & & & \\ a(2,0) & 0 & 0 & & & \\ a(3,0) & a(3,1) & 0 & 0 & & \\ a(4,0) & a(4,1) & a(4,2) & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

we have that $X = (\mathbf{I} - \mathbf{A})^{-1}F$, where $\mathbf{I}$ is the infinite identity matrix and $\mathbf{A}$ is lower triangular.

PROOF. (Sketch)
From Equations (3.1)–(3.4), we can conjecture an equivalent recurrence where (3.3) and (3.4) are replaced by the following equation (here δ_{nk} is the Kronecker delta function):

$$f(n,k) = \sum_{m=0}^{k-3} c(k,m) f(n-k+m) + \delta_{nk} F_{k-2}, \quad n \geq k \geq 2. \tag{3.5}$$

For this relation to hold, the coefficients $c(k,m)$ must satisfy the following recurrence relation:

$$c(k,m) = c(k-1,m-1) + c(k-1,m) + c(k-2,m), \quad k \geq 2,$$

or, equivalently,

$$c(k-1,m-1) = c(k,m) - c(k-1,m) - c(k-2,m), \quad k \geq 2,$$

with $c(2,0) = 1$ and $c(k,m) = 0$ in the case that $k < 2$, $m < 0$ or $m > k-2$. This implies that the generating function for the sequence $\{c(k,m)\}_{k\geq 0}$ (for each $m \geq 0$) is

$$\sum_{n\geq 0} c(k,m) x^k = \frac{x^{m+2}}{(1-x-x^2)^{m+1}}.$$

It follows that the coefficients $c(k,m) = a(k,m)$ in Equation (3.5) are convolved Fibonacci numbers [**17**] forming the so-called skew Fibonacci-Pascal triangle in the matrix $\mathbf{A} = (a(k,m))_{k,m\geq 0}$. In particular, the sequence of nonzero entries in column $m \geq 0$ of $\mathbf{A}$ is the m^{th} convolution of the sequence $\{F_n\}_{n\geq 0}$.

Combining the expansion of $f(n,n)$ from Equation (3.5) with Equation (3.2), we obtain

$$f(n) = \sum_{m=0}^{n-2} a(n,m) f(m) + F_{n-1},$$

which is equivalent to the matrix equation $X = \mathbf{A}X + F$. Since $\mathbf{I} - \mathbf{A}$ is clearly invertible, the result follows. □

Due to space restrictions, we omit a direct bijective proof of Theorem 3.10 that will be included in the full article.

REMARK 3.11. Note that $\mathbf{A}$ is a strictly lower triangular matrix with zero sub-diagonal. From this it follows that multiplication of a matrix $\mathbf{B}$ by $\mathbf{A}$ shifts the position of the highest nonzero diagonal in $\mathbf{B}$ down by two rows, so $(\mathbf{I} - \mathbf{A})^{-1} = \sum_{n\geq 0} \mathbf{A}^n$ as a Neumann series, and thus all nonzero entries of $(\mathbf{I} - \mathbf{A})^{-1}$ are positive integers.

Finally, one can explicitly compute

$$\mathbf{A} = \begin{bmatrix} 0 & & & & & & & & \\ 0 & 0 & & & & & & & \\ 1 & 0 & 0 & & & & & & \\ 1 & 1 & 0 & 0 & & & & & \\ 2 & 2 & 1 & 0 & 0 & & & & \\ 3 & 5 & 3 & 1 & 0 & 0 & & & \\ 5 & 10 & 9 & 4 & 1 & 0 & 0 & & \\ 8 & 20 & 22 & 14 & 5 & 1 & 0 & 0 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad \Longrightarrow \quad (\mathbf{I}-\mathbf{A})^{-1} = \begin{bmatrix} 1 & & & & & & & & \\ 0 & 1 & & & & & & & \\ 1 & 0 & 1 & & & & & & \\ 1 & 1 & 0 & 1 & & & & & \\ 3 & 2 & 1 & 0 & 1 & & & & \\ 7 & 6 & 3 & 1 & 0 & 1 & & & \\ 21 & 16 & 10 & 4 & 1 & 0 & 1 & & \\ 66 & 50 & 30 & 15 & 5 & 1 & 0 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

from which the first few values of the sequence $\{f(n)\}_{n\geq 0}$ are immediately calculable as

$$1, 1, 2, 4, 9, 23, 66, 209, 718, 2645, 10373, 43090, 188803, 869191, 4189511, 21077302, 110389321\ldots.$$

References

[1] M. H. Albert, S. Linton, and N. Ruškuc. "The Insertion Encoding or Permutations." *Electron. J. Combin.* **12** (2005), #R47, 31 pages.

[2] D. Aldous and P. Diaconis. "Longest Increasing Subsequences: From Patience Sorting to the Baik-Deift-Johansson Theorem." *Bull. Amer. Math. Soc.* **36** (1999), 413–432.

[3] S. Bespamyatnikh and M. Segal. "Enumerating longest increasing subsequences and patience sorting." *Inf. Process. Lett.* **76** (2000), 7–11.

[4] M. Bóna. *Combinatorics of Permutations.* Chapman & Hall/CRC Press, 2004.

[5] A. Burstein and I. Lankham. "Combinatorics of Patience Sorting Piles." *Proceedings of FPSAC 2005.* To appear in *Séminaire Lotharingien de Combinatoire*, Volume 54A. arXiv:math.CO/0506358.

[6] A. Burstein and I. Lankham. "A Geometric Form for the Extended Patience Sorting Algorithm." *Adv. Appl. Math.* **36** (2006), 106-117. arXiv:math.CO/0507031.

[7] A. Claesson. "Generalized Pattern Avoidance." *Europ. J. Combin.* **22** (2001), 961–971.

[8] A. Claesson and T. Mansour. "Counting Occurrences of a Pattern of Type $(1,2)$ or $(2,1)$ in Permutations." *Adv. Appl. Math.* **29** (2002), 293–310.

[9] A. Claesson and T. Mansour. "Enumerating Permutations Avoiding a Pair of Babson-Steingrímsson Patterns." *Ars Combinatoria* **77** (2005), 17–31..

[10] S. Dulucq, S. Gire, and O. Guibert. "A Combinatorial Proof of J. West's Conjecture." *Discrete Math.* **187** (1998), 71–96.

[11] S. Dulucq, S. Gire, and J. West. "Permutations with Forbidden Subsequences and Nonseparable Maps." *Discrete Math.* **153** (1996), 85–103.

[12] C. L. Mallows. "Problem 62-2, Patience Sorting." *SIAM Review* **4** (1962), 148–149; Solution in volume **5** (1963), 375–376.

[13] C. L. Mallows. "Problem 62-2." *SIAM Review* **5** (1963), 375–376.

[14] A. Marcus and G. Tardos. "Excluded Permutation Matrices and the Stanley-Wilf Conjecture." *J. Combin. Th. Ser. A* **107** (2004), 153–160.

[15] A. Price. *Packing densities of layered patterns.* Ph.D. thesis, University of Pennsylvania, 1997.

[16] B. Sagan. *The Symmetric Group, Second Edition.* Graduate Texts in Mathematics 203. Springer-Verlag, 2000.

[17] N. J. A. Sloane. Sequences **A037027** and **A038137**. *The On-Line Encyclopedia of Integer Sequences.* Available online at http://www.research.att.com/∼njas/sequences/

[18] R. Stanley. *Enumerative Combinatorics, Volume 2.* Studies in Advanced Mathematics 62. Cambridge University Press, 1999.

[19] X. G. Viennot. "Une forme géométrique de la correspondance de Robinson-Schensted." In *Combinatoire et Représenatation du Groupe Symétrique*, D. Foata, ed. Lecture Notes in Mathematics 579. Springer-Verlag, 1977, pp. 29–58.

[20] J. West. *Permutations with Forbidden Subsequences and Stack-sortable Permutations.* Ph.D. thesis, M.I.T., 1990.

[21] A. Woo and A. Yong. "When is a Schubert variety Gorenstein?" To appear in *Adv. in Math.* arXiv:math.CO/0409490.

Department of Mathematics, Iowa State University, Ames, IA 50011-2064, USA
E-mail address: `burstein@math.iastate.edu`
URL: `http://www.math.iastate.edu/burstein/`

Department of Mathematics, University of California, Davis, CA 95616-8633, USA
E-mail address: `issy@math.ucdavis.edu`
URL: `http://www.math.ucdavis.edy/∼issy/`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Area of Catalan Paths on a Checkerboard

Szu-En Cheng, Sen-Peng Eu, and Tung-Shan Fu

Abstract. It is known that the area of all Catalan paths of length n is equal to $4^n - \binom{2n+1}{n}$, which coincides with the number of inversions of all 321-avoiding permutations of length $n+1$. In this paper, a bijection between the two sets is established. Meanwhile, a number of interesting bijective results that pave the way to the required bijection are presented.

Résumé. Le fait que la somme des surfaces des chemins Catalan de longueur n est égale à $4^n - \binom{2n+1}{n}$, ce qui est aussi le nombre d'inversions dans toutes les permutations de longueur $n+1$ qui évitent le motif 321, est bien connu. Nous présentons dans cet article une bijection entre ces deux ensembles. Pour ce faire, nous établissons plusieurs résultats bijectifs intermédiaires intéressants.

1. Introduction

Among many other combinatorial structures, the nth Catalan number $c_n = \frac{1}{n+1}\binom{2n}{n}$ enumerates the number of lattice paths, called *Catalan paths of length* n, in the plane $\mathbb{Z}\times\mathbb{Z}$ from $(0,0)$ to (n,n) using north steps $(0,1)$ and east steps $(1,0)$ that never pass below the line $y=x$. Let $\mathcal{C}_n$ denote the set of Catalan paths of length n. A Catalan path is said to be *elevated* if it remains strictly above the line $y=x$ except at the start and end points. The *area* of a Catalan path is defined to be the number of triangles of the region enclosed by the path and the line $y=x$. For example, the area of the path shown in Figure 1 is 13. In [**8**], Merlini et al. derived that the area a_n of all Catalan paths of length n is $a_n = 4^n - \binom{2n+1}{n}$, which is also equal to $\sum_{k=1}^{n} 4^{n-k}c_k$ as shown in [**15**]. Shapiro et al. proved that the area of all elevated Catalan paths of length n is 4^{n-1} [**11**]. There is other literature concerning the area and moments of Catalan paths (e.g., [**3, 6, 9**]).

A permutation $\sigma = \sigma_1\cdots\sigma_n$ of $\{1,\ldots,n\}$, where $\sigma_i = \sigma(i)$, is called a 321-*avoiding permutation of length* n if there are no integers $i<j<k$ such that $\sigma_i > \sigma_j > \sigma_k$ (i.e., every decreasing subsequence is of length at most two). Let $S_n(321)$ denote the set of 321-avoiding permutations of length n. A pair (σ_i, σ_j) is called an *inversion* of σ if $i<j$ and $\sigma_i > \sigma_j$. What catches our attention is that, as reported by Deutsch in [**13**, A008549], the number sequence $\{a_n\}_{n\geq 0} = \{0, 1, 6, 29, 130, 562, \ldots\}$ counts the number of inversions of all 321-avoiding permutations of length $n+1$. The main purpose of this paper is to establish a bijection Π_n between the set of triangles under all Catalan paths of length n and the set of inversions of all 321-avoiding permutations of length $n+1$. The bijection is composed of two major stages (see Theorems 1.1 and 1.2).

To resolve this problem, we color the unit squares in the plane $\mathbb{Z}\times\mathbb{Z}$ in black and white like a checkerboard. A unit square B is colored black if the upper left corner (i,j) of B satisfies the condition that $i+j$ is odd, and white otherwise. For example, there are 1 black square and 3 white squares under the path shown in Figure 1. An intriguing observation is that the number of white squares under all Catalan paths of length $n+1$ is also equal to a_n (see Theorem 2.1). As the first stage of Π_n, the following bijection is one of the major results in this paper.

Theorem 1.1. *There is a bijection between the set of triangles under all Catalan paths of length n and the set of white squares under all Catalan paths of length $n+1$.*

2000 *Mathematics Subject Classification.* 05A15.

Key words and phrases. Catalan paths; permutations; inversions; polyominoes.

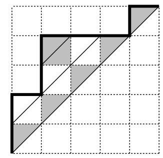


Figure 1. A Catalan path of length 5.

For the second stage of Π_n, we employ a variant of parallelogram polyominoes to establish the following bijection $\Psi_n : \mathcal{C}_n \to S_n(321)$, which is different from the one given by Billy et al. [**2**, page 361].

Theorem 1.2. *There is a bijection Ψ_n between the set $\mathcal{C}_n$ of Catalan paths of length n and the set $S_n(321)$ of 321-avoiding permutations of length n such that there is a one-to-one correspondence between the white squares under a path $\pi \in \mathcal{C}_n$ and the inversions of $\Psi_n(\pi) \in S_n(321)$.*

We organize this paper as follows. Regarding the plane as a checkerboard, we enumerate the black and white squares under Catalan paths in Section 2. The proofs of Theorems 1.1 and 1.2 are given in Sections 3 and 4, respectively. Finally, some enumerative results for variants of parallelogram polyominoes are given in Section 5.

2. Area of Catalan paths on a checkerboard

In this section, we shall enumerate the black and white squares under all Catalan paths of length n by the method of generating functions. The generating function $C = C(z) = \sum_{n\geq 0} c_n z^n$ for Catalan numbers $\{c_n\}_{n\geq 0}$ satisfies the equation $C = 1 + zC^2$. Another useful fact is $[z^n]C^t = \frac{t}{2n+t}\binom{2n+t}{n}$, which is known as the *ballot number* [**4**, p. 21]. Let N and E denote a north step and an east step, respectively. A *block* of a Catalan path is a section of the form $\mathsf{N}\mu\mathsf{E}$, where N is a north step leaving the line $y = x$, E is the first east step returning to the line $y = x$ afterward, and μ is a Catalan path of certain length (possibly empty). A *peak* (resp. *valley*) of a path is formed by a consecutive NE (resp. EN) pair.

Theorem 2.1. *For $n \geq 2$, the following results hold.*

(i) *The number of white squares under all Catalan paths of length n is $4^{n-1} - \binom{2n-1}{n-1}$.*
(ii) *The number of black squares under all Catalan paths of length n is $4^{n-1} - \binom{2n}{n-1}$.*
(iii) *The number of white squares under all elevated Catalan paths of length n is 4^{n-2}.*

Proof. Let $f_{n,k}$ (resp. $g_{n,k}$) denote the number of paths $\pi \in \mathcal{C}_n$ with k white squares (resp. black squares) under π. Define the generating functions $F(t,z) = \sum_{n,k\geq 0} f_{n,k}t^k z^n$, and $G(t,z) = \sum_{n,k\geq 0} g_{n,k}t^k z^n$. Taking partial derivative with respect to t and then setting $t = 1$, we have $\left(\frac{\partial F(t,z)}{\partial t}\right)_{t=1} = \sum_{n\geq 0}\left(\sum_{k\geq 0} k f_{n,k}\right) z^n$ and $\left(\frac{\partial G(t,z)}{\partial t}\right)_{t=1} = \sum_{n\geq 0}\left(\sum_{k\geq 0} k g_{n,k}\right) z^n$, which are the generating functions for the numbers in (i) and (ii), respectively.

A non-trivial path $\pi \in \mathcal{C}_n$ has a factorization $\pi = \mathsf{N}\mu\mathsf{E}\nu$, where E is the first east step that returns to the line $y = x$, and μ and ν are Catalan paths of certain lengths (possibly empty). Since, in the elevated path $\mathsf{N}\mu\mathsf{E}$, the black squares under μ become white and vice versa, we observe that the number of white squares under the first block $\mathsf{N}\mu\mathsf{E}$ of π is equal to the sum of the number of black squares under μ and the length of μ. Moreover, the number of black squares under the first block $\mathsf{N}\mu\mathsf{E}$ of π is equal to the number of white squares under μ. Hence $F(t,z)$ and $G(t,z)$ satisfy the following equations.

$$\begin{cases} F(t,z) = 1 + zG(t,tz)F(t,z), \\ G(t,z) = 1 + zF(t,z)G(t,z). \end{cases} \tag{2.1}$$

Let $F' = \left(\frac{\partial F(t,z)}{\partial t}\right)_{t=1}$ and $G' = \left(\frac{\partial G(t,z)}{\partial t}\right)_{t=1}$. Taking partial derivative with respect to t, setting $t = 1$, and taking into account that $F(1, z) = G(1, z) = C(z)$, we have

$$\begin{cases} F' = z((G' + C'z)C + F'C), \\ G' = z(F'C + G'C). \end{cases} \tag{2.2}$$

Since $C = 1 + zC^2$, $1 - zC = \frac{1}{C}$ and $C' = C^2 + 2zCC'$. Solving (2.2) with $C = \frac{1-\sqrt{1-4z}}{2z}$, we have

$$F' = \frac{z^2C'}{1-2zC} = \frac{1-2z-\sqrt{1-4z}}{2(1-4z)}, \quad \text{and} \quad G' = F' - z^2CC' = F' - \frac{z}{2}(C' - C^2).$$

It follows that

$$[z^n]F' = \frac{1}{2}[z^n]\frac{1}{1-4z} - [z^{n-1}]\frac{1}{1-4z} - \frac{1}{2}[z^n]\frac{1}{\sqrt{1-4z}} = 4^{n-1} - \binom{2n-1}{n-1},$$

and

$$[z^n]G' = [z^n]F' - \frac{1}{2}[z^{n-1}]C' + \frac{1}{2}[z^{n-1}]C^2 = 4^{n-1} - \binom{2n}{n-1}.$$

Hence (i) and (ii) follow.

Let $h_{n,k}$ denote the number of elevated Catalan paths τ of length n with k white squares under τ, and let $H(t, z) = \sum_{n,k\geq 0} h_{n,k}t^kz^n$. We observe that $H(t, z)$ satisfies the equation $H(t, z) = zG(t, tz)$. Let $H' = \left(\frac{\partial H(t,z)}{\partial t}\right)_{t=1}$. By the same method as above, we have $H' = z(G' + C'z)$. Hence $[z^n]H' = [z^{n-1}]G' + [z^{n-2}]C' = 4^{n-2}$, and (iii) follows. $\square$

Similarly, the area of a Catalan path is partitioned into regions of the four types: white up-triangles, white down-triangles, black up-triangles, and black down-triangles. For example, the area of the path in Figure 1 consists of 3 white up-triangles, 3 white down-triangles, 6 black up-triangles, and 1 black down-triangle. The following corollary is an immediate consequence of Theorem 2.1.

Corollary 2.2. *Among the area of all Catalan paths of length n, there are*
- (i) $4^{n-1} - \binom{2n-1}{n-1}$ *white up-triangles,*
- (ii) $4^{n-1} - \binom{2n-1}{n-1}$ *white down-triangles,*
- (iii) 4^{n-1} *black up-triangles, and*
- (iv) $4^{n-1} - \binom{2n}{n-1}$ *black down-triangles.*

Proof. It is clear that (i) and (ii) are equivalent to Theorem 2.1(i), and that (iv) is equivalent to Theorem 2.1(ii). Note that the number of black up-triangles under a path $\pi \in \mathcal{C}_n$ is equal to the number of white squares under the elevated path $\mathsf{N}\pi\mathsf{E} \in \mathcal{C}_{n+1}$. Hence (iii) follows from Theorem 2.1(iii). $\square$

Remarks: In [**1**, page 6], Barcucci et al. derived that the generating function for the number of inversions of all 321-avoiding permutations of length n is $\frac{1-2z-\sqrt{1-4z}}{2(1-4z)}$. Corollary 2.2(iii) has appeared in [**15**, Theorem A], which is obtained by making use of an enumerative result on parallelogram polyominoes in [**11**].

3. Proof of Theorem 1.1

Let $\mathcal{T}_n$ denote the set of ordered pairs (A, π), where $\pi \in \mathcal{C}_n$ and A is a triangle under π, and let $\mathcal{W}_{n+1}$ denote the set of ordered pairs (B, τ), where $\tau \in \mathcal{C}_{n+1}$ and B is a white square under τ. In this section, we shall establish a bijection $\Phi_n : \mathcal{T}_n \to \mathcal{W}_{n+1}$. Let $\mathcal{T}_n$ be partitioned into the following four subsets.

$T_1(n) = \{(A, \pi) \in \mathcal{T}_n |\ A \text{ is a black up-triangle under } \pi\}$,
$T_2(n) = \{(A, \pi) \in \mathcal{T}_n |\ A \text{ is a white up-triangle under } \pi\}$,
$T_3(n) = \{(A, \pi) \in \mathcal{T}_n |\ A \text{ is a white down-triangle under } \pi\}$,
$T_4(n) = \{(A, \pi) \in \mathcal{T}_n |\ A \text{ is a black down-triangle under } \pi\}$.

For any $(A,\pi) \in T_1(n) \cup T_2(n)$ (i.e., A is an up-triangle), A is said to be at position (i,j) if the upper left corner of A is (i,j), and A is said to be on the line $L : x+y = i+j$. For each up-triangle A, the *top triangle* of A is the up-triangle $\widehat{A}$ to the northwest of A at the intersection of π and L.

On the other hand, for any $(B,\tau) \in \mathcal{W}_{n+1}$, B is said to be at position (i,j) if the upper left corner of B is (i,j), and B is said to be on the line $L : x+y = i+j$ (note that $i+j$ is even). For each white square B, the *top box* of B is the white square $\widehat{B}$ to the northwest of B at the intersection of τ and L. Moreover, we say that $\widehat{B}$ is *falling* if the top edge of $\widehat{B}$ coincides with an east step of τ, and *rising* otherwise. For any $(B,\tau) \in \mathcal{W}_{n+1}$, B is called a *downhill square* (resp. *uphill square*) of τ if the top box of B is falling (resp. rising). Let $\mathcal{W}_{n+1}$ be partitioned into the following four subsets.

$$\begin{aligned}
W_1(n+1) &= \{(B,\tau) \in \mathcal{W}_{n+1} |\ B \text{ is a downhill square in the first block of } \tau\},\\
W_2(n+1) &= \{(B,\tau) \in \mathcal{W}_{n+1} |\ \text{the first block } \beta \text{ of } \tau \text{ is of length 1, i.e., } \beta = \mathsf{NE}\},\\
W_3(n+1) &= \{(B,\tau) \in \mathcal{W}_{n+1} |\ B \text{ is an uphill square in the first block of } \tau\},\\
W_4(n+1) &= \{(B,\tau) \in \mathcal{W}_{n+1} |\ \text{the first block } \beta \text{ of } \tau \text{ is of length} > 1, \text{ and } B \text{ is not in } \beta\}.
\end{aligned}$$

For each i ($1 \le i \le 4$), we shall establish a bijection $\Phi_{n,i} : T_i(n) \to W_i(n+1)$ (see Propositions 3.1-3.4). Then Φ_n is established by the refinement $\Phi_n|_{T_i(n)} = \Phi_{n,i}$, for $1 \le i \le 4$, and hence Theorem 1.1 is proved.

PROPOSITION 3.1. *There is a bijection $\Phi_{n,1}$ between $T_1(n)$ and $W_1(n+1)$.*

PROOF. Given a pair $(A,\pi) \in T_1(n)$, say A is at (i,j), we have $i+j = 2h-1$, for some h ($h \ge 1$). Let $\widehat{A}$ be the top triangle of A. We factorize π as $\pi = \mu\nu$, where μ goes from the origin to the upper left corner of $\widehat{A}$, and ν is the remaining part of π. Define a mapping $\Phi_{n,1}$ that carries (A,π) into $\Phi_{n,1}((A,\pi)) = (B,\tau)$, where $\tau = \mathsf{N}\mu\mathsf{E}\nu \in \mathcal{C}_{n+1}$ (i.e., with a north step N attached to the beginning and an east step E inserted between μ and ν) and B is the white square at $(i,j+1)$. Note that the top box $\widehat{B}$ of B is at the end point of μ, and that E is the top edge of $\widehat{B}$. Hence $\widehat{B}$ is a falling box and B is downhill. Hence $\Phi_{n,1}((A,\pi)) \in W_1(n+1)$.

To find $\Phi_{n,1}^{-1}$, given a pair $(B,\tau) \in W_1(n+1)$, say B is at (i,j), we have $i+j = 2h'$, for some h'. Since B is a downhill square, the top box $\widehat{B}$ of B is a falling box. We factorize τ as $\tau = \mathsf{N}\mu\mathsf{E}\nu$, where N is the first step of τ, E is the top edge of $\widehat{B}$, μ is the section between N and E, and ν is the remaining part of τ. Since B is in the first block of τ, μ remains above the line $y = x+1$ and hence $\mu\nu \in \mathcal{C}_n$. Hence $\Phi_{n,1}^{-1}((B,\tau)) = (A,\pi) \in T_1(n)$, where $\pi = \mu\nu$ and A is the black up-triangle at $(i,j-1)$. $\square$

For example, on the left of Figure 2 is a pair $(A,\pi) \in T_1(9)$, where A is at $(2,5)$. The top triangle $\widehat{A}$ of A in π is at $(1,6)$. Note that A is the second up-triangle on the line $x+y = 7$ from $\widehat{A}$. The corresponding pair $\Phi_{9,1}((A,\pi)) = (B,\tau) \in W_1(10)$ is shown on the right of Figure 2, where B is at $(2,6)$ and $\widehat{B}$ is at $(1,7)$. Note that B is the second square on the line $x+y = 8$ from $\widehat{B}$.

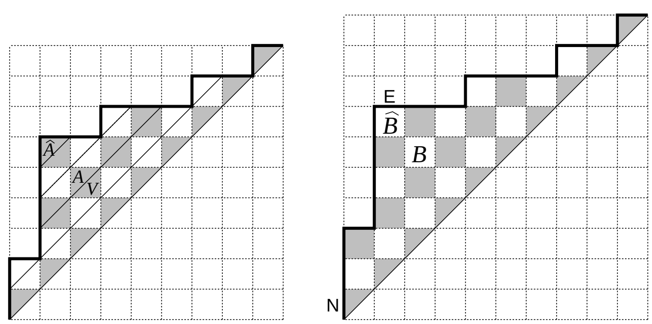


FIGURE 2. A pair $(A,\pi) \in T_1(9)$ and the corresponding pair $\Phi_{9,1}((A,\pi)) = (B,\tau) \in W_1(10)$.

PROPOSITION 3.2. *There is a bijection $\Phi_{n,2}$ between $T_2(n)$ and $W_2(n+1)$.*

PROOF. Given a pair $(A,\pi) \in T_2(n)$, say A is at (i,j), we have $i+j = 2h$, for some h ($h \ge 1$). Define a mapping $\Phi_{n,2} : T_2(n) \to W_2(n+1)$ that carries (A,π) into $\Phi_{n,2}((A,\pi)) = (B,\tau) \in W_2(n+1)$, where $\tau = \mathsf{NE}\pi \in \mathcal{C}_{n+1}$ and B is the white square at $(i+1,j+1)$. It is easy to find $\Phi_{n,2}^{-1}$ by a reverse process. $\square$

For example, on the left of Figure 3 is a pair $(A, \pi) \in T_2(9)$, where A is at $(4, 6)$. The corresponding pair $\Phi_{9,2}((A, \pi)) = (B, \tau) \in W_2(10)$ is shown on the right of Figure 3, where B is at $(5, 7)$.

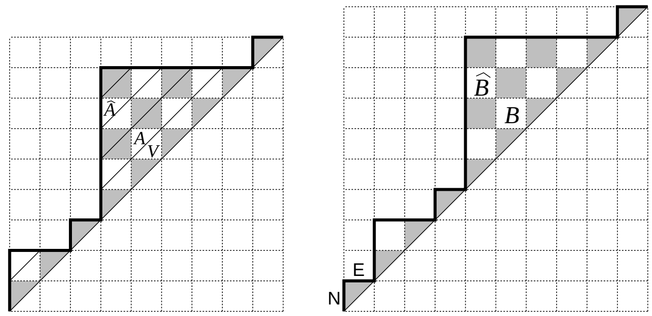


FIGURE 3. A pair $(A, \pi) \in T_2(9)$ and the corresponding pair $\Phi_{9,2}((A, \pi)) = (B, \tau) \in W_2(10)$.

PROPOSITION 3.3. *There is a bijection $\Phi_{n,3}$ between $T_3(n)$ and $W_3(n + 1)$.*

PROOF. Given a pair $(V, \pi) \in T_3(n)$, say the lower right corner of V is (i, j), we have $i + j = 2h$, for some h ($h \geq 1$). Let A be the white up-triangle at $(i - 1, j + 1)$. Clearly, $(A, \pi) \in T_2(n)$. We shall use the mapping $\Phi_{n,2}$ given in Proposition 3.2 as an intermediate stage to establish $\Phi_{n,3}$.

Let $\Phi_{n,2}((A, \pi)) = (B, \tau) \in W_2(n + 1)$. Then B is at $(i, j + 2)$. Let $\widehat{B}$ be the top box of B in τ, and let B be the kth square on the line $L : x + y = i + j + 2$ from $\widehat{B}$, for some k. We factorize τ as $\tau = \mathsf{NE}\mu\beta\nu$, where NE is the first block of τ, β is the block containing B, μ is the section between the first block and β, and ν is the remaining part of τ. Moreover, β is further factorized as $\beta = \alpha\gamma$, where α goes from the beginning of β to the upper left corner of $\widehat{B}$, and γ is the remaining part of β. Let p_α denote the end point of α. Define a mapping $\Phi_{n,3}$ that carries (V, π) into $\Phi_{n,3}((V, \pi)) = (C, \omega)$, where $\omega = \alpha\mathsf{N}\mu\mathsf{E}\gamma\nu$, $\widehat{C}$ is the top box at p_α in ω, and C is the kth square from $\widehat{C}$. Since α is followed by a north step, $\widehat{C}$ is a rising box and C is uphill. Moreover, C is in the first block $\alpha\mathsf{N}\mu\mathsf{E}\gamma$ of ω. Hence $\Phi_{n,3}((V, \pi)) \in W_3(n + 1)$.

To find $\Phi_{n,3}^{-1}$, given a pair $(C, \omega) \in W_3(n+1)$, say C is at (i, j), we have $i+j = 2h'$, for some h'. Let $\widehat{C}$ be the top box of C in ω, say $\widehat{C}$ is at (i', j'), and let C be the k'th square on the line $x+y = 2h'$ from $\widehat{C}$. First, we factorize ω as $\omega = \beta\nu$, where β is the first block of ω, and ν is the remaining part of ω. Since C is an uphill square in β, $\widehat{C}$ is a rising box and β has a factorization $\beta = \alpha\mathsf{N}\mu\mathsf{E}\gamma$, where α goes from the origin to the upper left corner of $\widehat{C}$, E is the first step after $\widehat{C}$ that returns to the line $y = x+j'-i'$, and γ is the remaining part of β. Let p_α denote the end point of α. Locate the pair (B, τ), where $\tau = \mathsf{NE}\mu\alpha\gamma\nu$, $\widehat{B}$ is the top box at p_α in τ, and B is the k'th square from $\widehat{B}$. Since the first block of τ is of length 1, $(B, \tau) \in W_2(n + 1)$. Let $\Phi_{n,2}^{-1}((B, \tau)) = (A, \pi) \in T_2(n)$. Then we retrieve the required pair $\Phi_{n,3}^{-1}((C, \omega)) = (V, \pi) \in T_3(n)$ from (A, π), where V is the white down-triangle that shares an edge with A. □

For example, given the pair $(V, \pi) \in T_3(9)$ shown on the left of Figure 3, where the lower right corner of V is $(5, 5)$. Let A be the white up-triangle at $(4, 6)$. The intermediate pair $\Phi_{9,2}((A, \pi)) = (B, \tau)$ is shown on the left of Figure 4. Factorize τ as $\tau = \mathsf{NE}\mu\beta\nu$, where $\mathsf{N} = 1$, $\mathsf{E} = 2$, $\mu = (3, \ldots, 8)$, $\beta = (9, \ldots, 18)$, and $\nu = (19, 20)$. Moreover, β is further factorized as $\beta = \alpha\gamma$, where $\alpha = (9, 10, 11, 12)$ and $\gamma = (13, \ldots, 18)$. The corresponding pair $\Phi_{9,3}((V, \pi)) = (C, \omega) \in W_3(10)$ is shown on the right of Figure 4, where $\omega = \alpha\mathsf{N}\mu\mathsf{E}\gamma\nu$, and C is at $(1, 3)$.

PROPOSITION 3.4. *There is a bijection $\Phi_{n,4}$ between $T_4(n)$ and $W_4(n + 1)$.*

PROOF. Given a pair $(V, \pi) \in T_4(n)$, say the lower right corner of V is (i, j), we have $i + j = 2h + 1$, for some h ($h \geq 1$). Let A be the up-triangle at $(i - 1, j + 1)$. Clearly, $(A, \pi) \in T_1(n)$. We shall use the mapping $\Phi_{n,1}$ given in Proposition 3.1 as an intermediate stage to establish $\Phi_{n,4}$. Let $\Phi_{n,1}((A, \pi)) = (B, \tau) \in W_1(n + 1)$. Then B is at $(i - 1, j + 2)$. Let $\widehat{B}$ be the top box of B in τ, and let B be the kth square on the line $L : x + y = i + j + 1$ from $\widehat{B}$, for some k. Since B is at $(i - 1, j + 2)$ and $j > i$, B is above the line

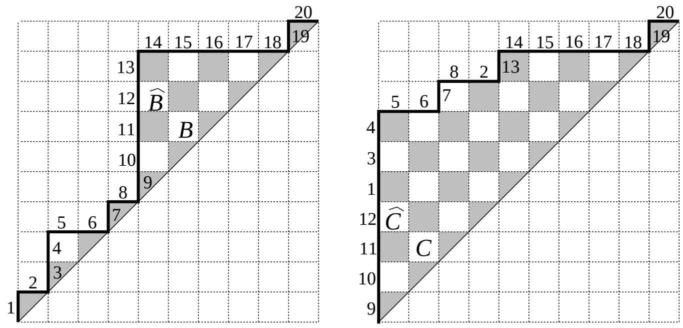


FIGURE 4. The pairs $\Phi_{9,2}((A,\pi)) = (B,\tau) \in W_2(10)$ and $\Phi_{9,3}((V,\pi)) = (C,\omega) \in W_3(10)$ that are associated with the pairs $(A,\pi) \in T_2(9)$ and $(V,\pi) \in T_3(9)$ shown on the left of Figure 3.

$y = x+2$. First, we factorize τ as $\tau = \beta\nu$, where β is the first block of τ and ν is the remaining part of τ. Next, β is further factorized as $\beta = \mathsf{NN}\mu_1\mu_2$, where μ_1 goes from $(0,2)$ to the first step after $\widehat{B}$ that returns to the line $L_2 : y = x+2$, and μ_2 is the remaining part of β. Form a new path $\beta' = \mathsf{NN}\mu_2\mu_1$ from β by switching μ_1 and μ_2. Note that $\mathsf{NN}\mu_2$ is the first block of β', and that B is in μ_1. Moreover, the section μ_1 of β' might have a valley on the line $L_1 : y = x-1$ (in front of $\widehat{B}$). There are two cases.

Case I. μ_1 has no valley on the line L_1. We define a mapping $\Phi_{n,4}$ that carries (V,π) into $\Phi_{n,4}((V,\pi)) = (C,\omega)$, where $\omega = \beta'\nu = \mathsf{NN}\mu_2\mu_1\nu$, and C is the white square B in μ_1. Since the first block $\mathsf{NN}\mu_2$ is of length at least 2, $\Phi_{n,4}((V,\pi)) \in W_4(n+1)$. It is worth mentioning that C is a downhill square since B is downhill in μ_1.

Case II. μ_1 has at least one valley on the line L_1. Then we factorize μ_1 as $\mu_1 = \lambda\mathsf{EN}\alpha\gamma$, where EN is the last valley on the line L_1, α goes from the end point of N to the upper left corner of $\widehat{B}$, and γ is the remaining part of μ_1. Let p_α be the end point of α. The mapping $\Phi_{n,4}$ is then defined by carrying (V,π) into $\Phi_{n,4}((V,\pi)) = (C,\omega)$, where $\omega = \mathsf{NN}\mu_2\alpha\mathsf{N}\lambda\mathsf{E}\gamma\nu$, $\widehat{C}$ is the top box at p_α in ω, and C is the kth square from $\widehat{C}$. Since the first block $\mathsf{NN}\mu_2$ of ω is of length at least 2 and since C is not in the first block, $\Phi_{n,4}((V,\pi)) \in W_4(n+1)$. Note that, since α is followed by a north step, $\widehat{C}$ is a rising box and C is uphill.

To find $\Phi_{n,4}^{-1}$, given a pair $(C,\omega) \in W_4(n+1)$, say C is at (i,j), for some $i \ge 2, j \ge 4$. First, we factorize ω as $\omega = \mathsf{NN}\mu_2\beta\nu$, where $\mathsf{NN}\mu_2$ is the first block of ω, β is the section that ends with the block containing C, and ν is the remaining part of ω. There are two cases.

Case i. C is a downhill square. We locate the pair (B,τ), where $\tau = \mathsf{NN}\beta\mu_2\nu$, and B is the square C in β. We observe that B is a downhill square in the first block $\mathsf{NN}\beta\mu_2$ of ω. Hence $(B,\tau) \in W_1(n+1)$.

Case ii. C is an uphill square. The top box $\widehat{C}$ of C in β is a rising box, say $\widehat{C}$ is at (i',j'). Let C be the k'th square on the line $x+y = i+j$ from $\widehat{C}$. We further factorize β as $\beta = \alpha\mu_1\mathsf{E}\gamma$, where α goes from the beginning of β to the upper left corner of $\widehat{C}$, E is the first east step that goes from the line $y = x+j'-i'$ to the line $y = x+j'-i'-1$, and γ is the remaining part of β. Let p_α denote the end point of α. Since $\widehat{C}$ is a rising box, μ_1 starts with a north step. Factorize μ_1 as $\mu_1 = \mathsf{N}\lambda\mathsf{E}$, and let $\mu_1' = \lambda\mathsf{EN}$. We locate the pair (B,τ), where $\tau = \mathsf{NN}\mu_1'\alpha\mathsf{E}\gamma\mu_2\nu$, $\widehat{B}$ is the top box at p_α in τ, and B is the k'th square from $\widehat{B}$. Since α is followed by an east step, $\widehat{B}$ is a falling box and B is a downhill square in the first block $\mathsf{NN}\mu_1'\alpha\mathsf{E}\gamma\mu_2$ of τ. Hence $(B,\tau) \in W_1(n+1)$.

For both cases, let $\Phi_{n,1}^{-1}((B,\tau)) = (A,\pi) \in T_1(n)$. Then we retrieve the required pair $\Phi_{n,4}^{-1}((C,\omega)) = (V,\pi) \in T_4(n)$ from (A,π), where V is the black down-triangle that shares an edge with A. $\square$

For example, given the pair $(V,\pi) \in T_4(9)$ shown on the left of Figure 2, where the lower right corner of V is $(3,4)$. Let A be the up-triangle at $(2,5)$. The intermediate pair $\Phi_{9,1}((A,\pi)) = (B,\tau) \in W_1(10)$ is shown on the left of Figure 5. First, factorize $\tau = \beta\nu$, where $\beta = (1,\dots,18)$ and $\nu = (19,20)$. Next, β is further factorized as $\beta = \mathsf{N}_1\mathsf{N}_2\mu_1\mu_2$, where $\mathsf{N}_1 = 1$, $\mathsf{N}_2 = 2$, $\mu_1 = (3,\dots,14)$ and $\mu_2 = (15,16,17,18)$. Let $\beta' = \mathsf{N}_1\mathsf{N}_2\mu_2\mu_1$. On the right of Figure 5 is the path $\beta'\nu$. We observe that $\mathsf{N}_1\mathsf{N}_2\mu_2$ is the first block of β', and that μ_1 has no valley on the line $L_1 : y = x-1$. Hence we have the corresponding pair $\Phi_{9,4}((V,\pi)) = (C,\omega) \in W_4(10)$, where $\omega = \beta'\nu = \mathsf{N}_1\mathsf{N}_2\mu_2\mu_1\nu$ and C is at $(5,7)$.

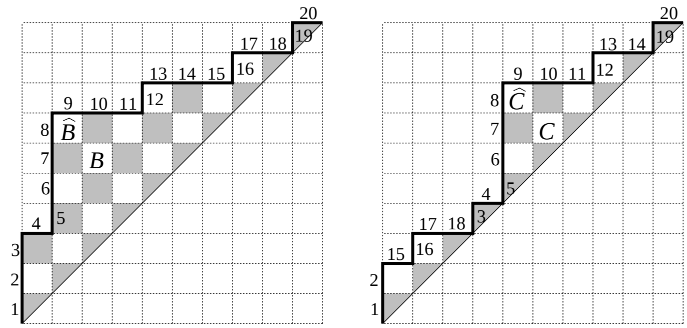


FIGURE 5. The pairs $\Phi_{9,1}((A,\pi)) = (B,\tau) \in W_1(10)$ and $\Phi_{9,4}((V,\pi)) = (C,\omega) \in W_4(10)$ that are associated with the pairs $(A,\pi) \in T_1(9)$ and $(V,\pi) \in T_4(9)$ shown on the left of Figure 2.

For the latter case, consider the pair $(V,\pi) \in T_4(11)$ shown on the left of Figure 6, where the lower right corner of V is $(7,8)$. Let A be the up-triangle at $(6,9)$. The intermediate pair $\Phi_{11,1}((A,\pi)) = (B,\tau) \in W_1(12)$ is shown on the right of Figure 6. First, τ is factorized as $\tau = \beta\nu$, where $\beta = (1,\dots,22)$ and $\nu = (23,24)$. Next, β is factorized as $\beta = \mathsf{N}_1\mathsf{N}_2\mu_1\mu_2$, where $\mu_1 = (3,\dots,18)$ and $\mu_2 = (19,20,21,22)$. Let $\beta' = \mathsf{N}_1\mathsf{N}_2\mu_2\mu_1$. On the left of Figure 7 is the path $\beta'\nu$. We observe that $\mathsf{N}_1\mathsf{N}_2\mu_2$ is the first block of β', and that μ_1 has two valleys on the line $L_1 : y = x - 1$. Hence μ_1 is further factorized as $\mu_1 = \lambda\mathsf{E}_3\mathsf{N}_3\alpha\gamma$, where $\mathsf{E}_3 = 11$ and $\mathsf{N}_3 = 12$ form the last valley on the line L_1 of μ_1, $\lambda = (3,\dots,10)$, $\alpha = (13,14,15,16)$, and $\gamma = (17,18)$. With N_3 moved in front of λ, we have $\mathsf{N}_3\lambda\mathsf{E}_3 = (12,3,4,\dots,11)$. The corresponding pair $\Phi_{11,4}((V,\pi)) = (C,\omega) \in W_4(12)$ is shown on the right of Figure 7, where $\omega = \mathsf{N}_1\mathsf{N}_2\mu_2\alpha\mathsf{N}_3\lambda\mathsf{E}_3\gamma\nu$ and C is at $(4,6)$.

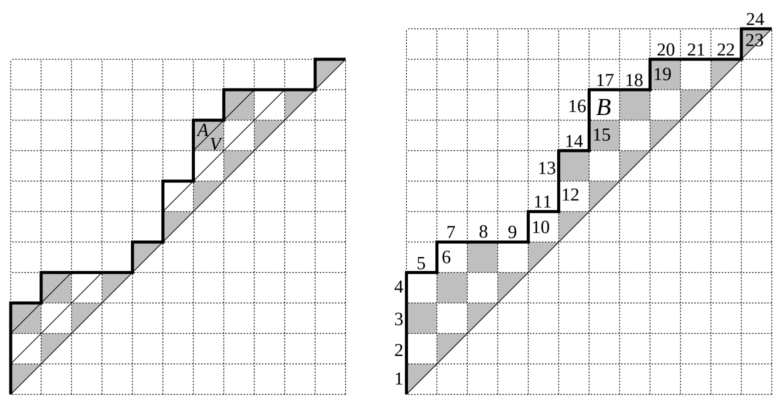


FIGURE 6. A pair $(V,\pi) \in T_4(11)$ and the corresponding pair $\Phi_{11,1}((A,\pi)) = (B,\tau) \in W_1(12)$.

4. Proof of Theorem 1.2

In this section, making use of a variant of parallelogram polyominoes, we shall prove Theorem 1.2 in two stages (see Propositions 4.1 and 4.3).

A *shortened polyomino* is formed by a pair (P,Q) of paths using north steps $(0,1)$ and east steps $(1,0)$ that start from the origin, end in a common point, and satisfy the following conditions

(H1) P never goes below Q, and
(H2) there are no north steps of P and Q overlapped.

The *perimeter* of a polyomino is twice of the length of its paths, and its *area* is the number of unit squares enclosed. As another occurrence of Catalan numbers, it is known that the number of shortened polyominoes of perimeter $2n$ is c_n (see [**7**, Section 5]). The shortened polyominoes of perimeter 6 are shown in Figure 8.

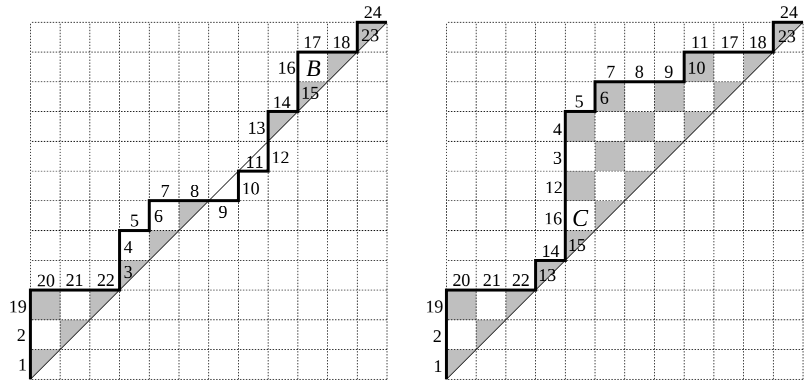


FIGURE 7. The intermediate path $\beta'\nu$ and the corresponding pair $\Phi_{11,4}((V,\pi)) = (C,\omega) \in W_4(12)$.

Making use of a similar argument to the one in [**15**, Theorem A], we prove the following proposition. Here, the end point of a step is said to be at *level* h if it is on the line $y = x + h$, for some integer h.

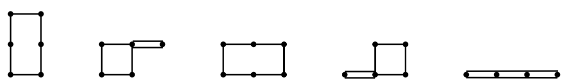


FIGURE 8. The shortened polyominoes with perimeter 6.

PROPOSITION 4.1. *There is a bijection Ω_n between the set $\mathcal{C}_n$ of Catalan paths of length n and the set $\mathcal{H}_n$ of shortened polyominoes of perimeter $2n$ such that there is a one-to-one correspondence between the white squares under a path $\omega \in \mathcal{C}_n$ and the squares in $\Omega_n(\omega) \in \mathcal{H}_n$.*

PROOF. Given a path $\omega \in \mathcal{C}_n$, let P (resp. Q) be the path formed by the even steps (resp. odd steps) of ω, and let Q^* be the path obtained from Q by interchanging north steps and east steps. Define a mapping Ω_n by carrying ω into $\Omega_n(\omega) = (P, Q^*)$. Let $P = p_1 \cdots p_n$ and $Q^* = q_1 \cdots q_n$. Clearly, P and Q^* have the same number of north steps (as well as east steps), and P always remains above Q^* since the distance between the end points of p_i and q_i $(1 \le i \le n)$ is one half of the level of the end point of p_i in ω. Moreover, whenever two steps in (P, Q^*) overlap, they are east steps since their corresponding steps in ω form a peak at level 1. Hence $\Omega_n(\omega) \in \mathcal{H}_n$. To find Ω_n^{-1}, it is simply to reverse the procedure.

We observe that each white square under ω is on the line $x + y = 2h$, for some h $(1 \le h \le n-1)$, and that the number of white squares under ω on the line $x + y = 2h$ is equal to the number of squares on the line $x + y = h$ in $\Omega_n(\omega)$. Hence there is a one-to-one correspondence between the set of white squares under ω and the set of squares in $\Omega_n(\omega)$ such that the kth square on the line $x + y = 2h$ from its top box under ω corresponds to the kth square on the line $x + y = h$ (from upper left to lower right) in $\Omega_n(\omega)$. $\square$

We remark that the actual distance between the end points of p_i and q_i in (P, Q^*) has a factor $\sqrt{2}$, but we omit it.

For example, given the pair $(C, \omega) \in \mathcal{W}_{10}$ shown on the right of Figure 5. The shortened polyomino $\Omega_{10}(\omega) = (P, Q^*)$ is shown on the left of Figure 9, where $P = \mathsf{NNEENNENEE}$ consists of the even steps of ω and $Q^* = \mathsf{ENNEEENNNE}$ is obtained from the odd steps $Q = \mathsf{NEENNNEEEN}$ of ω by interchanging north steps and east steps. The white square C under ω is carried into the square D in $\Omega_{10}(\omega)$.

Let us turn to the second half of the proof of Theorem 1.2. Let S_n be the set of permutations of $[n] := \{1, \ldots, n\}$. We write $\sigma = \sigma_1 \cdots \sigma_n \in S_n$, where $\sigma_i = \sigma(i)$. For a $\sigma \in S_n$, an *excedance* (resp. *weak excedance*) of σ is an integer $i \in [n-1]$ such that $\sigma_i > i$ (resp. $\sigma_i \ge i$). Here the element σ_i is called an *excedance letter* (resp. *weak excedance letter*). *Non-weak excedances* and *non-weak excedance letters* are defined in the obvious way, in terms of i and σ_i, such that $\sigma_i < i$. Let $E(\sigma)$ be the set of excedances of σ, and let $\mathsf{inv}(\sigma)$ be the number of inversions of σ. The following characterization of 321-avoiding permutations was given by R. Simion [**12**, Lemma 5.6] (see also [**10**, Proposition 2.3]).

LEMMA 4.2. *A permutation σ is 321-avoiding if and only if*

$$\mathit{inv}(\sigma) = \sum_{k \in E(\sigma)} (\sigma_k - k).$$

PROPOSITION 4.3. *There is a bijection Υ_n between the set $\mathcal{H}_n$ of shortened polyominoes of perimeter $2n$ and the set $S_n(321)$ of 321-avoiding permutations of length n such that there is a one-to-one correspondence between the squares in a polyomino $(P, Q) \in \mathcal{H}_n$ and the inversions of $\Upsilon_n((P, Q)) \in S_n(321)$.*

PROOF. Given a shortened polyomino $(P, Q) \in \mathcal{H}_n$, let $P = p_1 \cdots p_n$ and $Q = q_1 \cdots q_n$. Let the steps $p_1, \ldots, p_n$ of P be labeled from 1 to n. For each i $(1 \le i \le n)$, we assign the ith step q_i of Q the label z_i of the opposite step across the polyomino. The mapping Υ_n is defined by carrying (P, Q) into $\Upsilon_n((P, Q)) = z_1 \cdots z_n$. Since the labels of the north steps (resp. east steps) of Q are increasing, every decreasing subsequence of $\Upsilon_n((P, Q))$ is of length at most two. Hence $\Upsilon_n((P, Q)) \in S_n(321)$.

To find Υ_n^{-1}, we shall retrieve a shortened polyomino $\Upsilon_n^{-1}(\sigma)$ for any $\sigma = \sigma_1 \cdots \sigma_n \in S_n(321)$. Let $\{j_1, \ldots, j_t\}$ be the set of weak excedances of σ (i.e., $\sigma(j_i) \ge j_i$, for $1 \le i \le t$). For each i $(1 \le i \le t)$, put an east step E_i at height $y = \sigma(j_i) - i$ as the top of the ith column of $\Upsilon_n^{-1}(\sigma)$. The upper path of $\Upsilon_n^{-1}(\sigma)$ goes from $(0, 0)$ to the end point of E_t containing $\mathsf{E}_1, \ldots, \mathsf{E}_t$. On the other hand, for each i $(1 \le i \le t)$, put an east step E'_i at height $y = j_i - i$ as the bottom of the ith column of $\Upsilon_n^{-1}(\sigma)$. The lower path of $\Upsilon_n^{-1}(\sigma)$ goes from $(0, 0)$ to the end point of E_t containing $\mathsf{E}'_1, \ldots, \mathsf{E}'_t$. Since $\sigma(j_i) \ge j_i \ge i$ $(1 \le i \le t)$, $\Upsilon_n^{-1}(\sigma) \in \mathcal{H}_n$ is well-defined.

Note that there are $\sigma(j_i) - j_i$ squares in the ith column of $\Upsilon_n^{-1}(\sigma)$, and that, by Lemma 4.2, $\mathsf{inv}(\sigma) = \sum_{i=1}^{t} (\sigma(j_i) - j_i)$. Hence the number of inversions of σ is equal to the number of squares in $\Upsilon_n^{-1}(\sigma)$. Moreover, the columns (resp. rows) of $\Upsilon_n^{-1}(\sigma)$ are labeled with weak excedance letters (resp. non-weak excedance letters) increasingly. Since each square D in $\Upsilon_n^{-1}(\sigma)$ is the intersection of the column with label σ_i and the row with label σ_j, for some excedance i and non-weak excedance j, there is one-to-one correspondence between the squares in $\Upsilon_n^{-1}(\sigma)$ and the inversions of σ such that D is carried into the inversion (σ_i, σ_j). □

For example, in Figure 9, the labeling of the shortened polyomino (P, Q^*) on the left is shown in the center. The corresponding permutation $\sigma = \Upsilon_{10}((P, Q^*)) = 312479568a$ $(a = 10)$ can be obtained from the labeling of the lower path Q^*. Note that the square D in (P, Q^*) is carried into the inversion $(\sigma_6, \sigma_7) = (9, 5)$ of $\Upsilon_{10}((P, Q^*))$. To show $\Upsilon_{10}^{-1}(\sigma)$, note that the weak excedances of σ are $\{1, 4, 5, 6, 10\}$, i.e., $\sigma_1 = 3$, $\sigma_4 = 4$, $\sigma_5 = 7$, $\sigma_6 = 9$, and $\sigma_{10} = 10$. The east steps on the upper path and lower path of $\Upsilon_{10}^{-1}(\sigma)$ are shown on the right of Figure 9.

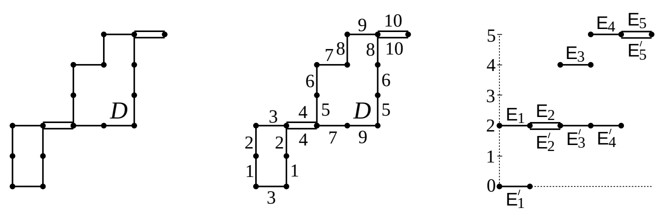


FIGURE 9. The shortened polyomino $\Omega_{10}(\omega)$ associated with the path $\omega \in \mathcal{C}_{10}$ in Figure 5, and its labeling.

By the composition $\Psi_n = \Upsilon_n \circ \Omega_n$, Theorem 1.2 is proved. Hence, by Theorems 1.1 and 1.2, we establish the required bijection between the area of all Catalan paths of length n and the inversions of all 321-avoiding permutations of length $n + 1$.

5. Some enumerative results for parallelogram polyominoes

In the previous section, we introduced a variant of parallelogram polyominoes, called shortened polyominoes. A *parallelogram polyomino* is a pair of non-intersecting paths that starts from the origin and ends in a common point. A *shrunk polyomino* is a pair of paths that start from the origin and end in a common point such that one path never goes below the other. In fact, a shortened polyomino of perimeter $2n$ can be obtained from a parallelogram polyomino (P, Q) of perimeter $2n+2$ by deleting the initial (north) step of the upper path P and deleting the final (north) step of the lower path Q. Moreover, a shrunk polyomino of perimeter $2n-2$ can be obtained from a shortened polyomino (P', Q') of perimeter $2n$ by further deleting the final (east) step of the upper path P' and deleting the first (east) step of the lower path Q'. Figure 10 shows polyominoes of the three types for the case of $n = 3$. Refer also to [**14**, Exercise 6.19(l)(m)].

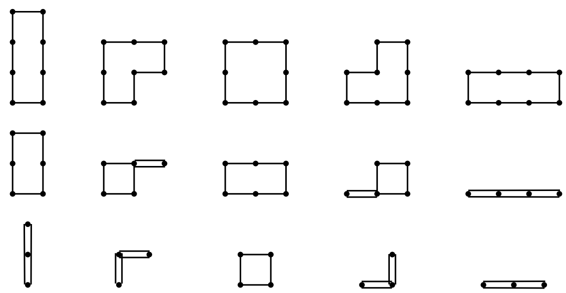


Figure 10. The polyominoes of three kinds for the case $n = 3$.

A bijection Ω'_n between Catalan paths of length n and parallelogram polyominoes of perimeter $2n+2$ can be obtained from the bijection Ω_n in Proposition 4.1 as follows. Given a path $\omega \in \mathcal{C}_n$, let $(P, Q^*) = \Omega_n(\omega) \in \mathcal{H}_n$ be the corresponding shortened polyomino. The bijection Ω'_n is defined by $\Omega'_n(\omega) = (\mathsf{N}P, Q^*\mathsf{N})$, which is obtained from $\Omega_n(\omega)$ with a north step attached to the beginning of the upper path and a north step attached to the end of the lower path. We remark that this bijection is different from the one given by Delest and Viennot in [**5**, Section 4] and the one given by Reifegerste in [**10**, Theorem 3.10]. The following proposition is also an immediate consequence of the bijection Ω_n.

Proposition 5.1. *There is a bijection Θ_n between the set $\mathcal{C}_n$ of Catalan paths of length n and the set $\mathcal{R}_n$ of shrunk polyominoes of perimeter $2n-2$ such that there is a one-to-one correspondence between the black squares under a path $\pi \in \mathcal{C}_n$ and the squares in $\Theta_n(\pi) \in \mathcal{R}_n$.*

Proof. Given a path $\pi \in \mathcal{C}_n$, consider the shortened polyomino $\Omega_n(\pi) = (P, Q^*)$ under the mapping Ω_n in Proposition 4.1. Let $P = p_1 \cdots p_n$ and $Q^* = q_1 \cdots q_n$. There is an immediate bijection $\Theta_n : \mathcal{C}_n \to \mathcal{R}_n$ that carries π into $\Theta_n(\pi) = (P', Q^{*\prime}) \in \mathcal{R}_n$, where $P' = p_1 \cdots p_{n-1}$ and $Q^{*\prime} = q_2 \cdots q_n$. Moreover, the number of black squares under π on the line $x + y = 2h+1$, $(1 \le h \le n-2)$ is equal to the distance between the end points of p_h and q_{h+1} in $(P', Q^{*\prime})$. Hence there is a one-to-one correspondence between the black squares under π and the squares in $\Theta_n(\pi)$. □

The following bijective result can be obtained by the same argument as in the proof of Proposition 4.1, which appeared implicitly in [**15**, Theorem A].

Proposition 5.2. *There is a bijection Λ_n between the set $\mathcal{E}_n$ of elevated Catalan paths of length $n+1$ and the set $\mathcal{P}_n$ of parallelogram polyominoes of perimeter $2n+2$ such that there is a one-to-one correspondence between the white squares under a path $\pi \in \mathcal{E}_n$ and the squares in $\Lambda_n(\pi) \in \mathcal{P}_n$.*

By Theorem 2.1 and Propositions 4.1, 5.1, and 5.2, we deduce the enumerative results on the area of the various polyominoes.

Theorem 5.3. *For $n \ge 2$, the following results hold.*

(i) *The area of all shortened polyominoes of perimeter $2n$ is $4^{n-1} - \binom{2n-1}{n-1}$.*

(ii) *The area of all shrunk polyominoes of perimeter* $2n-2$ *is* $4^{n-1} - \binom{2n}{n-1}$.
(iii) *The area of all parallelogram polyominoes of perimeter* $2n+2$ *is* 4^{n-1}.

A *2-Motzkin path of length* n is a lattice path from $(0,0)$ to $(n,0)$ that never goes below the x-axis, using up steps $(1,1)$, down steps $(1,-1)$, and level steps $(1,0)$, where the level steps can be either of two kinds: straight and wavy. The *area* of a 2-Motzkin path is defined to be the sum of the heights of the end points of all steps. By a simple substitution, there is a bijection between the set $\mathcal{M}_n$ of 2-Motzkin paths of length n and the set $\mathcal{R}_{n+1}$ of shrunk polyominoes of perimeter $2n$. Given a $\tau \in \mathcal{M}_n$, for each i $(1 \le i \le n)$, we associate the ith step t_i of τ with a pair (p_i, q_i) of steps, where

$$(p_i, q_i) = \begin{cases} (\mathsf{N}, \mathsf{E}) & \text{if } t_i \text{ is an up step} \\ (\mathsf{E}, \mathsf{N}) & \text{if } t_i \text{ is a down step} \\ (\mathsf{N}, \mathsf{N}) & \text{if } t_i \text{ is a straight level step} \\ (\mathsf{E}, \mathsf{E}) & \text{if } t_i \text{ is a wavy level step.} \end{cases}$$

The corresponding shrunk polyomino of τ is the pair (P,Q) of paths, where $P = p_1 \cdots p_n$ and $Q = q_1 \cdots q_n$. It is straightforward to verify that the height of the end point of t_i in τ is equal to the distance between p_i and q_i in (P,Q). By Theorem 5.3(ii), we have the following result.

COROLLARY 5.4. *The area of all 2-Motzkin paths of length* n *is* $4^n - \binom{2n+2}{n}$.

Acknowledgments

The authors would like to thank Professor Louis Shapiro for useful suggestions during his visit to Academia Sinica, Taiwan, and thank the referees for helpful remarks. The authors are partially supported by National Science Council, Taiwan under grants NSC 93-2115-M-390-006 (S.-E. Cheng), NSC 94-2115-M-390-005 (S.-P. Eu), and NSC 94-2115-M-251-001 (T.-S. Fu)

References

[1] E. Barcucci, A. Del Lungo, E. Pergola, R. Pinzani, Some permutations with forbidden subsequences and their inversion numbers, *Discrete Math.* **234** (2001), 1–15.
[2] S. C. Billey, W. Jockusch, R. P. Stanley, Some combinatorial properties of Schubert polynomials, *J. Algebraic Combin.* **2** (1993), 345–374.
[3] E. Chapman, Moments of Dyck paths, *Discrete Math.* **204** (1999), 113–117.
[4] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht, Boston, 1974.
[5] M. P. Delest, X. G. Viennot, Algebraic languages and polyomino enumeration, *Theoret. Comput. Sci.* **34** (1984), 169–206.
[6] E. Deutsch, Dyck paths enumeration, *Discrete Math.* **204** (1999), 167–202.
[7] E. Deutsch, L. W. Shapiro, A survey of the Fine numbers, *Discrete Math.* **241** (2001), 241–265.
[8] M. Merlini, R. Sprugnoli, M. C. Verri, The area determined by under-diagonal lattice path, Proc. CAAP, vol. 1095, Lecture Notes in Computer Science, 1996, pp. 59–71.
[9] E. Pergola, Two bijections for the area of Dyck paths, *Discrete Math.* **241** (2001), 435–447.
[10] A. Reifegerste, Excedances and descents of bi-increasing permutations, preprint, at arXiv:math.CO/0212247.
[11] L. W. Shapiro, D. Rogers, W.-J. Woan, The Catalan numbers, the Lebesgue integral, and 4^{n-2}, *Amer. Math. Monthly* **104** (1997), 926–931.
[12] R. Simion, Combinatorial statistics on non-crossing partitions, *J. Combin. Theory Ser. A* **66** (1994), 270–301.
[13] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences, published electronically at http://www.research.att.com/∼njas/sequences/.
[14] R. P. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, Cambridge, 1999.
[15] W.-J. Woan, Area of Catalan paths, *Discrete Math.* **226** (2001), 439–444.

DEPARTMENT OF APPLIED MATHEMATICS, NATIONAL UNIVERSITY OF KAOHSIUNG, KAOHSIUNG 811, TAIWAN
E-mail address: `chengszu@nuk.edu.tw`

DEPARTMENT OF APPLIED MATHEMATICS, NATIONAL UNIVERSITY OF KAOHSIUNG, KAOHSIUNG 811, TAIWAN
E-mail address: `speu@nuk.edu.tw`

MATHEMATICS FACULTY, NATIONAL PINGTUNG INSTITUTE OF COMMERCE, PINGTUNG 900, TAIWAN
E-mail address: `tsfu@npic.edu.tw`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Conjugacy in Permutation Representations of the Symmetric Group

Extended Abstract

Yona Cherniavsky and Mishael Sklarz

Abstract. Although the conjugacy classes of the general linear group are known, it is not obvious (from the canonic form of matrices) that two permutation matrices are similar if and only if they are conjugate as permutations in the symmetric group, i.e. that conjugacy classes of S_n do not unite under the natural representation. We prove this fact, and give its application to the enumeration of fixed points under a natural action of $S_n \times S_n$. We also consider the permutation representations of S_n which arise from the action of S_n on k-tuples, and classify which of them unite conjugacy classes and which do not.

Résumé. Bien que les classes de conjugaison du groupe linéaire général soient bien connues, il n'est pas évident (à partir de la forme canonique des matrices) que deux oermutations sont similaires si et seulement si elles sont conjuguées comme permutations du groupe symétrique, i.e. que les classes de conjugaison de S_n ne s'unissent pas sous la représentation naturelle.Nous prouvons ici ce fait et nous l'appliquons à l'énumération des points fixes pour une action naturelle de $S_n \times S_n$. We étudions aussi la représentation par permutations de S_n qui découle de l'action de S_n sur les k-uplets, et nous distinguons celles qui unissent les classes de conjugaisons.

1. Introduction

In this extended abstract we study the action of S_n on ordered k-tuples. Denote by ρ_k the corresponding permutation representation over an arbitrary field $\mathbb{F}$. The following problem was presented to us by Lubotzky and Roichman.

Problem 1. *For which $1 \leq k \leq n$ and for which fields $\mathbb{F}$ does the following hold:*
For any two permutations $\pi, \sigma \in S_n$, $\rho_k(\pi)$ is conjugate to $\rho_k(\sigma)$ in $GL(n, \mathbb{F})$ if and only if π and σ are conjugate in S_n.

This problem arises in the enumeration of invertible matrices with respect to a certain natural action of $S_n \times S_n$, see [**BC**] and Section 4 below.

For $k = n$, i.e. the regular representation, a negative solution to Problem 1 was essentially given by Burnside (See [**B**] p. 23-24). In Section 2 it is shown that for $k = 1$ the answer is positive. A full solution is given in Section 3: We find that ρ_1 and ρ_2 do not unite any classes, that ρ_3 unites classes only when n is even, and that ρ_k for $k \geq 4$ always unites some classes. These results do not depend on the choice of the field $\mathbb{F}$. Finally, our results are applied in Section 4 to the enumeration of fixed points of a natural action of $S_n \times S_n$ on invertible matrices. This is an extended abstract: Proofs and full details can be found in [**CS**].

2. The Natural Representation of S_n

There is a natural embedding of S_n in $GL(n, \mathbb{F})$ where $\mathbb{F}$ is any field. Consider a permutation $\pi \in S_n$ as an $n \times n$ matrix obtained from the identity matrix by permutations of the rows. More explicitly: for every permutation $\pi \in S_n$ we identify π with the matrix:

2000 *Mathematics Subject Classification.* Primary 20C30; Secondary 05E15.
Key words and phrases. conjugacy classes, symmetric group, permutation representations, characters, fixed points.

$$[\pi]_{i,j} = \begin{cases} 1 & i = \pi(j) \\ 0 & \text{otherwise} \end{cases}$$

This representation can also be realized as the permutation representation which is obtained from the natural action of S_n on $\{1, 2, \ldots, n\}$ defined by $\pi \cdot i = \pi(i)$.

Our first result is that this representation does not unite conjugacy classes of S_n. We shall use the following well known fact:

Fact 2.1. *If σ is a cycle of length n, then σ^k consists of (n,k) cycles, each of length $n/(n,k)$.*[1]

Proposition 2.1. *Let $\mathbb{F}$ be a field of characteristic 0. The conjugacy classes of S_n do not unite in $GL(n,\mathbb{F})$. In other words, if π and σ are permutations with similar matrices in $GL(n,\mathbb{F})$, then they are conjugate in S_n too.*

Proof. Let π and σ be permutations which are similar as matrices. First of all, we note that for any k, π^k and σ^k are also similar.

Each cycle of length k in π contributes the term $x^k - 1$ into the characteristic polynomial of the permutation matrix. Under the above restriction on char($\mathbb{F}$) it seems reasonable that the cycle structure of a permutation can be recovered from the characteristic polynomial of the corresponding permutation matrix. However, our proof utilizes the trace of the permutation matrix and the traces of its powers.

Denote by $c_d(\pi)$ the number cycles with length equal to d in π. We shall use induction on d to prove that $c_d(\pi) = c_d(\sigma)$, for all d, and this will show that π and σ are conjugate.

Since π and σ are similar as matrices, we have trace(π) = trace(σ). However, the trace function counts the 1's on the diagonal (here we use the restriction on char($\mathbb{F}$)), and each such 1 corresponds to a fixed point of the permutation, so trace(π) = $c_1(\pi)$. Therefore, $c_1(\pi) = c_1(\sigma)$, i.e. π and σ have the same number of fixed points. This is the base of our induction.

Now let d be an arbitrary number, and suppose that $c_k(\pi) = c_k(\sigma)$ for all $k < d$. From Lemma 2.1 it follows that a k-cycle in π ends up as a product of k 1-cycles in π^d if and only if k divides d. Therefore, we can conclude that

$$\mathrm{trace}(\pi^d) = \sum_{k|d} k \cdot c_k(\pi) = d \cdot c_d(\pi) + \sum_{k|d, k<d} k \cdot c_k(\pi).$$

Now, by our induction hypothesis, for all proper divisors $k \mid d$ we have $c_k(\pi) = c_k(\sigma)$. On the other hand, trace(π^d) = trace(σ^d). This implies that $c_d(\pi) = c_d(\sigma)$, and completes the induction argument.

We have shown that π and σ have the same cycle structure, so they are conjugate as permutations. □

Note that if $\mathbb{F}$ is such that char($\mathbb{F}$) $< n$ then the trace of a permutation matrix no longer gives the number of fixed points of the permutation, so a more devious route is necessary.

In this case it is impossible to recover the cycle structure of a permutation from the characteristic polynomial of the corresponding permutation matrix: for example, if char($\mathbb{F}$) = 2 we have $x^4+1 = (x^2+1)^2 = (x+1)^4$, i.e. one cycle of length 4, two cycles of length 2 and four cycles of length 1 all have the same characteristic polynomial.

However, in [**CS**] we extend Proposition 2.1, and prove the following:

Theorem 2.2. *Let $\mathbb{F}$ be an arbitrary field. The conjugacy classes of S_n do not unite in $GL(n,\mathbb{F})$. In other words, if π and σ are permutations with similar matrices in $GL(n,\mathbb{F})$, then they are conjugate in S_n too.*

The proof is obtained by considering certain eigenspaces of powers of π and σ. See [**CS**] for full details.

It should be noted that this property of the natural representation seems to be very "delicate". For example, in the natural representations of the signed permutation groups this property fails to hold. In particular, in B_2, the permutations $\sigma = (1,2)$ and $\tau = (1,\bar{1})$ are not conjugate, and yet the matrices associated with them, namely

$$P(\sigma) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad P(\tau) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

are similar matrices.

[1] We use (n,k) to denote the greatest common divisor of n and k.

3. Other Permutation Representations

3.1. Representations Arising from the Action of S_n on k-tuples. In Section 2 we proved that the natural representation of S_n does not unite conjugacy classes. On the other hand, it is well known (see [**B**] p. 23-24) that the regular representation of S_n (indeed, of any group) unites all elements of equal order. The natural representation can be seen as the permutation representation obtained from the natural action of S_n on the set $\{(1), (2), \ldots, (n)\}$ of 1-tuples. On the other hand, the regular representation can be seen as the permutation representation which arises from the action of S_n on all $n!$ ordered n-tuples of numbers from $\{1, 2, \ldots, n\}$. In this section we wish to address the representations in between: the representation arising from the action of S_n on pairs, triplets, etc. and to see where the representations start uniting conjugacy classes.

At first we shall confine ourselves to the complex field, and prove our results there. The results for general fields will follow from these results.

We begin with a general theorem, which holds true for any representation of any finite group.

Theorem 3.1. *Let G be a group, and $\sigma, \tau \in G$. Let $T : G \to GL(d, \mathbb{C})$ a representation of G, with character χ. Then $T(\sigma) \sim T(\tau)$ as matrices if and only if $\chi(\sigma^k) = \chi(\tau^k)$ for all k.*

Proof is given in [**CS**].

Note that the fact that the regular representation unites all elements of equal order can be derived from this theorem: If χ is the character of the regular representation, then

$$\chi(\sigma^k) = \begin{cases} |G| & |\sigma| \mid k, \\ 0 & \text{otherwise} \end{cases}$$

so obviously $\chi(\sigma^k) = \chi(\tau^k)$ for all k if and only if σ and τ have the same order.

The criterion which we just presented is still rather complicated to use for general groups, but it can be simplified in our case, because of the following simple fact.

Fact 3.2. *Let $\sigma \in S_n$, with $|\sigma| = m$.*

- *If k is relatively prime to m. Then $\sigma^k \sim \sigma$.*
- *For any k, $\sigma^k \sim \sigma^{(m,k)}$.*

Claim 3.3. *Let $T : S_n \to GL(d, \mathbb{C})$ be a representation of the symmetric group, with character χ, and $\sigma, \tau \in S_n$ elements of order m. Then $T(\sigma) \sim T(\tau)$ if and only if $\chi(\sigma^k) = \chi(\tau^k)$ for $k \mid m$.*

Corrolary 3.4. *If σ and τ are of prime order p, then $T(\sigma) \sim T(\tau)$ if and only if $\chi(\sigma) = \chi(\tau)$.*

Definition 3.5. Let $\sigma, \tau \in S_n$ be elements of equal order m, such that if $k \neq 1$ and $k \mid m$ then $\sigma^k \sim \tau^k$. It follows from 3.3 that $T(\sigma) \sim T(\tau)$ if and only if $\chi(\sigma) = \chi(\tau)$. We call such elements *almost similar*. In fact, it is sufficient to require that $\sigma^p \sim \tau^p$ for all prime divisors of m.

We next show that almost similar elements are typical examples of elements that are united by representations, in the following sense:

Theorem 3.6. *Let $T : S_n \to GL(d, \mathbb{C})$ be a representation. If T unites some two conjugacy classes, then there must exist a pair of almost similar elements which it unites.*

Having proved this, we now have a criterion to check whether a representation unites classes: It is sufficient to show that all pairs of almost similar elements remain non united, i.e. that the character of the representation takes different values on them.

Using this criterion, we show

Theorem 3.7.

1. *The natural representation does not unite classes.*
2. *The representation arising from the action on pairs does not unite classes.*
3. *The representation arising from the action on triplets unites classes iff n is even.*
4. *Representations arising from the action on k-tuples, with $k \geq 4$, always unite some classes.*

3.2. Representations Arising from the Action of S_n on Subsets. The natural route to follow now would be to try and generalize these results to other permutation representations, and in particular to those arising from the action of S_n on k-subsets of $\{1, 2, \ldots, n\}$. The general answer eludes us at present, and seems to be pretty unsatisfactory. However, we have managed to show that the representation arising from the action of S_n on *all* subsets of $[n]$ does in fact not unite any classes. In all this section, we shall omit proofs, and refer the interested reader to [**CS**] for full details.

Theorem 3.8. *The action of S_n on the power set $2^{[n]}$ of $[n]$ does not unite classes.*

Consider now the action of S_n on *even* sized subsets of $[n]$. If n is even, then this action unites some classes. For example, $(1,2)(3,4)\ldots(n-1,n)$ and $(1)(2)(3,4)\ldots(n-1,n)$ get united. (They are almost similar, and both fix $2^{n/2}$ sets.)

However, if n is odd, then this representation does not unite classes.

Theorem 3.9. *Let n be odd. The action of S_n on the set of even-sized subsets of $[n]$ does not unite classes.*

Finally, we conclude this section by exploring the behavior of the representation arising from the action of S_n on *odd* sized subsets of $[n]$.

Theorem 3.10. *The action of S_n on the set of odd-sized subsets of $[n]$ does not unite classes. This does not depend on n's parity.*

3.3. General Fields. The proofs in the two previous sections apply only to the complex field $\mathbb{C}$, (in fact, to all fields with characteristic 0.) We shall now show that the same applies to any field. We shall base ourselves on Theorem 2.1 from Section 2, where we proved that the natural representation does not unite classes, regardless the base field.

Lemma 3.11. *Let $f : G \to H$ and $g : H \to K$ be group homomorphisms.*

1. *If f and g both do not unite classes, then also gf does not unite them.*
2. *If gf does not unite classes, then neither does f.*

Theorem 3.12. *Let T be any permutation representation of S_n. If T does not unite classes when considered a representation into $GL(m, \mathbb{C})$, then it does not unite classes when considered as a representation into $GL(m, \mathbb{F})$, for any field $\mathbb{F}$.*

Proof. Any permutation representation can be factored into $S_n \to S_m \to GL(m, \mathbb{C})$, where the first homomorphism is the permutation representation and the second is the natural representation. Now, suppose T does not unite classes. By Lemma 3.11, neither does the permutation representation $S_n \to S_m$. We already know that the natural representation does not unite classes, whatever the field. Tacking these two homomorphisms together gives us the representation in any field, and another appeal to Lemma 3.11 proves that it still doesn't unite any classes. $\square$

4. The action of $S_n \times S_n$ on invertible matrices

In this section we present an application of Theorem 2.1.

Definition 4.1. Let $\mathbb{F}$ be any field. We define an action of $S_n \times S_n$ on the group $GL(n, \mathbb{F})$ by

$$(\pi, \sigma) \bullet A = \pi A \sigma^{-1} \text{ where } (\pi, \sigma) \in S_n \times S_n \text{ and } A \in GL(n, \mathbb{F}) \tag{1}$$

Definition 4.2. Let M be a finite subset of $GL(n, \mathbb{F})$, invariant under the action of $S_n \times S_n$ defined above. We denote by α_M the permutation representation of $S_n \times S_n$ obtained from the action (1) . In the sequel we identify the action (1) with the permutation representation α_M associated with it.

Now we define a generalization of the conjugacy representation of S_n

We present a conjugacy representation of S_n on a subset M of $GL(n, \mathbb{F})$.

Definition 4.3. Denote by β the permutation representation of S_n obtained by the following action on M.

$$\pi \circ A = (\pi, \pi) \bullet A = \pi A \pi^{-1} \tag{2}$$

The connection between α_M and β_M is given by the following easily seen claim:

CLAIM 4.4. *Consider the diagonal embedding of S_n into $S_n \times S_n$. Then*

$$\beta_M = \alpha_M \downarrow_{S_n}^{S_n \times S_n} . \qquad \square$$

COROLLARY 4.5. *For every finite set $M \subseteq GL(n, \mathbb{F})$ invariant under the action (1) of $S_n \times S_n$ defined above:*
If π and σ are conjugate in S_n then

$$\chi_{\alpha_M}((\pi, \sigma)) = \chi_{\alpha_M}((\pi, \pi)) = \chi_{\beta_M}(\pi) = \#\{A \in M \mid \pi A = A\pi\} .$$

If π is not conjugate to σ in S_n then

$$\chi_{\alpha_M}((\pi, \sigma)) = 0 .$$

PROOF. If π and σ are conjugate in S_n then (π, σ) is conjugate to (π, π) in $S_n \times S_n$. Since the character is a class function, we have:

$$\chi_{\alpha_M}(\pi, \sigma) = \chi_{\alpha_M}(\pi, \pi) = \#\{A \in M \mid \pi A \pi^{-1} = A\} = \#\{A \in M \mid \pi A = A\pi\}$$

i.e. the value of the character of α_M calculated on the element (π, σ) with π conjugate to σ in S_n is equal to the number of matrices in M which commute with the permutation matrix π.

Now, we know that the character of a permutation representation counts the number of fixed points, so:

$$\chi_{\alpha_M}(\pi, \sigma) = \#\{A \in M \mid \pi A \sigma^{-1} = A\} = \#\{A \in M \mid \pi = A\sigma A^{-1}\}.$$

Note that $\pi = A\sigma A^{-1}$ means that π and σ are similar as invertible matrices. Thus, by Theorem 2.1, if π and σ are not conjugate in S_n they can not be conjugate in $GL(n, \mathbb{F})$ and we have:

$$\{A \in M \mid \pi = A\sigma A^{-1}\} = \varnothing$$

and so

$$\chi_{\alpha_M}(\pi, \sigma) = 0$$

if π and σ are not conjugate in S_n. $\square$

For an application of Corrolary 4.5 to the enumeration of fixed points, see [**BC**].

Acknowledgments. The authors are grateful to Alex Lubotzky, Yuval Roichman, Eli Bagno, Uzi Vishne and Boris Kunyavskii for helpful discussions.

References

[B] William Burnside, *Theory of Groups of Finite Order.* University Press, Cambridge 1897.

[BC] Eli Bagno and Yona Cherniavsky, *Permutation representations on invertible matrices.* Arxiv reference math.RT/0411556 To be published.

[CR] Charles Curtis and Irving Reiner, *Representation Theory of Finite Groups and Associative Algebras.* John Wiley and Sons, New York 1962.

[CS] Yona Cherniavsky and Mishael Sklarz, *Conjugacy in permutation representations of the symmetric group.* Arxiv reference math.CO/0411554

[FH] William Fulton and Joe Harris, *Representation Theory, a First Course.* Springer, New York 1991

[S] Bruce Sagan, *The Symmetric Group. Representations, Combinatorial Algorithms, and Symmetric Functions.* Springer, New York 2001

[Se] Jean-Pierre Serre, *Linear Representations of Finite Groups.* Springer, New York 1977

DEPARTMENT OF MATHEMATICS, BAR ILAN UNIVERSITY, RAMAT GAN 52900, ISRAEL
E-mail address: `cherniy@math.biu.ac.il`

DEPARTMENT OF MATHEMATICS, BAR ILAN UNIVERSITY, RAMAT GAN 52900, ISRAEL
E-mail address: `sklarzm@math.biu.ac.il`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Enumeration of Bruhat intervals between nested involutions in S_n

Alessandro Conflitti

Abstract. We build a chain

$$\mathrm{id} = \vartheta_0 < \vartheta_1 < \cdots < \vartheta_{\lfloor \frac{n}{2} \rfloor - 1} < \vartheta_{\lfloor \frac{n}{2} \rfloor}$$

of nested involutions in the Bruhat ordering of S_n, with $\vartheta_{\lfloor \frac{n}{2} \rfloor}$ the maximal element for the Bruhat order, and we study the cardinality of the Bruhat intervals $[\vartheta_j, \vartheta_k]$ for all $0 \leq j < k \leq \lfloor \frac{n}{2} \rfloor$, and the number of permutations incomparable with ϑ_t, for all $0 \leq t \leq \lfloor \frac{n}{2} \rfloor$.

Résumé. Nous construisons une chaîne

$$\mathrm{id} = \vartheta_0 < \vartheta_1 < \cdots < \vartheta_{\lfloor \frac{n}{2} \rfloor - 1} < \vartheta_{\lfloor \frac{n}{2} \rfloor}$$

des involutions nichées dans l'ordre de Bruhat de S_n, avec $\vartheta_{\lfloor \frac{n}{2} \rfloor}$ l'élément maximal pour l'ordre de Bruhat, et nous étudions la cardinalité des intervalles de Bruhat $[\vartheta_j, \vartheta_k]$ pour tout les $0 \leq j < k \leq \lfloor \frac{n}{2} \rfloor$, et le nombre de permutations incomparables avec ϑ_t, pour tout le $0 \leq t \leq \lfloor \frac{n}{2} \rfloor$.

1. Overview

For any $n \geq 2$, let S_n be the symmetric group of n elements equipped with the Bruhat ordering $\leq$; see e.g. [**3, 6, 8, 21, 22**]. One of the most celebrated combinatorial and algebraic problems is to study its Bruhat graph and its Bruhat intervals $[a, b] = \{z \in S_n : a \leq z \leq b\}$ for $a, b \in S_n$; see e.g. [**1, 7, 12, 15**]. These are intimately related with the Kazhdan–Lusztig polynomials of S_n and the algebraic geometry of Schubert varieties. See e.g. [**9, 10, 13, 14, 18**] and the references therein.

In this work we build a chain

$$\mathrm{id} = \vartheta_0 < \vartheta_1 < \cdots < \vartheta_{\lfloor \frac{n}{2} \rfloor - 1} < \vartheta_{\lfloor \frac{n}{2} \rfloor}$$

of nested involutions in the Bruhat ordering of S_n, with $\vartheta_{\lfloor \frac{n}{2} \rfloor}$ the maximal element for the Bruhat order (see Definition 3.1 for the exact definition of ϑ_t, $t = 0, \ldots, \lfloor \frac{n}{2} \rfloor$), and we study the cardinality of the Bruhat intervals $[\vartheta_j, \vartheta_k]$ for all $0 \leq j < k \leq \lfloor \frac{n}{2} \rfloor$. Moreover we study the number of permutations incomparable with ϑ_t, for all $0 \leq t \leq \lfloor \frac{n}{2} \rfloor$. Our results imply and generalize the result of [**27**], where a closed formula for the cardinality of $[\vartheta_0, \vartheta_1]$ is proved. This problem is related to the explicit computation of Kazhdan–Lusztig polynomials for some classes of elements. See e.g. [**24, 25**] and the references therein.

The importance of the set $\{\vartheta_t : t = 0, \ldots, \lfloor \frac{n}{2} \rfloor\}$ lies in the fact that involutions of the symmetric group and, more generally, of Coxeter groups, are elements having nice algebraic properties, see [**28, 29, 30, 31**]. In particular, in [**28**] it is proved that the maximal length element of any conjugacy class in S_n containing involutions is one of the ϑ_t for some $t = 0, \ldots, \lfloor \frac{n}{2} \rfloor$.

2. Preliminaries

In this section we collect together some definitions, notation and results that will be used in the following. We follow [**11, 20, 32**] for combinatorics and poset notation and terminology.

2000 *Mathematics Subject Classification.* 05A15, 05E15, 20F55, 33C05.
Key words and phrases. symmetric group, Bruhat intervals, nested involutions.

For $x \in \mathbb{R}$ we let $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : n \leq x\}$; for $n \in \mathbb{N}$ we let $[n] = \{t \in \mathbb{N} : 1 \leq t \leq n\} = \{1, \ldots, n\}$, and $[0] = \emptyset$. For any complex number a, we define the rising factorial as $(a)_0 = 1$ and $(a)_m = \prod_{j=0}^{m-1}(a+j)$ for any $m \in \mathbb{N} \setminus \{0\}$. The cardinality of a set $\mathcal{X}$ will be denoted by $\#\mathcal{X}$.

For any $n \geq 2$, let S_n be the symmetric group of permutations of n objects, viz. the set of all bijections

$$\sigma : [n] \xrightarrow{\sim} [n].$$

If $\sigma \in S_n$ then we write $\sigma = [a_1, \ldots, a_n]$ to mean that $\sigma(j) = a_j$ for $j \in [n]$. Sometimes we also write σ in *disjoint cycle form* and we usually omit writing the 1–cycles of σ. Given $\sigma, \tau \in S_n$ we let $\sigma\tau = \sigma \circ \tau$ (composition of functions) so that, for example $(1,2)(2,3) = (1,2,3)$. For any $\sigma = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ we say that a pair $(i,j) \in [n] \times [n]$ is an *inversion* of σ if $i < j$ and $a_i > a_j$, and we denote the number of inversions of σ by $\operatorname{inv}(\sigma)$.

We set

- : $\mathcal{E}_n = \{(j, j+1) : j \in [n-1]\}$,
- : $T_n = \{(i,j) : 1 \leq i < j \leq n\}$, the set of transpositions in S_n,
- : $D(\sigma) = \{\tau \in \mathcal{E}_n : \operatorname{inv}(\sigma\tau) < \operatorname{inv}(\sigma)\}$, the *descent set* of $\sigma \in S_n$.

We recall the definition of Bruhat order on S_n:

Definition 2.1. Let $n \geq 2$. For any $u, v \in S_n$, $u < v$ in Bruhat order if and only if there exist $k \in \mathbb{N}$ and $t_1, \ldots, t_k \in T_n$ such that

$$\begin{aligned} &v = ut_1 \cdots t_k, \\ &\operatorname{inv}(ut_1 \cdots t_{j+1}) > \operatorname{inv}(ut_1 \cdots t_j) \qquad \text{for any } \; j \in [k-1]. \end{aligned}$$

It is easy to see that $[n, \ldots, 1]$ is the maximum element in S_n for the Bruhat order.

Now we state a criterion for deciding when two permutations are comparable in the Bruhat ordering, which was achieved in [**5**].

Theorem 2.2. *Let $n \geq 2$, and for any $\sigma, \tau \in S_n$, let $\sigma[j,k]$ be the j-th entry in the increasing rearrangement of $\{\sigma(1), \ldots, \sigma(k)\}$ for all $1 \leq j \leq k \leq n-1$, and define $\tau[j,k]$ similarly. Then the following are equivalent:*

1. *$\sigma \leq \tau$ in the Bruhat order,*
2. *$\sigma[j,k] \leq \tau[j,k]$, for all $k \in D(\sigma)$ and $1 \leq j \leq k$,*
3. *$\sigma[j,k] \leq \tau[j,k]$, for all $k \in \{1, \ldots, n-1\} \setminus D(\tau)$ and $1 \leq j \leq k$.*

3. Main Results

Definition 3.1. Let $n \geq 2$. We define

$$\begin{aligned} \vartheta_t &= \prod_{j=0}^{t-1}(j+1, n-j) = (1,n)\cdots(t, n-t+1) \\ &= [n, \ldots, n-t+1, t+1, \ldots, n-t, t, \ldots, 1] \in S_n \end{aligned}$$

for all $0 \leq t \leq \left\lfloor \frac{n}{2} \right\rfloor$.

Obviously ϑ_t is an involution for all $0 \leq t \leq \left\lfloor \frac{n}{2} \right\rfloor$, and

$$\mathrm{id} = \vartheta_0 < \vartheta_1 < \cdots < \vartheta_{\left\lfloor \frac{n}{2} \right\rfloor - 1} < \vartheta_{\left\lfloor \frac{n}{2} \right\rfloor} = \max\{\sigma \in S_n\}$$

in the Bruhat order of S_n.

Definition 3.2. Let $n \geq 2$, and $0 \leq t \leq n-1$. We define

$$\mathcal{F}_t(n) = \begin{cases} \{\sigma \in S_n : \sigma \leq \vartheta_t\} = [\vartheta_0, \vartheta_t] & \text{if } t \leq \left\lfloor \frac{n}{2} \right\rfloor, \\ S_n & \text{if } t \geq \left\lfloor \frac{n}{2} \right\rfloor; \end{cases}$$

and

$$\mathbf{F}_t(n) = \#\mathcal{F}_t(n),$$

setting $\mathbf{F}_t(0) = \mathbf{F}_t(1) = 1$.

LEMMA 3.3. *Let $n \geq 2$, $0 \leq t \leq n-1$, and $j \in [t+1]$. The number of permutations $\sigma \in \mathcal{F}_t(n)$ with the constraint that there exists a subset $A = \{\alpha_1, \dots, \alpha_j\} \subset [t+1]$ and an array $B = (\beta_1, \dots, \beta_j)$ with pairwise distinct coordinates $\beta_k \in [t+1]$ for all $k \in [j]$ such that*

$$\sigma(\alpha_l) = \beta_l \qquad \textit{for all } l \in [j]$$

equals

$$j!\binom{t+1}{j}^2 \mathbf{F}_t(n-j).$$

PROOF. Of course we can always assume $t \geq 1$ otherwise the result is trivial.
For any fixed $A = \{\alpha_1, \dots, \alpha_j\}$ and $B = (\beta_1, \dots, \beta_j)$ with the desired properties, let

$$\mathcal{Z}_t^n[j](A,B) = \{\sigma \in \mathcal{F}_t(n) : \sigma(\alpha_k) = \beta_k \text{ for all } k \in [j]\}.$$

We note that from Theorem 2.2 and Definition 3.1 we get that $\mathcal{Z}_t^n[j](A,B) \neq \emptyset$ for all possible choices of A and B.

Consider the order–preserving bijections

$$\begin{aligned} \varphi &: [n] \setminus A \xrightarrow{\sim} [n-t], \\ \psi &: [n] \setminus B \xrightarrow{\sim} [n-t]. \end{aligned}$$

Then from Theorem 2.2 there is a bijection

$$f : \mathcal{Z}_t^n[j](A,B) \xrightarrow{\sim} \mathcal{F}_t(n-j)$$

defined in the following way: we delete $\sigma(k)$ if $k \in A$, whereas for all $k \notin A$

$$\sigma(k) \xrightarrow{f} \psi(\sigma(\varphi(k))).$$

Noticing that there are $\binom{t+1}{j}$ ways for choosing A and $j!\binom{t+1}{j}$ ways for choosing B, the desired result follows. $\square$

THEOREM 3.4. *For any $n \geq 2$,*

$$\mathbf{F}_t(n) = \begin{cases} \sum_{j=1}^{t+1} (-1)^{j-1} j!\binom{t+1}{j}^2 \mathbf{F}_t(n-j) & \text{if } 0 \leq t \leq n-1, \\ n! & \text{if } t \geq n. \end{cases}$$

PROOF. Of course we can always assume $t \in [n-1]$ otherwise the result is trivial.
From Theorem 2.2 we see that if $\sigma \in \mathcal{F}_t(n)$ then

$$\{\sigma(k) : k \in [t+1]\} \bigcap [t+1] \neq \emptyset.$$

Let $k \in [t+1]$ and

$$R_k = \{\sigma \in \mathcal{F}_t(n) : \sigma(k) \in [t+1]\}.$$

Then by inclusion–exclusion we have

$$\mathbf{F}_t(n) = \#\left(\bigcup_{k \in [t+1]} R_k\right) = \sum_{j=1}^{t+1} (-1)^{j-1} \sum_{\substack{\mathcal{I} \subset [t+1] \\ \#\mathcal{I} = j}} \#\left(\bigcap_{z \in \mathcal{I}} R_z\right),$$

and the desired result follows from Lemma 3.3. $\square$

The following Corollary is immediate, and it gives a purely combinatorial proof of an identity for the factorial.

COROLLARY 3.5. *For any $n \geq 2$ and for all $\left\lfloor \frac{n}{2} \right\rfloor \leq k, t \leq n-1$,*

$$\sum_{j=1}^{k+1} (-1)^{j-1} j!\binom{k+1}{j}^2 (n-j)! = \sum_{j=1}^{t+1} (-1)^{j-1} j!\binom{t+1}{j}^2 (n-j)! = n!$$

PROOF. From Definition 3.2, $\mathbf{F}_k(n) = \mathbf{F}_t(n) = n!$ for all $\left\lfloor \frac{n}{2} \right\rfloor \leq k, t \leq n-1$. Taking in account Theorem 3.4, the desired result follows. $\square$

We note that this identity can be also proved using the theory of hypergeometric series and applying Chu–Vandermonde summation, see [**2, 17, 19, 23**]. In fact, it is equivalent to

$$\sum_{j=0}^{\infty} (-1)^j \, j! \binom{t+1}{j}^2 (n-j)! = 0$$

for all $\left\lfloor \frac{n}{2} \right\rfloor \leq t \leq n-1$, and we have that

$$\begin{aligned} \sum_{j=0}^{\infty} (-1)^j \, j! \binom{t+1}{j}^2 (n-j)! &= \left({}_2F_1 \left[\begin{matrix} -t-1, -t-1 \\ -n \end{matrix} ; 1 \right] \right) ((1)_n) \\ &= \frac{((1)_n) \left((1-n+t)_{1+t} \right)}{(-n)_{1+t}}; \end{aligned}$$

obviously if $\left\lfloor \frac{n}{2} \right\rfloor \leq t \leq n-1$ then $1-n+t \leq 0 \leq 1-n+2t$, therefore $(1-n+t)_{1+t} = 0$.

Now we give an explicit formula for the generating function of the sequence $\{\mathbf{F}_t(n)\}_{n \geq 2t}$ for any $t \geq 1$, and then, using it, we are able to prove a closed formula for the function $\mathbf{F}_t(n)$ for any $t \geq 1$ and any $n \geq 2t$.

THEOREM 3.6. *For any $t \geq 1$,*

$$\sum_{n \geq 2t} \mathbf{F}_t(n) X^n = X^{2t} \frac{\sum_{k=0}^{t} \left(\sum_{j=k+1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! \, (2t-j+k)! \right) X^k}{\sum_{j=0}^{t+1} (-1)^j \binom{t+1}{j}^2 j! X^j}.$$

PROOF. From Theorem 3.4 we get

$$\begin{aligned} \sum_{n \geq 2t} \mathbf{F}_t(n) X^n &= \sum_{n \geq 2t} \sum_{j=1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! \mathbf{F}_t(n-j) X^n \\ &= \sum_{j=1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! X^j \\ &\quad \cdot \left[\sum_{m=2t-j}^{2t-1} \mathbf{F}_t(m) X^m + \sum_{m \geq 2t} \mathbf{F}_t(m) X^m \right]. \end{aligned}$$

From Definition 3.2 we have $\mathbf{F}_t(n) = n!$ if $2t+1 \geq n$, thus

$$\sum_{m=2t-j}^{2t-1} \mathbf{F}_t(m) X^m = \sum_{m=2t-j}^{2t-1} m! X^m = \sum_{k=0}^{j-1} (2t-j+k)! X^{2t-j+k},$$

hence

$$\begin{aligned} &\left(\sum_{n \geq 2t} \mathbf{F}_t(n) X^n \right) \left(\sum_{j=0}^{t+1} (-1)^j \binom{t+1}{j}^2 j! X^j \right) \\ &= X^{2t} \left[\sum_{j=1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! \sum_{k=0}^{j-1} (2t-j+k)! X^k \right] \\ &= X^{2t} \cdot \sum_{k=0}^{t} \left(\sum_{j=k+1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! \, (2t-j+k)! \right) X^k, \end{aligned}$$

and the desired result follows. $\square$

THEOREM 3.7. *For any $t \geq 1$ and any $n \geq 2t$*

$$\mathbf{F}_t(n) = \sum_{z=0}^{\min\{t,n-2t\}} \left[\left(\sum_{j=z+1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j!\,(2t-j+z)! \right) \cdot \left(\sum_{\substack{\alpha\in\mathbb{N}^{t+1}\\ \Omega(\alpha)=n-2t-z}} \frac{(\|\alpha\|)!}{\prod_{k=1}^{t+1}(\alpha_k!)} \left(\prod_{j=1}^{t+1} \left((-1)^{j-1} \binom{t+1}{j}^2 j! \right)^{\alpha_j} \right) \right) \right],$$

where for any multi–index $\alpha = (\alpha_1, \dots \alpha_{t+1}) \in \mathbb{N}^{t+1}$ we set $\|\alpha\| = \sum_{j=1}^{t+1} \alpha_j$ and $\Omega(\alpha) = \sum_{j=1}^{t+1} j \cdot \alpha_j$.

PROOF. With an eye on Theorem 3.6, observe first that

$$\begin{aligned} & \frac{1}{\sum_{j=0}^{t+1} (-1)^j \binom{t+1}{j}^2 j! X^j} \\ & = \frac{1}{1-\sum_{j=1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! X^j} \\ & = \sum_{l\geq 0} \left(\sum_{j=1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! X^j \right)^l . \end{aligned} \tag{3.1}$$

Now, for any $r \geq 1$ and any multi–index $\alpha = (\alpha_1, \dots \alpha_r) \in \mathbb{N}^r$, we set $\|\alpha\| = \sum_{j=1}^{r} \alpha_j$ and $\Omega(\alpha) = \sum_{j=1}^{r} j \cdot \alpha_j$, and we recall that for any $r \geq 1$, $s \geq 1$, and $z_1, \dots, z_r \in \mathbb{R}$ we have

$$\left(\sum_{j=1}^{r} z_j \right)^s = \sum_{\substack{\alpha\in\mathbb{N}^r\\ \|\alpha\|=s}} \frac{s!}{\prod_{k=1}^{r}(\alpha_k!)} \left(\prod_{j=1}^{r} z_j^{\alpha_j} \right).$$

Therefore (3.1) equals

$$\begin{aligned} & \sum_{l\geq 0} \left(\sum_{j=1}^{t+1} (-1)^{j-1} \binom{t+1}{j}^2 j! X^j \right)^l \\ & = \sum_{l\geq 0} \sum_{\substack{\alpha\in\mathbb{N}^{t+1}\\ \|\alpha\|=l}} \frac{l!}{\prod_{k=1}^{t+1}(\alpha_k!)} \left[\prod_{j=1}^{t+1} \left((-1)^{j-1} \binom{t+1}{j}^2 j! X^j \right)^{\alpha_j} \right] \\ & = \sum_{\alpha\in\mathbb{N}^{t+1}} \frac{(\|\alpha\|)!}{\prod_{k=1}^{t+1}(\alpha_k!)} \left[\prod_{j=1}^{t+1} \left((-1)^{j-1} \binom{t+1}{j}^2 j! \right)^{\alpha_j} \right] X^{\Omega(\alpha)} \\ & = \sum_{v\geq 0} \left[\sum_{\substack{\alpha\in\mathbb{N}^{t+1}\\ \Omega(\alpha)=v}} \frac{(\|\alpha\|)!}{\prod_{k=1}^{t+1}(\alpha_k!)} \left(\prod_{j=1}^{t+1} \left((-1)^{j-1} \binom{t+1}{j}^2 j! \right)^{\alpha_j} \right) \right] X^v, \end{aligned} \tag{3.2}$$

and combining Theorem 3.6 and (3.2) we get

$$\begin{aligned}
&\sum_{n\geq 2t}\mathbf{F}_t\left(n\right)X^n\\
&= X^{2t}\cdot\left[\sum_{k=0}^{t}\left(\sum_{j=k+1}^{t+1}\left(-1\right)^{j-1}\binom{t+1}{j}^2 j!\left(2t-j+k\right)!\right)X^k\right]\\
&\cdot\sum_{v\geq 0}\left[\sum_{\substack{\alpha\in\mathbb{N}^{t+1}\\ \Omega(\alpha)=v}}\frac{\left(\|\alpha\|\right)!}{\prod_{k=1}^{t+1}\left(\alpha_k!\right)}\left(\prod_{j=1}^{t+1}\left(\left(-1\right)^{j-1}\binom{t+1}{j}^2 j!\right)^{\alpha_j}\right)\right]X^v\\
&= \sum_{l\geq 0}\left[\sum_{z=0}^{\min\{l,t\}}\left(\sum_{j=z+1}^{t+1}\left(-1\right)^{j-1}\binom{t+1}{j}^2 j!\left(2t-j+z\right)!\right)\right.\\
&\left.\cdot\left(\sum_{\substack{\alpha\in\mathbb{N}^{t+1}\\ \Omega(\alpha)=l-z}}\frac{\left(\|\alpha\|\right)!}{\prod_{k=1}^{t+1}\left(\alpha_k!\right)}\left(\prod_{j=1}^{t+1}\left(\left(-1\right)^{j-1}\binom{t+1}{j}^2 j!\right)^{\alpha_j}\right)\right)\right]X^{2t+l}.
\end{aligned}$$

The desired result follows. □

Now we show that knowing the cardinality of Bruhat intervals starting from the identity leads to knowing the cardinality of Bruhat intervals between two general nested involutions.

Theorem 3.8. *Let $n \geq 2$ and $0 \leq j < k \leq \left\lfloor \frac{n}{2} \right\rfloor$; then*

$$\#\left[\vartheta_j,\vartheta_k\right] = \mathbf{F}_{k-j}\left(n-2j\right).$$

Proof. In order to prove the statement we exhibit a bijection

$$\begin{array}{rcl} f:\left[\vartheta_j,\vartheta_k\right] & \xrightarrow{\sim} & \mathcal{F}_{k-j}\left(n-2j\right)\\ \sigma & \longmapsto & f_\sigma \end{array}$$

From Theorem 2.2 we see that if $\sigma \in \left[\vartheta_j,\vartheta_k\right]$ then $\sigma\left(l\right) = n+1-l$ for all $l \in [j]\bigcup\left([n]\setminus[n-j]\right)$. We set

$$f_\sigma\left(l\right) = \sigma\left(l+j\right)-j$$

for all $l \in [n-2j]$, and the desired result follows. □

Knowing the cardinality of Bruhat intervals starting from the identity and Bruhat intervals between two nested involutions leads to knowing the number of permutations less or equal than one of the two nested involutions and incomparable with the other one.

Theorem 3.9. *Let $n \geq 2$ and $0 \leq j < k \leq \left\lfloor \frac{n}{2} \right\rfloor$; then*

$$\begin{aligned}
&\#\{\sigma\in S_n : \sigma\leq\vartheta_k \textit{ and } \sigma \textit{ is incomparable with } \vartheta_j\}\\
&= \mathbf{F}_k\left(n\right)-\mathbf{F}_j\left(n\right)-\mathbf{F}_{k-j}\left(n-2j\right)+1.
\end{aligned}$$

Proof. We have

$$\begin{aligned}
&\{\sigma\in S_n : \sigma\leq\vartheta_k \text{ and } \sigma \text{ is incomparable with } \vartheta_j\}\\
&= \mathcal{F}_k\left(n\right)\setminus\left(\mathcal{F}_j\left(n\right)\bigcup\left[\vartheta_j,\vartheta_k\right]\right),
\end{aligned}$$

and $\mathcal{F}_j\left(n\right)\bigcap\left[\vartheta_j,\vartheta_k\right] = \{\vartheta_j\}$; the desired result follows. □

Corollary 3.10. *Let $n \geq 2$ and $0 \leq t \leq \left\lfloor \frac{n}{2} \right\rfloor$; then*

$$\#\{\sigma\in S_n : \sigma \textit{ is incomparable with } \vartheta_t\} = n!+1-\mathbf{F}_t\left(n\right)-\mathbf{F}_{\left\lfloor\frac{n}{2}\right\rfloor-t}\left(n-2t\right).$$

□

4. Remarks

Studying the cardinality of Bruhat intervals between similar nested involutions in different Coxeter systems leads to other challenging questions.

In particular, one can consider the chains

$$\mathrm{id} = \phi_0 < \phi_1 < \cdots < \phi_{n-1} < \phi_n,$$
$$\mathrm{id} = \psi_0 < \psi_1 < \cdots < \psi_{\lfloor \frac{n}{2} \rfloor - 1} < \psi_{\lfloor \frac{n}{2} \rfloor}$$

of nested involutions in the Bruhat ordering of B_n, the *hyperoctahedral group* of rank n (see [**6, 26**]), where

$$\begin{aligned} \phi_r &= \prod_{j=0}^{r-1} (-n+j, n-j), \\ \psi_t &= \prod_{j=0}^{t-1} (j+1, -n+j)\,(-j-1, n-j) \end{aligned}$$

for any $r = 0, \ldots, n-1$ and any $t = 0, \ldots, \lfloor \frac{n}{2} \rfloor$, and to investigate $\#\,[\phi_j, \phi_k]$ for all $0 \le j < k \le n-1$ and $\#\,[\psi_h, \psi_z]$ for all $0 \le h < z \le \lfloor \frac{n}{2} \rfloor$.

We note that in order to study enumeration of Bruhat intervals in a Coxeter system (W, S) (see [**6, 21, 22**] for comprehensive references about Coxeter systems) it is not required that $W < \infty$. In fact, the following fact is well–known, and we refer e.g. to [**6**] for a proof.

Proposition 4.1. *Let (W, S) be a Coxeter system, and $u, v \in W$. Bruhat intervals $[u, v] = \{z \in W : u \le z \le v\}$ are finite (even if $\#S = \infty$). In fact, $\#\,[u, v] \le 2^{l(v)}$, where $l\,(v)$ denotes the length of v.*

Therefore, another tempting choice to investigate the cardinality of Bruhat intervals between suitable involutions would be to consider $\tilde{A}_n$, the affine group of type $\tilde{A}$ and rank n; see [**4, 6, 16**].

References

[1] R. M. Adin, Y. Roichman, 'On Degrees in the Hasse Diagram of the Strong Bruhat Order', *preprint* (2005), 1–13, available at `http://front.math.ucdavis.edu/math.CO/0505020`.

[2] G. E. Andrews, R. Askey and R. Roy, *Special functions*, Cambridge University Press, Cambridge 1999.

[3] A. Björner, 'Orderings of Coxeter Groups', *Contemp. Math.* **34** (1984), 175–195.

[4] A. Björner and F. Brenti, 'Affine permutations of type A', *Electron. J. Combin.* **3** (1996), Research paper 18, 35 pp.

[5] A. Björner and F. Brenti, 'An improved tableau criterion for Bruhat order', *Electron. J. Combin.* **3** (1996), *no. 1*, Research paper 22, 5 pp.

[6] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics 231, Springer–Verlag, New York 2005.

[7] A. Björner and T. Ekedahl, 'On the shape of Bruhat intervals', *preprint* (2005), 1–15, available at `http://front.math.ucdavis.edu/math.CO/0508022`.

[8] M. Bóna, *Combinatorics of permutations*, Chapman & Hall/CRC, Boca Raton, Florida, 2004.

[9] F. Brenti, 'Kazhdan-Lusztig Polynomials: History, Problems, and Combinatorial Invariance', *Séminaire Lotharingien de Combinatoire* **49** (2003), Art. B49b, 30 pp.

[10] F. Brenti, 'The intersection cohomology of Schubert varieties is a combinatorial invariant', *European J. Combin.* **25** (2004), 1151–1167.

[11] C. Charalambides, *Enumerative combinatorics*, Chapman & Hall/CRC, Boca Raton, Florida, 2002.

[12] F. du Cloux, 'An abstract model for Bruhat intervals' *European J. Combin.* **21** (2000), 197–222.

[13] F. du Cloux, 'Computing Kazhdan-Lusztig polynomials for arbitrary Coxeter groups', *Experiment. Math.* **11** (2002), 371–381.

[14] F. du Cloux, 'Rigidity of Schubert closures and invariance of Kazhdan-Lusztig polynomials', *Adv. Math.* **180** (2003), 146–175.

[15] M. Dyer, 'On the "Bruhat graph" of a Coxeter system', *Compositio Math.* **78** (1991), 185–191.

[16] H. Eriksson and K. Eriksson, 'Affine Weyl groups as infinite permutations', *Electron. J. Combin.* **5** (1998), Research paper 18, 32 pp.

[17] N. J. Fine, *Basic hypergeometric series and applications*, American Mathematical Society, Providence, RI, 1988.

[18] W. Fulton, *Young tableaux*, Cambridge University Press, 1999.

[19] G. Gasper and M. Rahman, *Basic hypergeometric series*, Cambridge University Press, Cambridge, 1990.

[20] I. M. Gessel and R. P. Stanley, 'Algebraic enumeration', *Handbook of combinatorics*, Vol. 2, Elsevier, Amsterdam 1995, 1021–1061.

[21] H. Hiller, *Geometry of Coxeter Groups*, Pitman (Advanced Publishing Program), Boston, London, Melbourne, 1982.

[22] J. H. Humphreys, *Reflection groups and Coxeter groups*, Cambridge University Press, Cambridge 1990.

[23] C. Krattenthaler, 'HYP and HYPQ: Mathematica packages for the manipulation of binomial sums and hypergeometric series, respectively q-binomial sums and basic hypergeometric series', *J. Symbolic Comput.* **20** (1995), 737–744.
[24] M. Marietti, 'Closed product formulas for certain R-polynomials', *European J. Combin.* **23** (2002), 57–62.
[25] M. Marietti, 'Kazdhan-Lusztig polynomials for Boolean elements in linear Coxeter systems', *preprint* (2004).
[26] C. Musili, *Representions of Finite Groups*, Hindustan Book Agency, India, 1993.
[27] M. Pagliacci and N. Zagaglia Salvi, 'Enumeration of the permutations under $(1, n)$ in the Bruhat order', *preprint* (2005).
[28] S. B. Perkins and P. J. Rowley, 'Minimal and maximal length involutions in finite Coxeter groups', *Comm. Algebra* **30** (2002), 1273–1292.
[29] S. B. Perkins and P. J. Rowley, 'Lengths of involutions in Coxeter groups', *J. Lie Theory* **14** (2004), 69–71.
[30] R. W. Richardson, 'Conjugacy classes of involutions in Coxeter groups', *Bull. Austral. Math. Soc.* **26** (1982), 1–15.
[31] T. A. Springer, 'Some remarks on involutions in Coxeter groups', *Comm. Algebra* **10** (1982), 631–636.
[32] R. P. Stanley, *Enumerative Combinatorics Vol. 1*, Cambridge University Press, 1997.

Fakultät für Mathematik, Universität Wien, Nordbergstrasse 15, A-1090 Wien, Austria
E-mail address: `alessandro.conflitti@univie.ac.at`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Overpartitions, lattice paths and Rogers-Ramanujan identities

Sylvie Corteel and Olivier Mallet

Abstract. We define the notions of successive ranks and generalized Durfee squares for overpartitions. We show how these combinatorial statistics give extensions to overpartitions of combinatorial interpretations in terms of lattice paths of the generalizations of the Rogers-Ramanujan identities due to Burge, Andrews and Bressoud. All our proofs are combinatorial and use bijective techniques. Our result includes the Andrews-Gordon identities, the generalization of the Gordon-Göllnitz identities and Gordon's theorems for overpartitions.

Résumé. Nous définissons les notions de rangs successifs et de carré de Durfee généralisé pour les overpartitions. Nous montrons comment ces statistiques combinatoires permettent d'étendre aux overpartitions des interprétations combinatoires en termes de chemins des généralisations des identités de Rogers-Ramanujan dues à Burge, Andrews et Bressoud. Toutes nos preuves sont combinatoires et utilisent des techniques bijectives. Notre résultat englobe les identités d'Andrews-Gordon, les généralisations de l'identité de Gordon-Göllnitz et les theorèmes de Gordon pour les overpartitions.

1. Introduction

The starting point of this work is a result of Lovejoy of 2003 [**25**], called Gordon's theorem for overpartitions which states that

Theorem 1.1. [**25**] *Let $\overline{B}_k(n)$ denote the number of overpartitions of n of the form $(\lambda_1, \lambda_2, \dots, \lambda_s)$, where $\lambda_j - \lambda_{j+k-1} \geq 1$ if λ_{j+k-1} is overlined and $\lambda_j - \lambda_{j+k-1} \geq 2$ otherwise. Let $\overline{A}_k(n)$ denote the number of overpartitions of n into parts not divisible by k. Then $\overline{A}_k(n) = \overline{B}_k(n)$.*

An overpartition here is a partition where the final occurrence of a part can be overlined [**16**]. For example there exist 8 overpartitions of 3

$$(3),\ (\overline{3}),\ (2,1),\ (\overline{2},1),\ (2,\overline{1}),\ (\overline{2},\overline{1}),\ (1,1,1),\ (1,1,\overline{1}).$$

Overpartitions have been recently heavily studied under different names and guises. They can be called joint partitions [**9**], or dotted partitions [**11**] and they are also closely related to 2-modular diagrams [**28**], jagged partitions [**21, 22**] and superpartitions [**20**]. Results on (for example) combinatorics of basic hypergeometric series identities [**17, 32**], q-series [**22, 25, 26**], congruences of the overpartition function [**21, 29**] and supersymmetric functions [**20**] have been discovered.

Gordon's theorem was proved in 1961 and is the following

Theorem 1.2. [**24**] *Let $B_{k,i}(n)$ denote the number of partitions of n of the form $(\lambda_1, \lambda_2, \dots, \lambda_s)$, where $\lambda_j - \lambda_{j+k-1} \geq 2$ and at most $i-1$ of the parts are equal to 1. Let $A_{k,i}(n)$ denote the number of partitions of n into parts not congruent to $0, \pm i$ modulo $2k+1$. Then $A_{k,i}(n) = B_{k,i}(n)$.*

2000 *Mathematics Subject Classification.* Primary 11P81; Secondary 05A17.
Key words and phrases. Partitions, overpartitions, Rogers-Ramanujan identities, lattice paths.
The authors are partially supported by the ACI Jeunes Chercheurs "Partitions d'entiers à la frontière de la combinatoire, des q-séries et de la théorie des nombres".

This theorem is an extension of the famous Rogers-Ramanujan identities proved by Rogers in 1894 [**31**] which correspond to the cases $k = i = 2$ and $k = 2$, $i = 1$. It is still a well known open problem to find a natural bijective proof of these identities, even though an impressive number of nearly combinatorial proofs have been published. A recent example was presented at FPSAC last year [**10**]. Lovejoy's result can be seen as an analog of Gordon's theorem, as the conditions on the $\overline{B}_k(n)$ reduce to the conditions on the $B_{k,k}(n)$ if the overpartition has no overlined parts and is indeed a partition.

Other combinatorial interpretations related to Gordon's theorem were given by Andrews and these became the Andrews-Gordon identities :

THEOREM 1.3. [**4**] *Let $C_{k,i}(n)$ be the number of partitions of n whose successive ranks lie in the interval $[-i+2, 2k-i-1]$ and let $D_{k,i}(n)$ be the number of partitions of n with $i-1$ successive Durfee squares followed by $k-i$ successive Durfee rectangles. Then*

$$A_{k,i}(n) = B_{k,i}(n) = C_{k,i}(n) = D_{k,i}(n).$$

Details can be found in [**2**, Chapter 7]. It is well understood combinatorially that $B_{k,i}(n) = C_{k,i}(n) = D_{k,i}(n)$ and that result was established by some beautiful work of Burge [**14, 15**] using some recursive arguments. This work was reinterpreted by Andrews and Bressoud [**7**] who showed that Burge's argument could be rephrased in terms of binary words and that Gordon's theorem can be established thanks to these combinatorial arguments and the Jacobi Triple product identity [**23**]. Finally Bressoud [**12**] reinterpreted these in terms of ternary words and showed some direct bijections between the objects counted by $B_{k,i}(n)$, $C_{k,i}(n)$, $D_{k,i}(n)$ and the ternary words.

The purpose of this extended abstract is therefore to extend these works [**7, 12, 14, 15**] to overpartitions to try to generalize both Gordon's theorem for overpartitions and the Andrews-Gordon identities.

Our main result is the following and is proved totally combinatorially:

THEOREM 1.4.

- *Let $\overline{B}_{k,i}(n,j)$ be the number of overpartitions of n of the form $(\lambda_1, \lambda_2, \ldots, \lambda_s)$ with j overlined parts and where $\lambda_\ell - \lambda_{\ell+k-1} \geq 1$ if $\lambda_{\ell+k-1}$ is overlined and $\lambda_\ell - \lambda_{\ell+k-1} \geq 2$ otherwise and at most $i-1$ parts are equal to 1.*
- *Let $\overline{C}_{k,i}(n,j)$ be the number of overpartitions of n with j non-overlined parts in the bottom row of their Frobenius representation and whose successive ranks lie in $[-i+2, 2k-i-1]$.*
- *Let $\overline{D}_{k,i}(n,j)$ be the number of overpartitions of n with j overlined parts and $i-1$ successive Durfee squares followed by $k-i$ successive Durfee rectangles, the first one being a generalized Durfee square/rectangle.*
- *Let $\overline{E}_{k,i}(n,j)$ be the number of paths that use four kinds of unitary steps with special (k,i)-conditions, major index n, and j South steps.*

Then $\overline{B}_{k,i}(n,j) = \overline{C}_{k,i}(n,j) = \overline{D}_{k,i}(n,j) = \overline{E}_{k,i}(n,j)$.

We use the classical q-series notations : $(a)_\infty = (a;q)_\infty = \prod_{i=0}^{\infty}(1-aq^i)$, $(a)_n = (a)_\infty/(aq^n)_\infty$ and $(a_1,\ldots,a_k;q)_\infty = (a_1;q)_\infty \ldots (a_k;q)_\infty$. The generating function $\overline{\mathcal{E}}_{k,i}(a,q) = \sum_{n,j} \overline{E}_{k,i}(n,j) q^n a^j$ is :

THEOREM 1.5.

$$\overline{\mathcal{E}}_{k,i}(a,q) = \frac{(-aq)_\infty}{(q)_\infty} \sum_{n=-\infty}^{\infty} (-1)^n a^n q^{kn^2+(k-i+1)n} \frac{(-1/a)_n}{(-aq)_n}. \tag{1.1}$$

In some cases, we can use the Jacobi Triple Product identity [**23**]:

$$(-1/z, -zq, q; q)_\infty = \sum_{n=-\infty}^{\infty} z^n q^{\binom{n+1}{2}}$$

and show that this generating function has a very nice form. For example, Corollary 1.1.

$$\overline{\mathcal{E}}_{k,i}(0,q) = \frac{(q^i, q^{2k+1-i}, q^{2k+1}; q^{2k+1})_\infty}{(q)_\infty} \tag{1.2}$$

$$\overline{\mathcal{E}}_{k,i}(1/q,q^2) = \frac{(q^2;q^4)_\infty (q^{2i-1}, q^{4k+1-2i}, q^{4k}; q^{4k})_\infty}{(q)_\infty} \tag{1.3}$$

$$\overline{\mathcal{E}}_{k,i}(1,q) = \frac{(-q)_\infty}{(q)_\infty} \sum_{j=0}^{2(k-i)} (-1)^j (q^{i+j}, q^{2k-i-j}, q^{2k}; q^{2k})_\infty \tag{1.4}$$

$$\overline{\mathcal{E}}_{k,i}(1/q,q) = \frac{(-q)_\infty}{(q)_\infty} \left((q^i, q^{2k-i}, q^{2k}; q^{2k})_\infty + (q^{i-1}, q^{2k+1-i}, q^{2k}; q^{2k})_\infty\right) \tag{1.5}$$

Hence our result gives a general view of different problems on partitions and overpartitions and shows how they are related.

- The case $a \to 0$ corresponds to the Andrews-Gordon identities [**4**].
- The case $q \to q^2$ and $a \to 1/q$ corresponds to Andrew's generalization of the Gordon-Göllnitz identities [**5, 7**].
- The cases $a \to 1$ and $i = k$ and $a \to 1/q$ and $i = 1$ correspond to the two Gordon's theorems for overpartitions of Lovejoy [**25**].

Therefore our extension of the work on the Andrews-Gordon identities [**7, 12, 14, 15**] to the case of overpartitions includes these identities, but it also includes Andrew's generalization of the Gordon-Göllnitz identities and Gordon's theorems for overpartitions.

We start by some definitions in Section 2. In Section 3 we present the paths counted by $\overline{E}_{k,i}(n,j)$ and compute the generating function. In Section 4 we present a direct bijection between the paths counted by $\overline{E}_{k,i}(n,j)$ and the overpartitions counted by $\overline{C}_{k,i}(n,j)$. In Section 5 we present a recursive bijection between the paths counted by $\overline{E}_{k,i}(n,j)$ and the overpartitions counted by $\overline{B}_{k,i}(n,j)$. We also give a generating function proof. In Section 6, we present a combinatorial argument that shows that the paths counted by $\overline{E}_{k,i}(n,j)$ and the overpartitions counted by $\overline{D}_{k,i}(n,j)$ are in bijection. All these bijections are refinements of Theorem 1.4. The number of the peaks of the paths will correspond respectively to the number of columns of the Frobenius representations, the number of weighted pairs and the size of the generalized Durfee square. We conclude in Section 7 with open further questions.

Due to the length of this extended abstract, we will most of the time present the sketch of the proofs. More details can be found in [**19, 30**].

2. Definitions on overpartitions

We will define all the notions in terms of overpartitions. We refer to [**2**] for definitions for partitions. In all of the cases the definitions coincide when the overpartition has no overlined parts.

An overpartition of n is a non-increasing sequence of natural numbers whose sum is n in which the final occurrence (equivalently, the first occurrence) of a number may be overlined. Alternatively n can be called the weight of the overpartition. Since the overlined parts form a partition into distinct parts and the non-overlined parts form an ordinary partition, the generating function of overpartitions is $\frac{(-q)_\infty}{(q)_\infty}$.

The *multiplicity* of the part j of an overpartition, denoted by f_j, is the number of occurrences of this part. We overline the multiplicity if the part appears overlined. For example, the multiplicity of the part 4 in the overpartition $(6,6,5,4,4,\overline{4},3,\overline{1})$ is $f_4 = \overline{3}$. The multiplicity sequence is the sequence $(f_1, f_2, \ldots)$. For example the previous overpartition has multiplicity sequence $(\overline{1},0,1,\overline{3},1,2)$.

The Frobenius representation of an overpartition [**16, 27**] of n is a two-rowed array

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_N \\ b_1 & b_2 & \ldots & b_N \end{pmatrix}$$

where $(a_1,\ldots,a_N)$ is a partition into distinct nonnegative parts and $(b_1,\ldots,b_N)$ is an overpartition into nonnegative parts where the first occurrence of a part can be overlined and $N + \sum(a_i + b_i) = n$.

We now define the successive ranks.

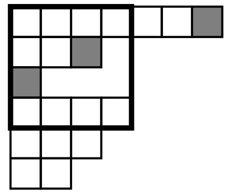


Figure 1. The generalized Durfee square of $\lambda = (\overline{7}, 4, 3, \overline{3}, 2, \overline{1})$ has side 4.

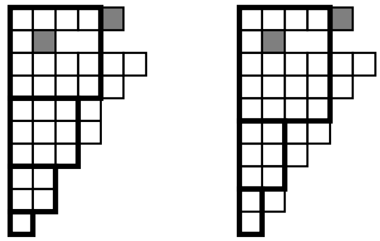


Figure 2. Successive Durfee squares and successive Durfee rectangles of $(6, 5, \overline{5}, 4, 4, 3, 2, 2, \overline{2}, 1)$.

Definition 2.1. The *successive ranks* of an overpartition can be defined from its Frobenius representation. If an overpartition has Frobenius representation $\begin{pmatrix} a_1 & a_2 & \cdots & a_N \\ b_1 & b_2 & \cdots & b_N \end{pmatrix}$ then its ith successive rank r_i is $a_i - b_i$ minus the number of non-overlined parts in $\{b_{i+1}, \ldots, b_N\}$.

This definition in an extension of Lovejoy's definition of the rank [**27**]. For example, the successive ranks of $\begin{pmatrix} 7 & 4 & 2 & 0 \\ \overline{3} & 3 & 1 & \overline{0} \end{pmatrix}$ are $(2, 0, 1, 0)$.

We say that the *generalized Durfee square* of an overpartition λ has side N if N is the largest integer such that the number of overlined parts plus the number of non-overlined parts greater or equal to N is greater than or equal to N (see Figure 1). Thanks to the Algorithm Z [**8**], we can easily show that there exists a bijection between overpartitions whose Frobenius representation has N columns and whose bottom line has j overlined parts and overpartitions with generalized Durfee square of size N and $N - j$ overlined parts. See [**19**] for details. The generating function of overpartitions with generalized Durfee square of size N where the exponent of q counts the weight and the exponent of a the number of overlined parts is

$$\frac{a^N q^{\binom{N+1}{2}} (-1/a)_N}{(q)_N (q)_N}.$$

Definition 2.2. The *successive Durfee squares* of an overpartition are its generalized Durfee square and the successive Durfee squares of the partition below the generalized Durfee square, if we represent the partition as in Figure 1, with the overlined parts above the non-overlined ones. We can also define similarly the *successive Durfee rectangles* by dissecting the overpartition with $d \times (d+1)$-rectangles instead of squares.

These definitions imply that

$$\sum_{n_1 \geq \ldots \geq n_{k-1} \geq 0} \frac{q^{\binom{n_1+1}{2} + n_i + \ldots + n_{k-1}} (-1/a)_{n_1} a^{n_1}}{(q)_{n_1}} \left(q^{n_2^2} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}_q \right) \left(q^{n_3^2} \begin{bmatrix} n_2 \\ n_3 \end{bmatrix}_q \right) \cdots \left(q^{n_{k-1}^2} \begin{bmatrix} n_{k-2} \\ n_{k-1} \end{bmatrix}_q \right) \tag{2.1}$$

where

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q)_n}{(q)_k (q)_{n-k}}$$

is the generating function of overpartitions with $i - 1$ successive Durfee squares followed by $k - i$ successive Durfee rectangles, the first one being a generalized Durfee square/rectangle.

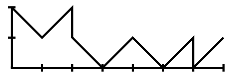


FIGURE 3. This path has four peaks : two NES peaks (located at $(2,2)$ and $(6,1)$) and two NESE peaks (located at $(4,1)$ and $(7,1)$). Its major index is $2+4+6+7=19$.

3. Paths and generating function

This part is an extension of papers of Andrews and Bressoud [**7, 12**] based on ideas of Burge [**12**]. We study paths in the first quadrant that use four kinds of unitary steps :

- North-East NE : $(x,y)\to(x+1,y+1)$,
- South-East SE : $(x,y)\to(x+1,y-1)$,
- South S : $(x,y)\to(x,y-1)$,
- East E : $(x,0)\to(x+1,0)$.

The *height* corresponds to the y-coordinate. A South step can only appear after a North-East step and an East step can only appear at height 0. The paths must end with a North-East or South step. A *peak* is a vertex preceeded by a North-East step and followed by a South step (in which case it will be called a *NES peak*) or by a South-East step (in which case it will be called a *NESE peak*). If the path ends with a North-East step, its last vertex is also a NESE peak. The *major index* of a path is the sum of the x-coordinates of its peaks (see Figure 3 for an example). When the paths have no South steps, this is the definition of the paths in [**12**].

Let $\overline{E}_{k,i}(n,j,N)$ be the number of such paths of major index n with N peaks, j South steps that start at height $k-i$ and whose height is less than k. Let $\overline{\mathcal{E}}_{k,i}(N)$ be the generating function of those paths, that is $\overline{\mathcal{E}}_{k,i}(N)=\overline{\mathcal{E}}_{k,i}(N,a,q)=\sum_{n,j}\overline{E}_{k,i}(n,j,N)a^jq^n$.

Then

PROPOSITION 3.1.

$$\begin{aligned}
\overline{\mathcal{E}}_{k,i}(N) &= q^N\overline{\mathcal{E}}_{k,i+1}(N)+q^N\overline{\Gamma}_{k,i-1}(N); \qquad i<k\\
\overline{\Gamma}_{k,i}(N) &= q^N\overline{\Gamma}_{k,i-1}(N)+(a+q^{N-1})\overline{\mathcal{E}}_{k,i+1}(N-1); \quad 0<i<k\\
\overline{\mathcal{E}}_{k,k}(N) &= \frac{q^N}{1-q^N}\overline{\Gamma}_{k,k-1}(N)
\end{aligned}$$

$$\overline{\mathcal{E}}_{k,i}(0)=1 \qquad \overline{\Gamma}_{k,0}(N)=0$$

PROOF. We prove that by induction on the length of the path. If the path is empty, then its major index is 0 and $N=0$. Moreover if $N=0$ the only path counted in $\overline{\mathcal{E}}_{k,i}(0)$ is the empty path. If the path is not empty, then we take off its first step. If $i<k$, then a path counted in $\overline{\mathcal{E}}_{k,i}(N)$ starts with a North-East (defined by $q^N\overline{\Gamma}_{k,i-1}(N)$) or a South-East step ($q^N\overline{\mathcal{E}}_{k,i+1}(N)$). If $i>0$, $\overline{\Gamma}_{k,i}(N)$ is the generating function of paths counted in $\overline{\mathcal{E}}_{k,i+1}(N)$ where the first North-East step was deleted. These paths can start with a North-East step ($q^N\overline{\Gamma}_{k,i-1}(N)$), a South step ($a\overline{\mathcal{E}}_{k,i+1}(N-1)$) or a South-East step ($q^{N-1}\overline{\mathcal{E}}_{k,i+1}(N-1)$). If $i=k$ then a path counted in $\overline{\mathcal{E}}_{k,k}(N)$ starts with a North-East ($q^N\overline{\Gamma}_{k,k-1}(N)$) or an East step ($q^N\overline{\mathcal{E}}_{k,k}(N)$). The height of the paths is less than k, therefore no path which starts at height $k-1$ can start with a North-East step and $\overline{\Gamma}_{k,0}(N)=0$. $\square$

These recurrences uniquely define the series $\overline{\mathcal{E}}_{k,i}(N)$ and $\overline{\Gamma}_{k,i}(N)$. We get that :

THEOREM 3.1.

$$\begin{aligned}
\overline{\mathcal{E}}_{k,i}(N) &= a^Nq^{\binom{N+1}{2}}(-1/a)_N\sum_{n=-N}^{N}(-1)^n\frac{q^{kn^2+n(k-i)-\binom{n}{2}}}{(q)_{N-n}(q)_{N+n}}\\
\overline{\Gamma}_{k,i}(N) &= a^Nq^{\binom{N}{2}}(-1/a)_N\sum_{n=-N}^{N-1}(-1)^n\frac{q^{kn^2+n(k-i)-\binom{n+1}{2}}}{(q)_{N-n-1}(q)_{N+n}}
\end{aligned}$$

The proof is omitted. It uses simple algebraic manipulation to prove that these generating functions satisfy the recurrence relations of Proposition 3.1.

We just need a proposition (which is in fact a aprticular case of the q-Gauss identity [**23**]) that can be proved combinatorially and analytically [**19**] to prove Theorem 1.5.

Proposition 3.2. *For any $n \in \mathbb{Z}$*

$$\sum_{N \geq |n|} \frac{(-azq)_n(-q^n/a)_{N-n} q^{\binom{N+1}{2}-\binom{n+1}{2}} z^{N-n} a^{N-n}}{(zq)_{N+n}(q)_{N-n}} = \frac{(-azq)_\infty}{(zq)_\infty}.$$

Summing on N using the previous proposition we get

$$\sum_{N \geq 0} \overline{\mathcal{E}}_{k,i}(N) = \frac{(-aq)_\infty}{(q)_\infty} \sum_{n=-\infty}^{\infty} (-1)^n a^n q^{kn^2+(k-i+1)n} \frac{(-1/a)_n}{(-aq)_n}.$$

This is equation (1.1).

4. Paths and successive ranks

This section is a generalization of Bressoud's correspondence for partitions presented in [**12**]. The aim of this section is the following:

Proposition 4.1. *There exists a one-to-one correspondence between the paths of major index n with j south steps counted by $\overline{E}_{k,i}(n,j)$ and the overpartitions of n zith j non-overlined parts in the bottom line of their Frobenius representation and whose successive ranks lie in $[-i+2, 2k-i-1]$ counted by $\overline{C}_{k,i}(n,j)$. This correspondence is such that the paths have N peaks if and only if the Frobenius representation of the overpartition has N columns.*

Given a lattice path which starts at $(0,a)$ and a peak (x,y) with u South steps to its left, we map this peak to the pair (s,t) where

$$s = (x+a-y+u)/2$$
$$t = (x-a+y-2-u)/2$$

if there are an even number of East steps to the left of the peak, and

$$s = (x+a+y-1+u)/2$$
$$t = (x-a-y-1-u)/2$$

if there are an odd number of East steps to the left of the peak. Moreover, we overline t if the peak is a NESE peak. In both cases, s and t are integers and we have $s+t+1 = x$. In the case of partitions treated in [**12**], u is always 0.

Let N be the number of peaks in the path and j the number of South steps of the paths. If the ith peak from the right has coordinates (x_i, y_i) and the corresponding pair is (s_i, t_i), then we show in [**19**] that the sequence $(s_1, s_2, \ldots, s_N)$ is a partition into distinct nonnegative parts and the sequence $(t_1, t_2, \ldots, t_N)$ is an overpartition into nonnegative parts with j non-overlined parts. Therefore $\begin{pmatrix} s_1 & s_2 & \cdots & s_N \\ t_1 & t_2 & \cdots & t_N \end{pmatrix}$ is the Frobenius representation of an overpartition whose weight is

$$\sum_{i=1}^{N}(s_i + t_i + 1) = \sum_{i=1}^{N} x_i$$

i.e. the major index of the corresponding path.

As an example, the path in Figure 4 corresponds to the partition $\begin{pmatrix} 14 & 11 & 6 & 4 & 2 \\ 7 & \overline{6} & \overline{5} & 4 & \overline{3} \end{pmatrix}$.

The peaks all have height at least one, thus for a peak (x,y) which is preceeded by an even number of East steps, we have :

$$1 \leq y = a+1+t-s+u$$
$$\Leftrightarrow s-t-u \leq a$$
$$\Leftrightarrow \text{the corresponding successive rank is} \leq a$$

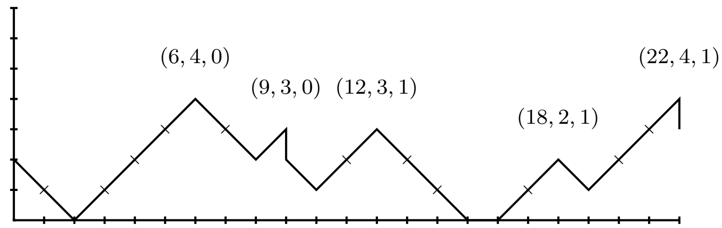


FIGURE 4. Illustration of the correspondence between paths and successive ranks. The values of x, y and u are given for each peak.

and if the peak is preceeded by and odd number of East steps, we have :

$$\begin{aligned} & 1 \le y = s - t - u - a \\ \Leftrightarrow\ & s - t - u \ge a + 1 \\ \Leftrightarrow\ & \text{the corresponding successive rank is} \ge a + 1 \end{aligned}$$

Thus, given a Frobenius representation of an overpartition and a nonnegative integer a, there is a unique corresponding path which starts at $(0, a)$.

In our paths, all peaks have height at most $k-1$ and $a = k-i$, therefore in the first case the successive rank $r \in [-i+2, k-i]$ and in the second case $r \in [k-i+1, 2k-i-1]$.

The map is easily reversible. This proves Proposition 4.1.

5. Paths and multiplicities

Recall that $\overline{B}_{k,i}(n,j)$ is the number of overpartitions λ of n with j overlined parts such that for all ℓ,

$$\begin{cases} \lambda_\ell - \lambda_{\ell+k-1} \le \begin{cases} 1 & \text{if } \lambda_{\ell+k-1} \text{ is overlined} \\ 2 & \text{otherwise} \end{cases} \\ f_1 < i \end{cases}$$

or equivalently

$$\begin{cases} \forall \ell,\ f_\ell + f_{\ell+1} \le \begin{cases} k+1 & \text{if a part } \ell \text{ is overlined} \\ k & \text{otherwise} \end{cases} \\ f_1 < i \end{cases}$$

The aim of this section is the following:

PROPOSITION 5.1. *There exists a one-to-one correspondence between the paths counted by $\overline{E}_{k,i}(n,j)$ and the overpartitions counted by $\overline{B}_{k,i}(n,j)$. This correspondence is such that the paths have N peaks if and only if the overpartition has N weighted pairs.*

We will first give a generating function proof of that proposition (without the refinement). Then we will give a combinatorial proof which is a generalization of Burge's correspondence for partitions presented in [**14**].

5.1. A generating function proof.

Let $\overline{\mathcal{B}}_{k,i}(a,q) = \sum_{n\ge 0} \overline{B}_{k,i}(n,j) a^j q^n$. We prove that

PROPOSITION 5.2.

$$\overline{\mathcal{B}}_{k,i}(a,q) = \overline{\mathcal{E}}_{k,i}(a,q)$$

PROOF. We generalize Lovejoy's proof of Theorem 1.1 of [**25**]. Let

$$\begin{aligned} J_{k,i}(a,x,q) &= H_{k,i}(a,xq,q) - axqH_{k,i-1}(a,xq,q) \\ H_{k,i}(a,x,q) &= \sum_{n=0}^{\infty} \frac{x^{kn} q^{kn^2+n-in} a^n (1-x^i q^{2ni})(axq^{n+1})_\infty (1/a)_n}{(q)_n (xq^n)_\infty}. \end{aligned}$$

Andrews showed in [**2**, p. 106-107] that for $2 \le i \le k$,

$$\begin{aligned} J_{k,i}(a,x,q) - J_{k,i-1}(a,x,q) &= (xq)^{i-1} J_{k,k-i+1}(a,xq,q) - a(xq)^{i-1} J_{k,k-i+2}(a,xq,q) \\ J_{k,1}(a,x,q) &= J_{k,k}(a,x,q). \end{aligned}$$

This functional equation of $J_{k,i}(a,x,q)$ implies that

$$\overline{\mathcal{B}}_{k,i}(a,q) = J_{k,i}(-a,1,q).$$

Hence

$$\begin{aligned}
\overline{\mathcal{B}}_{k,i}(a,q) &= \frac{(-aq)_\infty}{(q)_\infty}\sum_{n=0}^{\infty}(-1)^n a^n \frac{q^{kn^2+n(k-i+1)}(-1/a)_n(1-q^{(2n+1)i})}{(-aq)_{n+1}} \\
&\quad + aq\frac{(-aq)_\infty}{(q)_\infty}\sum_{n=0}^{\infty}(-1)^n a^n \frac{q^{kn^2+n(k-i+2)}(-1/a)_n(1-q^{(2n+1)(i-1)})}{(-aq)_{n+1}} \\
&= \frac{(-aq)_\infty}{(q)_\infty}\sum_{n=0}^{\infty}(-1)^n a^n \frac{q^{kn^2+n(k+1)}(-1/a)_n(q^{-in}+aq^{1-(i-1)n})}{(-aq)_{n+1}} \\
&\quad - \frac{(-aq)_\infty}{(q)_\infty}\sum_{n=0}^{\infty}(-1)^n a^n \frac{q^{kn^2+n(k+1)}(-1/a)_n(q^{(n+1)i}+aq^{(n+1)(i-1)+1})}{(-aq)_{n+1}} \\
&= \frac{(-aq)_\infty}{(q)_\infty}\left(\sum_{n=0}^{\infty}(-1)^n a^n \frac{q^{kn^2+n(k+1-i)}(-1/a)_n}{(-aq)_n} - \sum_{n=0}^{\infty}(-1)^n a^{n+1} \frac{q^{kn^2+n(k+i)+i}(-1/a)_{n+1}}{(-aq)_{n+1}}\right) \\
&= \frac{(-aq)_\infty}{(q)_\infty}\left(\sum_{n=0}^{\infty}(-1)^n a^n \frac{q^{kn^2+n(k+1-i)}(-1/a)_n}{(-aq)_n} + \sum_{n=-\infty}^{-1}(-1)^n a^{-n} \frac{q^{kn^2+n(k-i)}(-1/a)_{-n}}{(-aq)_{-n}}\right) \\
&= \frac{(-aq)_\infty}{(q)_\infty}\sum_{n=-\infty}^{\infty}(-1)^n a^n \frac{q^{kn^2+n(k+1-i)}(-1/a)_n}{(-aq)_n} = \overline{\mathcal{E}}_{k,i}(a,q)
\end{aligned}$$

□

5.2. A combinatorial proof. This part is a generalization of [**14**, Section 3]. Like Burge, we define operations on overpartitions represented by their multiplicity sequence.

The operation α is defined as follows. We divide the overpartition into $(\ell+1)$-tuples of the form $(f_m,\ldots,f_{m+\ell})$ with $\ell \geq 1$ starting at the smallest part. When we find a multiplicity $f_m > 0$, we open a parenthesis to its left. If f_m is not overlined then we close the parenthesis to the right of f_{m+1}. Otherwise, we look for the next non-overlined multiplicity, say f_p. If $f_p = 0$ then we close the parenthesis to its right, otherwise we close the parenthesis to the right of f_{p+1}. Then we look for the next positive multiplicity, and so on. Finally, for each $(\ell+1)$-tuple $(f_m,\ldots,f_{m+\ell})$, we do :

- $f_m \leftarrow f_m - 1$
- $f_{m+\ell} \leftarrow f_{m+\ell} + 1$
- if f_m is overlined, we remove its overlining and we overline the smallest non-overlined multiplicity in the $(\ell+1)$-tuple.

The operation β (resp. δ) consists in setting $f_0 = 1$ (resp. $f_0 = \overline{1}$) and applying α.

The inverse operation α^{-1} is performed by first dividing the overpartition into $(\ell+1)$-tuples of the form $(f_m,\ldots,f_{m+\ell})$, with $\ell \geq 1$ starting at the largest part, such that :

- $f_{m+\ell} > 0$
- f_m is not overlined
- f_{m+p} is overlined for $1 \leq p \leq \ell - 1$

(for an example, see the first line of Table 1, which corresponds to the overpartition $(5,5,\overline{5},\overline{4},3,\overline{2})$) and then doing for each $(\ell+1)$-tuple :

- if $f_{m+\ell} = \overline{1}$:
  - remove the overlining of $f_{m+\ell}$
  - underline f_m
- else if $\ell > 1$:
  - remove the overlining of $f_{m+\ell-1}$
  - underline f_m
- $f_{m+\ell} \leftarrow f_{m+\ell} - 1$
- $f_m \leftarrow f_m + 1$

Operation	N	i	0	1	2	3	4	5
α^{-1}	3	1	0	(0	$\overline{1}$)	(1	$\overline{1}$	$\overline{3}$)
δ^{-1}	3	2	(0	$\overline{1}$)	(0	$\overline{2}$)	(1	$\overline{2}$)
α^{-1}	2	2	0	0	(1	$\overline{1}$)	(2	$\overline{1}$)
α^{-1}	2	3	0	(0	$\overline{2}$)	(0	$\overline{3}$)	
α^{-1}	2	4	0	(1	$\overline{1}$)	(1	$\overline{2}$)	
β^{-1}	2	4	(0	$\overline{2}$)	0	(2	$\overline{1}$)	
δ^{-1}	2	3	(0	$\overline{1}$)	(0	$\overline{3}$)		
α^{-1}	1	3	0	0	(1	$\overline{2}$)		
α^{-1}	1	4	0	0	(2	$\overline{1}$)		
α^{-1}	1	4	0	(0	$\overline{3}$)			
α^{-1}	1	4	0	(1	$\overline{2}$)			
α^{-1}	1	4	0	(2	$\overline{1}$)			
β^{-1}	1	4	(0	$\overline{3}$)				
β^{-1}	1	3	(0	$\overline{2}$)				
δ^{-1}	1	2	(0	$\overline{1}$)				
	0	2	0					

TABLE 1. Reduction of the overpartition $(5, 5, \overline{5}, \overline{4}, 3, \overline{2})$.

If there is an $(\ell + 1)$-tuple $(f_0, \ldots, f_\ell)$, the operation α^{-1} will produce a zero part, which may be overlined or not. The operation β^{-1} (resp. δ^{-1}) consists in applying α^{-1} and removing the non-overlined (resp. overlined) zero part.

The inverse operations allow us to define a reduction process for overpartitions which is similar to Burge's reduction for partitions [**14**]. An example is shown on Table 1.

Let $\overline{B}_{k,i}(n, j, N)$ be the number of partitions counted by $\overline{B}_{k,i}(n, j)$ such that $N = \sum_{(\ell+1)-tuples} \ell$. We call N the number of weighted pairs (for partitions, we always have $\ell = 1$ and N is the number of pairs [**14**]). Let $\overline{\mathcal{B}}_{k,i}(N) = \sum_{n,j} \overline{B}_{k,i}(n, j, N) q^n a^j$. Starting with an overpartition counted in $\overline{\mathcal{B}}_{k,i}(N)$, when we apply the reduction the weight will decrease by N. We can only apply a β^{-1} or δ^{-1} if $i > 0$. We show in [**19**] that when we apply α^{-1} (resp. β^{-1}, (resp. δ^{-1})), N stays the same (resp. stays the same or decreases by 1 [in which case the next reduction is an α^{-1}], (resp. decreases by 1)) and i increases by 1 (resp. decreases by 1, (resp. stays the same)). These observations imply that $\overline{\mathcal{B}}_{k,i}(N)$ satisfies exactly the same recurrences relations as $\overline{\mathcal{E}}_{k,i}(N)$ defined in Proposition 3.1. Therefore $\overline{\mathcal{B}}_{k,i}(N) = \overline{\mathcal{E}}_{k,i}(N)$. This proves Proposition 5.1.

6. Paths and successive Durfee squares

We will prove here that

PROPOSITION 6.1.
$$\frac{q^{\binom{n_1+1}{2}+n_2^2+\cdots+n_{k-1}^2+n_i+\cdots+n_{k-1}}(-1/a)_{n_1}a^{n_1}}{(q)_{n_1-n_2}\cdots(q)_{n_{k-2}-n_{k-1}}(q)_{n_{k-1}}}$$
is the generating function of the paths counted by major index and number of South steps starting at height $k - i$, whose height is less than k and having n_j peaks of relative height $\geq j$ for $1 \leq j \leq k - 1$.

The relative height of a peak was defined by Bressoud in [**12**] when he proved that

LEMMA 6.1 (Bressoud).
$$\frac{q^{n_1^2+n_2^2+\cdots+n_{k-1}^2+n_i+\cdots+n_{k-1}}}{(q)_{n_1-n_2}\cdots(q)_{n_{k-2}-n_{k-1}}(q)_{n_{k-1}}}$$
is the generating function of the paths with no South steps starting at height $k - i$, whose height is less than k and having n_j peaks of relative height $\geq j$ for $1 \leq j \leq k - 1$.

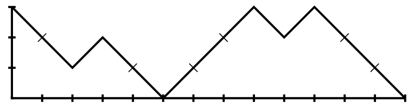


Figure 5. Example of a path.

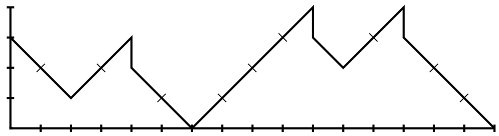


Figure 6. Effect of the "volcanic uplift".

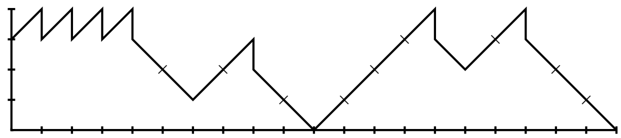


Figure 7. After adding the $n_1 - n_2 = 4$ NES peaks of relative height one.

An example of such a path, taken from [**12**], is shown on Figure 5.

A mountain in a path is a portion of the path that starts at the beginning of the path or at height 0 stays above the x-axis and ends at height 0. We recall Bressoud's definition of the relative height of a peak [**12**]. We first map each peak of the path to a pair (y, y') where y is the height of the peak and y' is defined as follows. In each mountain, we choose the leftmost peak of maximal height relative to that mountain. For this peak, y' is the minimal height over all vertices to its left. Then, if there are any unchosen peaks left, we cut all the mountains off at height one. This may divide some moutains into several mountains relative to height one. For each mountain relative to height one in which no peaks have been chosen, we choose the leftmost peak of maximal height relative to that mountain ; for this peak, y' is the greater of one and the minimal height over all vertices to its left. We continue cutting the mountains off at height 2, 3, etc. until all peaks have been chosen.

Definition 6.2. [**12**] The *relative height of a peak* is then defined by $y - y'$.

This definition extends naturally to overpartitions. We can now move on to the proof of Proposition 6.1.

Proof. We prove the proposition using Bressoud's result. We consider a path counted by

$$\frac{q^{n_2^2+\cdots+n_{k-1}^2+n_i+\cdots+n_{k-1}}}{(q)_{n_2-n_3}\cdots(q)_{n_{k-2}-n_{k-1}}(q)_{n_{k-1}}}$$

where $2 \le i \le k$. Thanks to Lemma 6.1, we know that this path starts at height $k-i$, its height is less than $k-1$ and having n_j peaks of relative height $\ge j-1$ for $2 \le j \le k-1$. We first insert a NES peak at each peak (see Figure 6). This "volcanic uplift" operation increases the weight of the path by

$$1 + 2 + \cdots + n_2 = \binom{n_2+1}{2}$$

and the relative height of each peak by one.

We then insert $n_1 - n_2$ NES peaks at the beginning of the path (see Figure 7). These new peaks have total weight $\binom{n_1-n_2+1}{2}$ and they increase the weight of each of the old peaks by $n_1 - n_2$. Altogether, the two operations introduce a factor

$$q^{\binom{n_2+1}{2}+\binom{n_1-n_2+1}{2}+n_2(n_1-n_2)} = q^{\binom{n_1+1}{2}}.$$

If $i = 2$, so that the path starts at $(0, k-2)$, we have the option to introduce an extra step at the beginning of the path, from $(0, k-1)$ to $(1, k-2)$. This introduces the factor q^{n_1}.

The factor $(-1)_{n_1}$ corresponds to a partition into distinct parts which lie in $[0, n_1 - 1]$. If this partition contains a part $j-1$ ($1 \le j \le n_1$), we transform the jth NES peak from the right into a NESE peak (see Figure 8). This operation increases the weight of the path by $j - 1$.

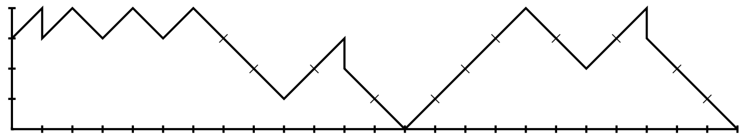


FIGURE 8. Effect of transforming some NES peaks into NESE peaks. The partition into distinct parts is $(5, 4, 3, 1)$.

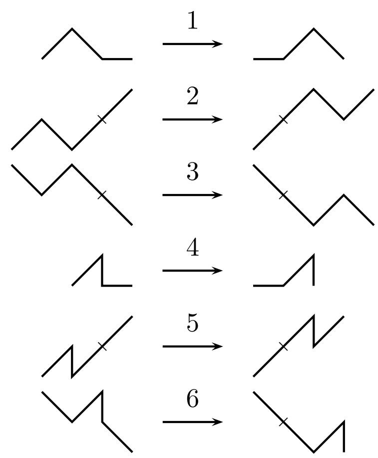


FIGURE 9. The rules for moving peaks.

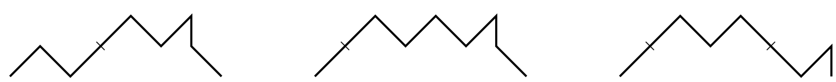


FIGURE 10. We want to move the leftmost peak to the right twice, but after the first move, we come up against a sequence of adjacent peaks. We then move the rightmost peak in this sequence.

The factor $\frac{1}{(q)_{n_1-n_2}}$ corresponds to a partition $(b_1, b_2, \ldots, b_{n_1-n_2})$ where $b_1 \geq b_2 \geq \ldots \geq b_{n_1-n_2} \geq 0$. For $1 \leq j \leq n_1 - n_2$, we move the jth peak of relative height one from the right b_j times according to the rules illustrated in Figure 9. See [**19**] for details.

When we move a peak, it can meet the next peak to the right. We say that a peak (x, y) meets a peak (x', y') if

$$x' - x = \begin{cases} 2 \text{ if } (x, y) \text{ is a NESE peak} \\ 1 \text{ if } (x, y) \text{ is a NES peak} \end{cases} .$$

If this happens, we abandon the peak we have been moving and move the next one. If we come up against a sequence of adjacent peaks, we move the rightmost peak in the sequence (see Figure 10).

It can be shown that the distribution of relative heights is not modified by the operations of Figure 9 and that the construction procedure is uniquely reversible. □

The multiple series

$$\sum_{n_1 \geq \ldots \geq n_{k-1} \geq 0} \frac{q^{\binom{n_1+1}{2} + n_2^2 + \cdots + n_{k-1}^2 + n_i + \cdots + n_{k-1}} (-1/a)_{n_1} a^{n_1}}{(q)_{n_1-n_2} \cdots (q)_{n_{k-2}-n_{k-1}} (q)_{n_{k-1}}}$$

can be re-expressed as (2.1), which is the generating function of overpartitions with $i-1$ successive Durfee squares followed by $k-i$ successive Durfee rectangles, the first one being a generalized Durfee square/rectangle.

7. Conclusion

We showed in this work how the combinatorial interpretation of the Andrews-Gordon identities can be generalized to the case of overpartitions, when the combinatorial statistics (successive ranks, generalized Durfee square, weighted pairs) are defined properly. There exist other generalizations of the Rogers-Ramanujan identities, see for example [**13**]. It was shown that the combinatorial interpretation in terms of lattice paths can also be done for these identities [**1, 12, 14, 15**]. Our work can also be extended in that direction and the results are presented in [**18**]. Finally there exists an extension of the concept of successive ranks for partitions due to Andrews, Baxter, Bressoud, Burge, Forrester and Viennot [**6**] and our goal now is to extend that notion to overpartitions.

References

[1] A. K. Agarwal and D. M. Bressoud, Lattice paths and multiple basic hypergeometric series. *Pacific J. Math.* **135** (1989) 209–228.
[2] G. E. Andrews, The theory of partitions. Cambridge University Press, Cambridge, 1998.
[3] G. E. Andrews, An analytic generalization of the Rogers-Ramanujan identities for odd moduli. *Proc. Nat. Acad. Sci USA* **71** (1974) 4082–4085.
[4] G. E. Andrews, Sieves in the theory of partitions. *Amer. J. Math.* **94**:1214-1230 (1972).
[5] G. E. Andrews, A generalization of the Göllnitz-Gordon partition theorems. *Proc. Amer. Math. Soc.* **8**:945–952 (1967).
[6] G. E. Andrews, R. J. Baxter, D. M. Bressoud, W. H. Burge, P. J. Forrester and G. Viennot, Partitions with prescribed hook differences, *Europ. J. Combinatorics* (1987) **8** 341–350.
[7] G. E. Andrews and D. Bressoud, On The Burge correspondence between partitions and binary words. Number Theory (Winnipeg, Man., 1983). *Rocky Mountain J. Math.* **15** (1985), no. 2, 225–233.
[8] G. E. Andrews and D. M. Bressoud, Identities in Combinatorics III : further aspects of ordered set sorting. *Discrete Math* **49** (1984) 222–236.
[9] C. Bessenrodt and I. Pak: Partition congruences by involutions, *European J. Comb.* **25** (2004) 1139–1149.
[10] C. Boulet and I. Pak, A combinatorial proof of the Rogers-Ramanujan identities, to appear in *J. Combin. Theory Ser. A* (2005) (Presented at FPSAC 2005).
[11] F. Brenti, Determinants of Super-Schur Functions, Lattice Paths, and Dotted Plane Partitions *Advances in Math.*, **98** (1993), 27-64.
[12] D. Bressoud, Lattice paths and the Rogers-Ramanujan identities. Number Theory, Madras 1987, 140–172, *Lecture Notes in Math.* **1395**, Springer, Berlin, 1989.
[13] D. Bressoud, A generalization of the Rogers-Ramanujan identities for all moduli. *J. Comb. Theory* **27**:64–68 (1979).
[14] W. H. Burge, A correspondence between partitions related to generalizations of the Rogers-Ramanujan identities. *Discrete Math.* **34** (1981), no. 1, 9–15.
[15] W. H. Burge, A three-way correspondence between partitions. *European J. Combin.* **3** (1982), no. 3, 195–213.
[16] S. Corteel, J. Lovejoy, Overpartitions. *Trans. Amer. Math. Soc.* **356** (2004), no. 4, 1623–1635 (Presented at FPSAC 2003).
[17] S. Corteel, J. Lovejoy, Frobenius partitions and the combinatorics of Ramanujan's ${}_1\psi_1$ summation, *J. Comb. Theory Ser. A* **97** (2002), 177–183.
[18] S. Corteel, J. Lovejoy and O. Mallet, Extension to overpartitions of Rogers-Ramanujan identities for even moduli, in preparation.
[19] S. Corteel and O. Mallet, Overpartitions, lattice paths and Rogers-Ramanujan identities, long version in preparation.
[20] P. Desrosiers, L. Lapointe, and P. Mathieu, A symmetric function theory in superspace, Presented at FPSAC 2005.
[21] J.-F. Fortin, P. Jacob, and P. Mathieu, Generating function for K-restricted jagged partitions. *Electron. J. Comb.* **12**, No.1, Research paper R12, 17 p., (2005).
[22] J.-F. Fortin, P. Jacob, and P. Mathieu, Jagged partitions, *Ramanujan Jal* (2005), to appear.
[23] G. Gaspar and M. Rahman, Basic hypergeometric series, Cambridge University Press(1990).
[24] B. Gordon, A combinatorial generalization of the Rogers-Ramanujan identities, *Amer. J. Math.* **83** (1961), 393–399 .
[25] J. Lovejoy, Gordon's theorem for overpartitions. *J. Combin. Theory Ser. A* **103** (2003), no. 2, 393–401.
[26] J. Lovejoy, Overpartition theorems of the Rogers-Ramanujan type *J. London Math. Soc.* **69** (2004), 562–574.
[27] J. Lovejoy, Rank and conjugation for the Frobenius representation of an overpartition. *Ann. Combin.* **9** (2005) 321–334.
[28] P. A. MacMahon, Combinatory Analysis, vol. 2, Cambridge Univ. Press, Cambridge, 1916.
[29] K. Mahlburg, The overpartition function modulo small powers of 2, *Discrete Mathematics* **286** (2004), no. 3, 263–267.
[30] O. Mallet, Rangs successifs, chemins et identités de type Rogers-Ramanujan, *Master thesis, Paris 6* (2005).
[31] L. J. Rogers, Second memoir on the expansion of certain infinite products, *Proc. London Math. Soc.* **25** (1894), 318–343.
[32] A. J. Yee, Combinatorial proofs of Ramanujan's ${}_1\psi_1$ summation and the q-Gauss summation, *J. Combin. Theory Ser. A*, **105** (2004), 63–77.

CNRS PRiSM, Université de Versailles, 45 Avenue des Etats-Unis, 78035 Versailles Cedex France
E-mail address: syl@prism.uvsq.fr

LIAFA, Université Denis Diderot, 2 place Jussieu, Case 7014, F-75251 Paris Cedex 05
E-mail address: mallet@liafa.jussieu.fr

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Macdonald polynomials at roots of unity

Francois Descouens and Hideaki Morita

Abstract. The aim of this note is to give some factorisation formulas for different versions of the Macdonald polynomials when the parameter t is specialized at roots of unity, generalizing those given in [**LLT1**] for Hall-Littlewood functions.

Résumé. Le but de cette note est de donner quelques formules de factorisations pour différentes versions des polynômes de Macdonald lorsque le paramètre t est spécialisé aux racines de l'unité. Ces formules généralisent celles données dans [**LLT1**] pour les fonctions de Hall-Littlewood.

1. Introduction

In [**LLT1**], Lascoux, Leclerc and Thibon give some factorisation formulas for the specialization of the parameter q at roots of unity for Hall-Littlewood functions. They also give a corollary of these formulas in terms of cyclic characters of the symmetric group. In this note, we give a generalization of these specializations for different versions of the Macdonald polynomials. We obtain similar formulas in terms of plethystic substitutions and cyclic characters. We also give in the last section a congruence for (q,t)-Kostka polynomials indexed by rectangles using Schur functions in the alphabet constituted by the powers of the parameter t. We will mainly follow the notations of [M].

Acknowlegdements: This note reports on work in progress in collaboration with Jean-Yves Thibon (Université de Marne-la-Vallée). All the computations about Macdonald polynomials have been done using the `MuPAD` package `MuPAD-Combinat` (see [**HT**] for more details on the project or the website *http://mupad-combinat.sourceforge.net/*).

2. Preliminaries

For a partition $\lambda = (\lambda_1, \ldots, \lambda_n)$, we write $l(\lambda)$ its length, $|\lambda|$ its weight, $m_i(\lambda)$ the multiplicity of the part of length i and $\lambda^{'}$ its conjugate partition. Let q and t be two indeterminates and $F = \mathbb{Q}(q,t)$. Let Λ_F be the ring of symmetric functions over the field F. Let us denote by $\langle \cdot, \cdot \rangle_{q,t}$ the inner product on Λ_F defined on the power sums products by

$$\langle \ p_\lambda(x) \ , \ p_\mu(x) \ \rangle_{q,t} = \delta_{\lambda\mu} z_\lambda(q,t),$$

where

$$z_\lambda(q,t) = \prod_{i\geq 1} (m_i)! \ i^{m_i(\lambda)} \ \prod_{i=1}^{l(\lambda)} \frac{1-q^{\lambda_i}}{1-t^{\lambda_i}}.$$

The special case $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\substack{q=0\\t=0}}$ is the usual inner product.

Key words and phrases. Macdonald polynomials, Symmetric functions, Plethysm.

Let $\{P_\lambda(x;q,t)\}_\lambda$ be the family of Macdonald polynomials obtained by orthogonalization of the Schur basis with respect to the inner product $\langle\cdot,\cdot\rangle_{q,t}$. Define a normalization of these functions by

$$Q_\lambda(x;q,t) = \frac{1}{\langle P_\lambda(x;q,t), P_\lambda(x;q,t)\rangle_{q,t}} P_\lambda(x;q,t).$$

It is clear from the previous definitions that the families $\{P_\lambda\}_\lambda$ and $\{Q_\mu\}_\mu$ are dual to each other with respect to the inner product $\langle\cdot,\cdot\rangle_{q,t}$ (c.f., [M, I, Section 4] and [M, VI, (2.7)]).

Proposition 2.1. *Let $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ be two sets of variables. The Macdonald polynomials $\{P_\lambda\}_\lambda$ and $\{Q_\lambda\}_\lambda$ satisfy the following Cauchy formula*

$$\sum_\lambda P_\lambda(x;q,t)Q_\lambda(y;q,t) = \prod_{i,j} \frac{(tx_iy_j;q)_\infty}{(x_iy_j;q)_\infty},$$

where $(a;q)_\infty$ is defined to be the infinite product $\prod_{r\geq 0}(1-aq^r)$.

Let $f(x) \in \Lambda_F$ be a symmetric function in the variables $x = (x_1, x_2, \dots)$. We consider the following algebra homomorphism

$$\begin{aligned} \tilde{}: \Lambda_F &\longrightarrow \Lambda_F \\ f(x) &\longmapsto \tilde{f}(x) = f\left(\frac{x}{1-t}\right). \end{aligned}$$

The images of the powersum functions $(p_k)_k \geq 1$ by this morphism are

$$\forall k \geq 1, \quad \tilde{p_k}(x) = \frac{1}{1-t^k} p_k(x).$$

We also define the algebra morphism

$$\begin{aligned} {}': \Lambda_F &\longrightarrow \Lambda_F \\ f(x) &\longmapsto f'(x) = f\left(\frac{1-q}{1-t}x\right). \end{aligned}$$

The images of the powersum functions are

$$\forall k \geq 1, \quad p_k^{'}(x) = \frac{1-q^k}{1-t^k} p_k(x).$$

Let us consider the following modified version of the Macdonal polynomial

$$Q'_\mu(x;q,t) = Q_\mu\left(\frac{1-q}{1-t}x;q,t\right).$$

We can see that the set $\{Q'_\mu\}_\mu$ is the dual basis of $\{P_\lambda\}_\lambda$ with respect to the usual inner product.

Proposition 2.2. *Let $x = (x_1, x_2, \dots)$ and $y = (y_1, y_2, \dots)$ be two sets of variables. The Macdonald polynomials $\{P_\lambda\}_\lambda$ and $\{Q_\lambda^{'}\}_\lambda$ satisfy the following Cauchy formula*

$$\sum_\lambda P_\lambda(x;q,t)Q'_\lambda(y;q,t) = \prod_{i,j} \frac{1}{1-x_iy_j}.$$

Let $J_\mu(x;q,t)$ be the symmetric function with two parameters defined by

$$J_\mu(x;q,t) = c_\mu(q,t)P_\mu(x;q,t) = c'_\mu(q,t)Q_\mu(x;q,t), \tag{2.1}$$

where

$$c_\mu(q,t) = \prod_{s\in\mu}(1-q^{a(s)}t^{l(s)+1}) \quad \text{and} \quad c'_\mu(q,t) = \prod_{s\in\mu}(1-q^{a(s)+1}t^{l(s)}).$$

The symmetric function $J_\mu(x;q,t)$ is called the integral form of $P_\mu(x;q,t)$ or $Q_\mu(x;q,t)$ [M, VI, Section 8]. Using this integral form, we can define an other modified version of the Macdonald polynomial and the (q,t)-Kostka polynomials $K_{\lambda,\mu}(q,t)$ by

$$\tilde{J}_\mu(x;q,t) = J_\mu\left(\frac{x}{1-t};q,t\right) = \sum_\lambda K_{\lambda,\mu}(q,t)s_\lambda.$$

Haiman, Haglund and Loehr consider a modified version of $\tilde{J}_\mu(x;q,t)$ and introduce others (q,t)-Kostka polynomials $\tilde{K}_{\lambda,\mu}(q,t)$ by

$$\tilde{H}_\mu(x;q,t) = t^{n(\mu)} \tilde{J}_\mu(x;q,t^{-1}) = \sum_\lambda \tilde{K}_{\lambda,\mu}(q,t) s_\lambda$$

where $n(\mu) = \sum_i (i-1)\mu_i$. In [**HHL**], they give a combinatorial interpretation of this modified version expanded on monomials by introducing a notion of major index and inversion on arbitrary fillings of μ by integers.

Remark 2.1. *Let μ and ρ be partitions of the same weight. We have*

$$X_\rho^\mu(q,t) = \langle\ \tilde{J}_\mu(q,t)\ ,\ p_\rho(x)\ \rangle,$$

where $X_\rho^\mu(q,t)$ is the Green polynomial with two variables, defined by

$$X_\rho^\mu(q,t) = \sum_\lambda \chi_\rho^\lambda K_{\lambda\mu}(q,t).$$

Here χ_ρ^λ is the value of the irreducible character of the symmetric group corresponding to the partition λ on the conjugancy class indexed by ρ.

3. Plethystic formula

We recall the definitions of some combinatorial quantities associated to a cell $s = (i,j)$ of a given partition. The arm length $a(s)$, arm-colength $a^{'}(s)$, leg length $l(s)$ and leg-colength $l^{'}(s)$ are respectively the number of cells at the east, at the west, at the south and at the north of the cell s (cf [M, VI, (6.14)])

$$a(s) = \lambda_i - j \quad , \quad a^{'}(s) = j-1,$$
$$l(s) = \lambda_j^{'} - i \quad , \quad l^{'}(s) = i-1.$$

We call plethysm of a symmetric function g by a powersum p_n, the following operation

$$p_n \circ g = g(x_1^n, x_2^n, \ldots).$$

As the powersums generate Λ_F, the operation $f \circ g$ is naturaly defined for any symmetric functions f and g (see [M, I, 8] for more details).

In this section, we shall show a plethystic formula for Macdonald polynomials when the second parameter t is specialized at primitive roots of unity.

Proposition 3.1. ([M,VI, (6.11')]) *Let l be a positive integer and λ a partition such that $l(\lambda) \leq l$. The Macdonald polynomials $P_\lambda(x;t,q)$ on the alphabet $x_i = t^i$ for $0 \leq i \leq l-1$ and $x_i = 0$ for all $i \geq l$ can be written*

$$P_\lambda(1,t,\ldots,t^{l-1};q,t) = t^{n(\lambda)} \prod_{s\in\lambda} \frac{1-q^{a^{'}(s)}t^{l-l^{'}(s)}}{1-q^{a(s)}t^{l(s)+1}}. \tag{3.1}$$

Proposition 3.2. *Let l be a positive integer and λ a partition such that $l(\lambda) \leq l$. For ζ a primitive l-th root of unity, the Macdonald polynomial P_λ satisfy the following specialization*

$$P_\lambda(1,\zeta,\zeta^2,\ldots,\zeta^{l-1};q,\zeta) = \begin{cases} (-1)^{(l-1)r} & \textit{if } \lambda = (r^l) \textit{ for some } r \geq 0, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Supplying zeros at the end of λ, we consider the partition λ as a sequence of length exactly equal to l, i.e., $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_l)$ for $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l \geq 0$. The multiplicity of 0 in λ is $m_0 = l - l(\lambda)$. We will denote by $\varphi_r(t)$ the polynomial

$$\varphi_r(t) = (1-t)(1-t^2)\ldots(1-t^r).$$

Let

$$f(t) = \frac{(1-t^l)(1-t^{l-1})\cdots(1-t^{l-l(\lambda)})(1-t^{l-l(\lambda)-1})\cdots(1-t^2)(1-t)}{\varphi_{m_0}(t)\varphi_{m_1}(t)\varphi_{m_2}(t)\cdots\cdots}$$

be the product of factors of the form $1-q^0t^\alpha$ for some $\alpha > 0$ in the formula (3.1). If we suppose that $f(\zeta) \neq 0$, the factor $1-t^l$ should be contained in one of $\varphi_{m_i}(t)$. This means that there exists $i \geq 0$ such that $m_i \geq l$. Since we consider λ as a sequence of length exactly l, this implies the condition $m_r = l$ for some $r \geq 0$. Thus, if $P_\lambda(1, \zeta, \zeta^2, \ldots, \zeta^{l-1}; q, \zeta) \neq 0$, the shape of λ should be (r^l).

Suppose now that $\lambda = (r^l)$. By Proposition 3.1, it follows that

$$\begin{aligned} P_\lambda(1, \zeta, \zeta^2, \ldots, \zeta^{l-1}; q, \zeta) &= \zeta^{n(\lambda)} \prod_{s \in \lambda} \frac{1 - q^{a'(s)} \zeta^{l-l'(s)}}{1 - q^{a(s)} \zeta^{1+l(s)}} \\ &= \zeta^{n(\lambda)} \prod_{(i,j) \in \lambda} \frac{1 - q^{j-1} \zeta^{l-(i-1)}}{1 - q^{r-j} \zeta^{l-i+1}} \\ &= \zeta^{n(\lambda)} \prod_{i=1}^{l} \prod_{j=1}^{r} \frac{1 - q^{j-1} \zeta^{l-i+1}}{1 - q^{r-j} \zeta^{l-i+1}} \end{aligned}$$

For each i, it is easy to see that

$$\prod_{j=1}^{r} \frac{1 - q^{j-1} \zeta^{l-i+1}}{1 - q^{r-j} \zeta^{l-i+1}} = 1.$$

Hence, we obtain

$$P_\lambda(1, \zeta, \zeta^2, \ldots, \zeta^{l-1}; q, \zeta) = \zeta^{n(\lambda)},$$

and it follows immediately from the definition of $n(\lambda)$ that

$$\zeta^{n(\lambda)} = \zeta^{l(l-1))r/2} = (-1)^{(l-1)r}.$$

□

Theorem 3.1. *Let l and r be two positive integers and ζ a primitive l-th root of unity. The Macdonald polynomials $Q'_{(r^l)}(x; q, t)$ satisfy the following specialization formula at $t = \zeta$*

$$Q'_{(r^l)}(x; q, \zeta) \quad = \quad (-1)^{(l-1)r} (p_l \circ h_r)(x).$$

Proof. Recall that

$$\sum_\lambda P_\lambda(x; q, t) Q'_\lambda(y; q, t) = \prod_{i,j} \frac{1}{1 - x_i y_j}.$$

If we let $x_i = \zeta^{i-1}$ for $i = 1, 2, \ldots, l$ and $x_i = 0$ for $i > l$ and $t = \zeta$, we obtain

$$\sum_\lambda P_\lambda(1, \zeta, \zeta^2, \ldots, \zeta^{l-1}; q, \zeta) Q'_\lambda(y; q, \zeta) = \prod_{j \geq 1} \prod_{i=1}^{l} \frac{1}{1 - \zeta^{i-1} y_j}. \tag{3.2}$$

By Proposition 3.2, the left hand side of (3.2) is equal to

$$\sum_{r \geq 0} (-1)^{(r-1)l} Q'_{(r^l)}(y; q, \zeta).$$

Since $\prod_{i=1}^{l}(1 - \zeta^{i-1} t) = 1 - t^l$, the right hand side of (3.2) coincides with

$$\sum_{r \geq 0} h_r(y^l),$$

where y^l denotes the set of variables $(y_1^l, y_2^l, \cdots)$. Comparing the degrees, we can conclude that

$$Q'_{(r^l)}(y; q, \zeta) = (-1)^{(l-1)r} h_r(y^l) = (-1)^{(l-1)r} (p_l \circ h_r)(y).$$

□

Example 3.2. *For $\lambda = (222)$ and $l = 3$, we can compute*

$$\begin{aligned} Q'_{(222)}(x; q, e^{\frac{2i\pi}{3}}) &= -s_{321} + s_{33} + s_{411} - s_{51} + s_6 + s_{222} \\ &= p_3 \circ h_2(x). \end{aligned}$$

In order to give similar formula for the modified versions of the integral form of the Macdonald polynomials, we give a formula for the specialization of the constant $c'_{(r^l)}(t, q)$ at t a primitive l-th root of unity.

LEMMA 3.3. *Let l and r be two positive integers and ζ a l-th primitive root of unity. The normalization constant satisfies the following specialization at $t = \zeta$*

$$c'_{(r^l)}(q,\zeta) = \prod_{i=1}^{r}(1-q^{il}).$$

PROOF. Recall the definition of the normalization constant

$$c'_{(r^l)}(q,t) = \prod_{s\in\mu}(1-q^{a(s)+1}t^{l(s)}) = \prod_{i=1}^{r}\prod_{j=1}^{l}(1-q^{r-i+1}t^{j}) = \prod_{i=1}^{r}\prod_{j=1}^{l}(1-q^{i}t^{j}).$$

Specializing t at ζ a l-th primitive root of unity, we obtain

$$c'_{(r^l)}(q,\zeta) = \prod_{i=1}^{r}\prod_{j=1}^{l}(1-q^{i}\zeta^{j}) = \prod_{i=1}^{r}(1-q^{il}).$$

□

COROLLARY 3.1. *With the same notation as in Theorem 3.1, the modified integral form of the Macdonald polynomials $\tilde{J}_\mu(x;q,t)$ satisfy a similar formula at $t=\zeta$, a primitive l-th root of unity*

$$\tilde{J}_{(r^l)}(x;q,\zeta) \quad = \quad (-1)^{(l-1)r}\prod_{i=1}^{r}(1-q^{il})\ p_l \circ h_r\left(\frac{x}{1-q}\right).$$

PROOF. Using the definition (2.1) of the integral form of the Macdonald polynomials

$$\begin{aligned}\tilde{J}_{(r^l)}(x;q,t) = J_{(r^l)}\left(\frac{x}{1-t};q,t\right) &= c'_{(r^l)}(q,t)Q'_{(r^l)}\left(\frac{x}{1-q};q,t\right)\\ &= c'_{(r^l)}(q,t)\left(Q'_{(r^l)}(\ .\ ;q,t)\circ\frac{1}{1-q}p_1\right)(x).\end{aligned}$$

By specializing the previous egality at a primitive l-th root of unity ζ, we obtain with Theorem 3.1 and the associativity of the plethysm

$$\begin{aligned}\tilde{J}_{(r^l)}(x;q,\zeta) &= c'_{(r^l)}(q,\zeta)(-1)^{r(l-1)}\left((p_l\circ h_r)\circ\frac{1}{1-q}p_1\right)(x)\\ &= c'_{(r^l)}(q,\zeta)(-1)^{r(l-1)}p_l\circ h_r\left(\frac{x}{1-q}\right).\end{aligned}$$

Using the formula of Lemma 3.3, we obtain the formula. □

COROLLARY 3.2. *With the same notations than in Theorem 3.1, the modified Macdonald polynomials $\tilde{H}_\mu(x;q,t)$ satisfy the following specialization at $t=\zeta$, a primitive l-th root of unity*

$$\tilde{H}_{(r^l)}(x;q,\zeta) = \prod_{i=1}^{r}(1-q^{il})\ p_l\circ h_r\left(\frac{x}{1-q}\right).$$

PROOF. The result follows from corollary 3.1 and $\zeta^{n(r^l)} = \zeta^{rl(l-1)/2} = (-1)^{(l-1)r}$. □

EXAMPLE 3.4. *For $\lambda = (222)$ and $l = 3$, we can compute*

$$\begin{aligned}\tilde{J}_{(222)}(x;q,e^{\frac{2i\pi}{3}}) &= q^3(s_{111111}-s_{21111}+s_{3111})-(q^3+1)(s_{321}-s_{33}-s_{222})+s_{411}-s_{51}+s_6\\ &= (1-q^3)(1-q^6)\ p_3\circ h_2\left(\frac{x}{1-q}\right).\end{aligned}$$

REMARK 3.5. *As the Madonald polynomials $P_{(r^l)}(x;q,t)$ indexed by rectangles satisfy the following specialization at $t=\zeta$, a primitive l-th root of unity,*

$$\left.\frac{1}{\langle P_{(r^l)}(x;q,t)\ ,\ P_{(r^l)}(x;q,t)\rangle_{q,t}}\right|_{t=\zeta} = 0$$

we obtain the following specializations

$$Q_{(r^l)}(x;q,\zeta)=0 \quad \textit{and} \quad J_{(r^l)}(x;q,\zeta)=0.$$

4. Pieri formula at roots of unity

In order to prove the factorization formulas, we prepare an auxiliary result (Proposition 4.1) on the coefficients of Pieri formula at root of unity (cf [M, VI, (6.24 ii)])

$$Q^{'}_{\mu}(x;q,t)g^{'}_{r}(x;q,t)=\sum_{\lambda}\psi_{\lambda/\mu}(q,t)\ Q^{'}_{\lambda}(x;q,t). \tag{4.1}$$

Let λ and μ be partitions such that λ/μ is a horizontal (r-)strip θ. Let $C_{\lambda/\mu}$ (resp. $R_{\lambda/\mu}$) be the union of columns (resp. rows) of λ that intersects with θ, and $D_{\lambda/\mu}=C_{\lambda/\mu}-R_{\lambda/\mu}$ the set theoretical difference. Then it can be seen from the definition that for each cell s of $D_{\lambda/\mu}$ (resp. $D_{\tilde{\lambda}/\tilde{\mu}}$) there exists a unique connected component of θ (resp. $\tilde{\theta}$), which lies in the same row as s. We denote the corresponding component by θ_s (resp. $\tilde{\theta}_s$).

Suppose that l and r are positive integers. Set $\tilde{\lambda}=\lambda\cup(r^l)$ and $\tilde{\mu}=\mu\cup(r^l)$. We shall consider the difference between $D_{\tilde{\lambda}/\tilde{\mu}}$ and $D_{\lambda/\mu}$. It can be seen that there exists a projection

$$p=p_{\lambda/\mu}:D_{\tilde{\lambda}/\tilde{\mu}}\longrightarrow D_{\lambda/\mu}.$$

The cardinality of the fiber of each cell $s=(i,j)\in D_{\lambda/\mu}$ is exactly one or two. Let J_s denote the set of second coordinates of the cells in θ_s. If all elements of J_s are all strictly larger than r, the fiber $p^{-1}(s)$ consists of a single element $s=(i,j)$. If all elements of J_s are strictly smaller than r, then the fiber $p^{-1}(s)$ consists of a single element $\tilde{s}:=(i,j+l)$. In the case where J_s contains r, then the fiber $p^{-1}(s)$ consists of exactly two elements $s=(i,j)$ and $\tilde{s}=(i,j+l)$. For the case where $r\in J_s$, we have the followig lemma, which follows immediately from the definition of the projection $p=p_{\lambda/\mu}$.

Lemma 4.1. *Let $s=(i,j)$ be a cell of $D_{\lambda/\mu}$ and $\tilde{s}=(i,j+l)$ be a cell of $D_{\tilde{\lambda}/\tilde{\mu}}$ such that $r\in J_s$. The arm length, the arm-colength, the leg length and the leg-colength satisfy the following properties* :

(1) $a_{\tilde{\mu}}(s)=a_{\tilde{\lambda}}(\tilde{s})$,
(2) $l_{\tilde{\mu}}(s)-l_{\tilde{\lambda}}(\tilde{s})=l$,
(3) $a_{\tilde{\mu}}(\tilde{s})=a_{\mu}(s)$,
(4) $l_{\tilde{\mu}}(\tilde{s})=l_{\mu}(s)$,
(5) $a_{\tilde{\lambda}}(s)=a_{\lambda}(s)$,
(6) $l_{\tilde{\lambda}}(s)-l_{\lambda}(s)=l$.

Proposition 4.1. *Let λ and μ be two partitions such that $\mu\subset\lambda$ and $\theta=\lambda-\mu$ a horizontal strip. Let r and l be positive integers and ζ a primitive root of unity. Then it follows that*

$$\psi_{\lambda\cup(r^l)/\mu\cup(r^l)}(q,\zeta)=\psi_{\lambda/\mu}(q,\zeta).$$

Proof. Recall that for a cell s of the partition ν,

$$\psi_{\lambda/\mu}(q,t)=\prod_{s\in D_{\lambda/\mu}}\frac{b_{\mu}(s)}{b_{\lambda}(s)},$$

where

$$b_{\nu}(s)=\frac{1-q^{a_{\nu}(s)}t^{l_{\nu}(s)+1}}{1-q^{a_{\nu}(s)+1}t^{l_{\nu}(s)}}$$

If $s=(i,j)\in D_{\lambda/\mu}$ satisfies the condition $r>j$ for all $j\in J_s$, then the fiber $p^{-1}(s)$ of the projection p is $\{\tilde{s}=(i,j+l)\}$, and we have $a_{\mu}(s)=a_{\tilde{\mu}}(s),a_{\lambda}(s)=a_{\tilde{\lambda}}(s)$ and $l_{\mu}(s)+l=l_{\tilde{\mu}}(s),l_{\lambda}(s)+l=l_{\tilde{\lambda}}(s)$. It is clear from these identities that $b_{\mu}(s)/b_{\lambda}(s)=b_{\tilde{\mu}}(s)/b_{\tilde{\lambda}}(s)$ at $t=\zeta$ in this case. Suppose that s satisfies $j>r$ for all $j\in J_s$. In this case, the fiber $p^{-1}(s)$ consisits of a single element $\{s=(i,j)\}$, and we have $a_{\mu}(s)=a_{\tilde{\mu}}(s)$

and $a_\lambda(s) = a_{\tilde{\lambda}}(s)$ and $l_\mu(s) = l_{\tilde{\mu}}(s)$ and $l_\lambda(s) = l_{\tilde{\lambda}}(s)$. Hence we have $b_\mu(s)/b_\lambda(s) = b_{\tilde{\mu}}(s)/b_{\tilde{\lambda}}(s)$. Consider the case where $r \in J_s$. In this case, the fiber $p^{-1}(s)$ consists of two elements $\{s, \tilde{s}\}$. Let consider

$$\prod_{u \in p^{-1}(s)} \frac{b_{\tilde{\mu}}(u)}{b_{\tilde{\lambda}}(u)} = \frac{1-q^{a_{\tilde{\mu}}(s)}t^{l_{\tilde{\mu}}(s)+1}}{1-q^{a_{\tilde{\mu}}(s)+1}t^{l_{\tilde{\mu}}(s)}} \frac{1-q^{a_{\tilde{\lambda}}(s)+1}t^{l_{\tilde{\lambda}}(s)}}{1-q^{a_{\tilde{\lambda}}(s)}t^{l_{\tilde{\lambda}}(s)+1}} \frac{1-q^{a_{\tilde{\mu}}(\tilde{s})}t^{l_{\tilde{\mu}}(\tilde{s})+1}}{1-q^{a_{\tilde{\mu}}(\tilde{s})+1}t^{l_{\tilde{\mu}}(\tilde{s})}} \frac{1-q^{a_{\tilde{\lambda}}(\tilde{s})+1}t^{l_{\tilde{\lambda}}(\tilde{s})}}{1-q^{a_{\tilde{\lambda}}(\tilde{s})}t^{l_{\tilde{\lambda}}(\tilde{s})+1}}$$

By items (1) and (2) of Lemma 4.1, it follows that

$$\left\{\frac{1-q^{a_{\tilde{\mu}}(s)}t^{l_{\tilde{\mu}}(s)+1}}{1-q^{a_{\tilde{\mu}}(s)+1}t^{l_{\tilde{\mu}}(s)}}\right\}^{-1}\Bigg|_{t=\zeta} = \frac{1-q^{a_{\tilde{\lambda}}(\tilde{s})+1}t^{l_{\tilde{\lambda}}(\tilde{s})}}{1-q^{a_{\tilde{\lambda}}(\tilde{s})}t^{l_{\tilde{\lambda}}(\tilde{s})+1}}\Bigg|_{t=\zeta}.$$

It also follows from item (3) and (4) of Lemma 4.1,

$$\frac{1-q^{a_{\tilde{\mu}}(\tilde{s})}t^{l_{\tilde{\mu}}(\tilde{s})+1}}{1-q^{a_{\tilde{\mu}}(\tilde{s})+1}t^{l_{\tilde{\mu}}(\tilde{s})}}\Bigg|_{t=\zeta} = \frac{1-q^{a_\mu(s)}t^{l_\mu(s)+1}}{1-q^{a_\mu(s)+1}t^{l_\mu(s)}}\Bigg|_{t=\zeta},$$

and from item (5) and (6),

$$\frac{1-q^{a_{\tilde{\lambda}}(s)+1}t^{l_{\tilde{\lambda}}(s)}}{1-q^{a_{\tilde{\lambda}}(s)}t^{l_{\tilde{\lambda}}(s)+1}}\Bigg|_{t=\zeta} = \frac{1-q^{a_\lambda(s)+1}t^{l_\lambda(s)}}{1-q^{a_\lambda(s)}t^{l_\lambda(s)+1}}\Bigg|_{t=\zeta}.$$

Therefore, it follows that

$$\prod_{u \in p^{-1}(s)} \frac{b_{\tilde{\mu}}(u)}{b_{\tilde{\lambda}}(u)} = \frac{b_\mu(s)}{b_\lambda(s)}.$$

Combining these, the assertion follows. □

5. Factorization formulas

In this section, we shall show factorization formulas for different kinds of Macdonald polynomials at roots of unity.

THEOREM 5.1. *Let l be a positive integer and ζ a primitive l-th root of unity. Let $\mu = (1^{m_1}2^{m_2}\cdots n^{m_n})$ be a partition of a positive integer n. For each i, let $m_i = lq_i + r_i$ with $0 \le r_i \le l-1$ and let $\bar{\mu} = (1^{r_1}2^{r_2}\cdots m^{r_n})$. The function $Q^{'}_\mu$ satisfy the following factorisation formula at $t = \zeta$*

$$Q^{'}_\mu(x;q,\zeta) \quad = \quad \left(Q^{'}_{(1^l)}(x;q,\zeta)\right)^{q_1} \left(Q^{'}_{(2^l)}(x;q,\zeta)\right)^{q_2} \cdots \left(Q^{'}_{(n^l)}(x;q,\zeta)\right)^{q_n} Q^{'}_{\bar{\mu}}(x;q,\zeta).$$

PROOF. We shall show that the $\mathbb{C}$-linear map defined by

$$\begin{array}{rcl} f_r : \Lambda_F & \longrightarrow & \Lambda_F \\ Q^{'}_\mu(x;q,\zeta) & \longmapsto & Q^{'}_{\mu\cup(r^l)}(x;q,\zeta) \end{array}$$

is an $\Lambda_{\mathbb{C}(q)}$-linear map. Let ζ be a primitive l-th root of unity. From (3.1), we have

$$Q^{'}_\mu(x;q,\zeta)g^{'}_k(x;q,\zeta) = \sum_\lambda \psi_{\lambda/\mu}(q,\zeta)Q^{'}_\lambda(x;q,\zeta)$$

where the sum is taken over the partitions λ such that $\lambda - \mu$ is a horizontal (k-)strip. Using the result of Proposition 4.1, it follows that

$$\begin{array}{rcl} Q^{'}_{\mu\cup(r^l)}(x;q,\zeta)g^{'}_k(x;q,\zeta) & = & \sum_\lambda \psi_{\lambda\cup(r^l)/\mu\cup(r^l)}(q,\zeta)Q^{'}_{\lambda\cup(r^l)}(x;q,\zeta) \\ & = & \sum_\lambda \psi_{\lambda/\mu}(q,\zeta)Q^{'}_{\lambda\cup(r^l)}(x;q,\zeta). \end{array}$$

Consequently, for each $r \ge 1$, the multiplication by g_k commutes with the morphism f_r. Since $\{g_k(x;q,\zeta_l)\}_{k\ge1}$ generate the algebra $\Lambda_{\mathbb{C}(q)}$ [M, VI, (2.12)], the map f_r is $\Lambda_{\mathbf{C}(q)}$-linear. This implies that

$$\begin{array}{rcl} \forall F \in \Lambda_{\mathbb{C}_q},\ f_r(F(x)) & = & F(x)f_r(1) \\ & = & F(x)Q^{'}_{(r^l)}(x;q,\zeta). \end{array}$$

□

Corollary 5.1. *With the same notation as in Theorem 3.1, we have*

$$\tilde{J}_\mu(x;q,\zeta) \quad = \quad \left(\tilde{J}_{(1^l)}(x;q,\zeta)\right)^{q_1} \left(\tilde{J}_{(2^l)}(x;q,\zeta)\right)^{q_2} \cdots \left(\tilde{J}_{(n^l)}(x;q,\zeta)\right)^{q_n} \tilde{J}_{\bar{\mu}}(x;q,\zeta).$$

Proof. If we define

$$\Psi_{\lambda/\mu}(q,t) := \psi_{\lambda/\mu}(q,t)\frac{c'_\mu(q,t)}{c'_\lambda(q,t)},$$

then the Pieri formula for the integral form $\tilde{J}_\mu(x;q,t)$ is written as follows

$$\tilde{J}_\mu(x;q,t)\tilde{g_k}(x;q,t) = \sum_\lambda \Psi_{\lambda/\mu}(q,t)\tilde{J}_\lambda(x;q,t),$$

where the sum is over the partitions λ such that $\lambda - \mu$ is a horizontal (k-)strip.
Let a positive integer r be arbitrarily fixed, and $\tilde{\nu}$ denote the partition $\nu \cup (r^l)$. Since we have already shown that $\psi_{\tilde{\lambda}/\tilde{\mu}}(q,\zeta) = \psi_{\lambda/\mu}(q,\zeta)$, it suffices to show that

$$\frac{c'_{\tilde{\mu}}(q,\zeta)}{c'_{\tilde{\lambda}}(q,\zeta)} = \frac{c'_\mu(q,\zeta)}{c'_\lambda(q,\zeta)}.$$

We shall actually show that

$$\frac{c'_{\tilde{\mu}}(q,\zeta)}{c'_\mu(q,\zeta)} = \frac{c'_{\tilde{\lambda}}(q,\zeta)}{c'_\lambda(q,\zeta)}.$$

It follows from the definition that

$$\begin{aligned} \frac{c'_{\tilde{\mu}}(q,t)}{c'_\mu(q,t)} &= \frac{\prod_{s\in\tilde{\mu}}(1-q^{a_{\tilde{\mu}}(s)+1}t^{l_{\tilde{\mu}}(s)})}{\prod_{s\in\mu}(1-q^{a_{\tilde{\mu}}(s)+1}t^{l_{\tilde{\mu}}(s)})} \\ &= \frac{\prod_{\substack{s\in\tilde{\mu}\\ s\notin(r^l)}}(1-q^{a_{\tilde{\mu}}(s)+1}t^{l_{\tilde{\mu}}(s)})}{\prod_{s\in\mu}(1-q^{a_{\tilde{\mu}}(s)+1}t^{l_{\tilde{\mu}}(s)})} \prod_{s\in(r^l)\subset\tilde{\mu}}(1-q^{a_{\tilde{\mu}}(s)+1}t^{l_{\tilde{\mu}}(s)}). \end{aligned}$$

The Young diagram of the partition $\tilde{\mu}$ is the disjoint union of the cells $\{\tilde{s} \in \tilde{\mu} | s \in \mu\}$ and (r^l). For each $s \in \mu$, we have as seen in previous Theorem that $a_{\tilde{\mu}}(\tilde{s}) = a_\mu(s)$, and $l_{\tilde{\mu}}(\tilde{s}) = l_\mu(s)$ or $l_\mu(s)+l$. Hence at $t = \zeta$, we have

$$\frac{c'_\mu(q,\zeta)}{c'_{\tilde{\mu}}(q,\zeta)} = \prod_{s\in(r^l)\subset\tilde{\mu}}(1-q^{a_{\tilde{\mu}}(s)+1}\zeta^{l_{\tilde{\mu}}(s)}). \tag{3.1}$$

$$\frac{c'_\lambda(q,\zeta)}{c'_{\tilde{\lambda}}(q,\zeta)} = \prod_{s\in(r^l)\subset\tilde{\lambda}}(1-q^{a_{\tilde{\lambda}}(s)+1}\zeta^{l_{\tilde{\lambda}}(s)}). \tag{3.2}$$

Although there is a difference between the positions where the block (r^l) is inserted in the Young diagram of μ and λ, (3.1) and (3.2) coincide at $t = \zeta$, since $a_{\tilde{\mu}}(s) = a_{\tilde{\lambda}}(s)$ for each $s \in (r^l)$. Thus we have

$$\frac{c'_\mu(q,\zeta)}{c'_{\tilde{\mu}}(q,\zeta)} = \frac{c'_\lambda(q,\zeta)}{c'_{\tilde{\lambda}}(q,\zeta)}.$$

□

Let $\nu = (\nu_1, \ldots, \nu_p)$ be a partition. For some $l \geq 0$, we denote by ν^l the partition where each part of ν is repeated l times. We can give a more explicit expression for the factorisation formula in the special case where $\mu = \nu^l$.

Corollary 5.2. *Let ν be a partition and l a positive integer. We have the following special cases for the factorisation formulas*

$$Q'_{\nu^l}(X;q,\zeta) \quad = \quad (-1)^{(l-1)|\nu|} p_l \circ h_\nu(x), \tag{5.1}$$

$$\tilde{J}_{\nu^l}(X;q,\zeta) \quad = \quad (-1)^{(l-1)|\nu|} \prod_{j=1}^{l(\nu)} \prod_{i=1}^{\nu_j} (1-q^{il})\ p_l \circ h_\nu\left(\frac{x}{1-q}\right). \tag{5.2}$$

EXAMPLE 5.2. *For $\lambda = (222111)$ and $k = 3$, we can compute*

$$\begin{aligned} Q'_{222111}(x;q,e^{2i\pi/3}) &= -s_{22221} - s_{321111} + s_{3222} + s_{33111} - s_{3321} + 3s_{333} + s_{411111} \\ &\quad -2s_{432} + 2s_{441} - s_{51111} + 2s_{522} - 2s_{54} + s_{6111} - 2s_{621} + 2s_{63} \\ &\quad +s_{711} - s_{81} + s_9 + s_{222111} \\ &= p_3 \circ h_{21}(x). \end{aligned}$$

6. A generalization of the plethystic formula

In this section, using the factorisation formula given in Theorem 5.1, we shall give a generalization of the plethystic formula obtained by specializing Macdonald polynomials at roots of unity in Theorem 3.1. For λ a partition, let consider the following map which is the plethystic substitution by the powersum p_λ

$$\begin{aligned} \Psi_\lambda : \Lambda_F &\longrightarrow \Lambda_F \\ f &\longmapsto p_\lambda \circ f \end{aligned}$$

LEMMA 6.1. *Let λ and μ be two partitions, the maps Ψ_λ and Ψ_μ satisfy the following multiplicative property*

$$\Psi_\lambda(f)\,\Psi_\mu(f) = \Psi_{\lambda\cup\mu}(f).$$

PROPOSITION 6.1. *Let d be an integer such that $d|l$ and ζ_d be a primitive d-th root of unity,*

$$Q'_{(r^l)}(x;q,\zeta_d) = (-1)^{\frac{rl(d-1)}{d}} p_d^{l/d} \circ h_r(x).$$

PROOF. Let d and l be two integers such that d divide l. Let $\mu = (r^l)$ the rectangle partition with parts of length r. Using the factorisation formula described in Theorem 5.1, we can write

$$Q'_{(r^l)}(x;q,\zeta_d) = \left(Q'_{(r^d)}(x;q,\zeta_d)\right)^{l/d}. \tag{6.1}$$

With the specialization formula at root of unity written in Theorem 3.1, we have

$$\begin{aligned} \left(Q'_{(r^d)}(x;q,\zeta_d)\right)^{l/d} &= \left((-1)^{(d-1)r} p_d \circ h_r(x)\right)^{l/d} \\ &= (-1)^{\frac{lr(d-1)}{d}} \left(p_d \circ h_r(x)\right)^{l/d} \end{aligned}$$

Using the Lemma 6.1, we obtain

$$\left(Q'_{(r^d)}(x;q,\zeta_d)\right)^{l/d} = (-1)^{\frac{lr(d-1)}{d}} p_d^{l/d} \circ h_r(x)$$

Finally, we obtain by (6.1)

$$Q'_{(r^l)}(x;q,\zeta_d) = (-1)^{\frac{lr(d-1)}{d}} p_d^{l/d} \circ h_r(x)$$

□

Using the same proof, we can write a similar specialization for integral forms of the Macdonald Polynomials.

COROLLARY 6.1. *The Macdonald polynomials $\tilde{J}_\mu(x;q,t)$ satisfy the same generalization than the $Q'_\mu(x;q,t)$*

$$\tilde{J}_{(r^l)}(x;q,\zeta_d) = (-1)^{\frac{rl(d-1)}{d}} \left(\prod_{i=1}^{r}(1-q^{il})\right)^{l/d} p_d^{l/d} \circ h_r\left(\frac{x}{1-q}\right).$$

Example 6.2. *For $\lambda = (222222)$ (i.e $r = 2$ and $l = 6$) and $d = 3$ we can compute*

$$\begin{aligned} Q^{'}_{(222222)}(x; q, e^{2i\pi/3}) &= -s_{322221} + s_{33222} + 2s_{333111} - 2s_{33321} + 2s_{3333} + s_{422211} - 2s_{432111} \\ &\quad + s_{43221} + 2s_{441111} - s_{4422} + 4s_{444} + s_{522111} - 2s_{52221} + s_{53211} - 2s_{54111} \\ &\quad + s_{5421} - 4s_{543} + 3s_{552} - s_{621111} + 2s_{6222} + s_{63111} - 2s_{6321} + 4s_{633} \\ &\quad + s_{6411} - 3s_{651} + 3s_{66} + s_{711111} - 2s_{732} + 2s_{741} - s_{81111} + 2s_{822} - 2s_{84} \\ &\quad + s_{9111} - 2s_{921} + 2s_{93} + s_{\underline{10}11} - s_{\underline{11}1} + s_{\underline{12}} + s_{222222} \\ &= p_3^2 \circ h_2(x) = p_{(33)} \circ h_2(x). \end{aligned}$$

7. Macdonald polynomials at roots of unity and cyclic characters of the symmetric group

In the following, we will denote the symmetric group of order k by $\mathfrak{S}_k$. Let $\Gamma \subset \mathfrak{S}_k$ be a cyclic subgroup generated by an element of order r. As Γ is a commutative subgroup its irreducible representations are one-dimensional vector space. The corresponding maps $(\gamma_j)_{j=0\ldots r-1}$ can be defined by

$$\begin{aligned} \gamma_j : \quad \Gamma &\longrightarrow \quad GL(\mathbb{C}) \simeq \mathbb{C}^* \\ \tau &\longmapsto \quad \zeta_r^j \end{aligned}$$

where ζ_r is a r-th primitive root of unity (See [**S**] for more details). In [**F**], Foulkes considered the Frobenius characteristic of the representations of $\mathfrak{S}_k$ induced by these irreducible representations and obtained an explicit formula that we will give in the next Proposition.

Let k and n be two positive integers such that $u = (k, d)$ (the greatest common divisor between k and n) and $d = u \cdot m$. Let us define the Ramanujan (or Von Sterneck) sum $c(k, d)$ by

$$c(k, d) = \frac{\mu(m)\phi(d)}{\phi(m)}$$

where μ is the Moebius function and ϕ the Euler totient. The quantity $c(k, d)$ corresponds to the sum of the k-th powers of the primitive d-th roots of unity (the previous expression was first given by Holder, see [**HW**]).

Proposition 7.1. *Let τ be a cyclic permutation of length k and Γ the maximal cyclic subgroup of $\mathfrak{S}_k$ generated by τ. Let j be a positive integer less than k. The Frobenius characteristic of the representation of $\mathfrak{S}_k$ induced by the irreducible representation of Γ, $\gamma_j : \tau \longmapsto \zeta_r^j$, is given by*

$$l_k^{(j)}(x) = \frac{1}{k} \sum_{d|k} c(j, d)\ p_d^{k/d}(x).$$

Example 7.1. For $\mathfrak{S}_6$ and $k = 2$, the corresponding cyclic character $l_6^{(2)}$ can be written

$$\begin{aligned} l_6^{(2)} &= \frac{1}{6}\left(p_{111111} + p_{222} - p_{33} - p_6\right) \\ &= s_{51} + 2s_{42} + s_{411} + 3s_{321} + 2s_{3111} + s_{222} + s_{2211} + s_{21111}. \end{aligned}$$

Theorem 7.2. *Let r and l be two positive integers. The specialization of the Macdonald polynomials $Q^{'}_{(r^l)}(x; q, t)$ at $t = \zeta$, a primitive l-th root of unity, is equivalent to*

$$Q^{'}_{(r^l)}(x; q, t) \mod 1 - t^l \quad = \quad \sum_{j=0}^{l-1} t^j\ (l_l^{(j)} \circ h_r)(x).$$

Proof. We will first give a generalization of the Moebius inversion formula due to E. Cohen (see [**C**] for the original work and [**D**] for a simpler proof). Let

$$P(q) = \sum_{k=0}^{n-1} a_k q^k$$

be a polynomial of degree less than $n - 1$ with coefficients a_k in $\mathbb{Z}$. P is said to be even modulo n if

$$(i, n) = (j, n) \implies a_i = a_j.$$

LEMMA 7.3. *The polynomial P is even modulo n if and only if for every divisor d of n, the residue of P modulo the d-th cyclotomic polynomial Φ_d is a constant r_d in $\mathbb{Z}$. In this case, one has*

$$\begin{aligned} (i)\quad a_k &= \frac{1}{n}\sum_{d|n} c(k,d)\ r_d \\ (ii)\quad r_d &= \sum_{t|n} c(n/d,t)\ a_{n/t}. \end{aligned}$$

Let d be an integer such that $d|l$. By expanding $Q'_{(r^l)}(x;q,t)$ (and more generally $Q'_\lambda(x;q,t)$) on the Schur basis, we can define a kind of (q,t)-Kostka polynomials $K'_{\mu,(r^l)}(q,t)$

$$Q'_{(r^l)}(x;q,t) = \sum_\mu K'_{\mu,(r^l)}(q,t)\ s_\mu(x).$$

Let μ be a partition and d an integer such that $d|l$. $P^q_\mu(t) = \sum_{j=0}^{l-1} a_j(q)t^j$ is the residue modulo $1-t^l$ of the (q,t)-Kostka polynomial $K'_{\mu,(r^l)}(q,t)$ if and only if for all ζ_d primitive d-th root of unity

$$P^q_\mu(\zeta_d) = K'_{\mu,(r^l)}(q,\zeta_d).$$

Using Theorem 5.1 , one has

$$P^q_\mu(\zeta_d) = (-1)^{(d-1)rl/d}\ \langle\ p_d^{l/d}\circ h_r(x)\ ,\ s_\mu(x)\ \rangle.$$

So, $P(\zeta_d)\in\mathbb{Z}$ since the entries of the transition matrix between the powersum to the Schur functions are all integers. Using the Lemma 7.3, we obtain

$$\begin{aligned} a_j(q) &= \frac{1}{l}\sum_{d|l} c(j,d)\ \langle\ p_d^{l/d}\circ h_r(x)\ ,\ s_\mu(x)\ \rangle \\ &= \langle\ l_l^{(j)}\circ h_r(x)\ ,\ s_\mu(x)\ \rangle. \end{aligned}$$

□

COROLLARY 7.1. *For two positive integers r and l, the same residue formulas occurs for the modified Macdonald polynomials $\tilde{J}_{(r^l)}(x;q,t)$ and $J'_{(r^l)}(x;q,t)$*

$$\tilde{J}_{(r^l)}(x;q,t) \mod 1-t^l \quad = \quad \prod_{i=1}^{r}(1-q^{il})\ \sum_{j=0}^{l-1} t^j\ (l_l^{(j)}\circ h_r)\left(\frac{x}{1-q}\right).$$

8. Congruences for (q,t)-Kostka polynomials

For a given partition λ, let denote by $\tilde{s}^{(q)}_\lambda$ the symmetic function defined as follows

$$\tilde{s}^{(q)}_\lambda(x) = s_\lambda\left(\frac{x}{1-q}\right).$$

We also define on the power sums products the internal product between two symmetric functions. For λ and μ two partitions, we have ([M, I, (7.12)])

$$p_\lambda \star p_\mu = \delta_{\lambda,\mu} z_\lambda p_\lambda.$$

PROPOSITION 8.1. *Let r and l be two positive integers and μ a partition of weight nl. Let denote by $\Phi_l(t)$ the cyclotomic polynomial of order l. The (q,t)-Kostka polynomial $\tilde{K}_{\mu,(r^l)}(q,t)$ satisfy the following congruence*

$$\tilde{K}_{\mu,(r^l)}(q,t) = \prod_{i=1}^{r}(1-q^{il})\ \tilde{s}^{(q)}_\mu(1,t,t^2,\ldots,t^{l-1}) \mod \Phi_l(t).$$

And more generally, for all partition ν of weight r

$$\tilde{K}_{\mu,\nu^l}(q,t) = \prod_{j=1}^{l(\nu)}\prod_{i=1}^{\nu_j}(1-q^{il})\ \widetilde{h_{l\nu}\star s_\mu}^{(q)}(1,t,t^2,\ldots,t^{l-1}) \mod \Phi_l(t),$$

where $l\nu$ denote the partition $(l\nu_1,\ldots,l\nu_p)$.

References

[C] E. Cohen, *A class of arithmetical functions*, Proc. Nat. Acad. Sci. USA **41** (1955), 939-944.

[D] J. Désarmenien, *Etude modulo n des statistiques mahoniennes*, Actes du 22eme Séminaire Lotharingien de Combinatoire, IRMA Strasbourg (1990), 27-35.

[F] H.O. Foulkes, *Characters of symmetric groups induced by characters of cyclic subgroups*, (Proc. Conf. Comb. Math. Inst. Oxford 1972), Inst. Math. Appl., Southend-on-Sea, 1972, 141-154.

[HHL] J. Haglund, M. Haiman and N. Loehr *A Combinatorial formula for Macdonald Polynomials*, (arXiv:math.CO/0409538 v1).

[HT] F. Hivert and N. Thiéry, *MuPAD-Combinat, an Open Source Package for Research in Algebraic Combinatorics* Séminaire Lotharingien de Combinatoire **51** (2004). *http://mupad-combinat.sourceforge.net/*

[HW] G. M Hardy and E. M Wright, *An Introduction to the theory of numbers*, 5th ed, Oxford, Clarendon Press, 1979.

[LLT] A. Lascoux, B. Leclerc, J.-Y. Thibon, *Green polynomials and Hall-Littlewood functions at roots of unity*, Europ. J. Combinatorics (**15**), 173-180.

[LLT1] A. Lascoux, B. Leclerc, J. -Y. Thibon, *Fonctions de Hall-Littlewood et Polynomes de Kostka-Foulkes aux racines de l'unité*, Comptes Rendus de l'Academie des Sciences, Paris, (**316**) 1993, 1-6.

[LLT2] A. Lascoux, B. Leclerc, J. -Y. Thibon, *Ribbon Tableaux, Hall-Littlewood functions, quantum affine algebras and unipotent varieties*, Journal of Mathematical Physics (**38**) (1997), 1041-1068.

[M] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed., Oxford University Press, 1995.

[S] J. P. Serre, *Representations lineaires des groupes finis*, Hermann, Paris, 1977.

Institut Gaspard Monge, Université de Marne-la-Vallée 77454 Marne-la-Valle, France
E-mail address: `francois.descouens@univ-mlv.fr`

School of Science, Tokai University Hiratsuka 259-1292, Japan
E-mail address: `moritah@msg.biglobe.ne.jp`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The m-colored composition poset

Brian Drake and T. Kyle Petersen

Abstract. We generalize Björner and Stanley's poset of compositions to m-colored compositions. Their work draws many analogies between their (1-colored) composition poset and Young's lattice of partitions, including links to (quasi-)symmetric functions and representation theory. Here we show that many of these analogies hold for any number of colors. While many of the proofs for Björner and Stanley's poset were simplified by showing isomorphism with the subword order, we remark that with 2 or more colors, our posets are not isomorphic to a subword order.

Résumé. Nous généralisons le poset de Björner et de Stanley des compositions aux compositions m-coloré. Leur travail met à jour de nombreuses analogies entre leur poset sur les compositions (une couleur) et le trellis de Young sur les partitions, y compris des liens avec les fonctions (quasi-)symétriques et la théorie de représentations. Dans ce travail, nous prouvons que plusieurs de ces analogies sont vraies quel que soit le nombre de couleurs. Tandis que plusieurs des preuves dans le travail de Björner et de Stanley se simplifiaient en montrant un isomorphisme avec l'ordre sur les mots et les sous-mots, nous remarquons qu'avec deux couleurs ou plus, nos posets ne sont pas isomorphes à un tel ordre.

1. Introduction

This paper explores a generalization of the poset of compositions introduced in recent work of Björner and Stanley [**3**], which draws several analogies between their poset and Young's lattice of partitions. We recall some key facts about Young's lattice.

A partition $\lambda = (\lambda_1, \lambda_2, \ldots)$ of n, denoted $\lambda \vdash n$, is a sequence of nonnegative integers $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ such that $\sum \lambda_i = n$. The set of all partitions of all integers $n \geq 0$ forms a lattice under the partial order given by inclusion of Young diagrams: $\lambda \leq \mu$ if $\lambda_i \leq \mu_i$ for all i. This lattice is called *Young's lattice*, Y, which has several remarkable properties, including the following list given in [**3**].

- Y1. Y is a graded poset, where a partition $\lambda \vdash n$ has rank n.
- Y2. The number of saturated chains from the minimal partition $\emptyset$ to λ is the number f^λ of Young tableaux of shape λ.
- Y3. The number of saturated chains from $\emptyset$ to rank n is the number of involutions in the symmetric group $\mathfrak{S}_n$.
- Y4. Let s_λ denote a Schur function. Pieri's rule [**5**] gives
$$s_1 s_\lambda = \sum_{\lambda \prec \mu} s_\mu,$$
where $\lambda \prec \mu$ means that μ covers λ in Y.
- Y5. Since Y is in fact a distributive lattice, every interval $[\lambda, \mu]$ is EL-shellable and hence Cohen-Macaulay.
- Y6. Y is the Bratteli diagram for the tower of algebras $K\mathfrak{S}_0 \subset K\mathfrak{S}_1 \subset \cdots$, where $K\mathfrak{S}_n$ denotes the group algebra of $\mathfrak{S}_n$ over K, a field of characteristic zero.

1991 *Mathematics Subject Classification.* Primary 06A07; Secondary 05A99, 52B22.
Key words and phrases. colored composition, poset, CL-shellable.

For each of the properties listed above, Björner and Stanley give an analogous property for their poset of compositions. Here we will consider more generally the poset of m-colored compositions, and show that this poset has properties analogous to Y1–5 above. Finding an analog of property Y6 is an open problem.

Recall that a composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is an ordered tuple of positive integers, called the *parts* of α. We write $l(\alpha) = k$ for the number of parts of α. If the sum of the parts of α is n, i.e., $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_k = n$, then we say α is a composition of n, written $\alpha \models n$. Let $\mathrm{Comp}(n)$ denote all the compositions of n, and define the set of all compositions

$$\mathcal{C} := \bigcup_{n \geq 0} \mathrm{Comp}(n),$$

where $\emptyset$ is the unique composition of 0.

Björner and Stanley give $\mathcal{C}$ a partial order defined by the following covering relations. We say β *covers* α, written $\alpha \prec \beta$, if $\alpha < \beta$ and there is no β' such that $\alpha < \beta' < \beta$. We can obtain all compositions β that cover α by adding 1 to a part of α or adding 1 to a part and splitting that part into two parts. In other words, for some j we can write β as:

(1) $(\alpha_1, \ldots, \alpha_{j-1}, \alpha_j + 1, \alpha_{j+1}, \ldots, \alpha_k)$, or
(2) $(\alpha_1, \ldots, \alpha_{j-1}, h + 1, \alpha_j - h, \alpha_{j+1}, \ldots, \alpha_k)$ for some $0 \leq h \leq \alpha_j - 1$.

We can generalize this poset to a family of posets indexed by the number of "colors" $m \geq 1$. An m-colored composition is an ordered tuple of colored positive integers, say $\alpha = (\varepsilon_1\alpha_1, \varepsilon_2\alpha_2, \ldots, \varepsilon_k\alpha_k)$, where the α_s are positive integers and if ω is a primitive m-th root of unity, $\varepsilon_s = \omega^{i_s}$, $0 \leq i_s \leq m-1$. We say the part $\varepsilon_s\alpha_s$ has color ε_s, and we write $\alpha \models_m n$ if $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_k = n$. For example, if $m = 3$, then $\alpha = (\omega 2, 1, \omega^2 1, 3)$ is a 3-colored composition of $2 + 1 + 1 + 3 = 7$.

Note that there are m^k ways to color any ordinary composition of n with k parts, leading us to conclude that there are

$$\sum_{k=1}^{n} \binom{n-1}{k-1} m^k = m(m+1)^{n-1}$$

m-colored compositions of n (so if $m = 1$, we have 2^{n-1} ordinary compositions). Let $\mathrm{Comp}^{(m)}(n)$ denote the set of all m-colored compositions of n, and define

$$\mathcal{C}^{(m)} := \bigcup_{n \geq 0} \mathrm{Comp}^{(m)}(n),$$

where $\emptyset$ is again the unique composition of 0.

We can define a partial order on $\mathcal{C}^{(m)}$ with many of the same properties of $\mathcal{C}$. The covering relations are as follows. We have β covers α if, for some j, we can write β as:

(1) $(\varepsilon_1\alpha_1, \ldots, \varepsilon_{j-1}\alpha_{j-1}, \varepsilon_j(\alpha_j + 1), \varepsilon_{j+1}\alpha_{j+1}, \ldots, \varepsilon_k\alpha_k)$,
(2) $(\varepsilon_1\alpha_1, \ldots, \varepsilon_{j-1}\alpha_{j-1}, \varepsilon_j(h + 1), \varepsilon_j(\alpha_j - h), \varepsilon_{j+1}\alpha_{j+1}, \ldots, \varepsilon_k\alpha_k)$ for some $0 \leq h \leq \alpha_j - 1$, or
(3) $(\varepsilon_1\alpha_1, \ldots, \varepsilon_{j-1}\alpha_{j-1}, \varepsilon_j h, \varepsilon' 1, \varepsilon_j(\alpha_j - h), \varepsilon_{j+1}\alpha_{j+1}, \ldots, \varepsilon_k\alpha_k)$ where $\varepsilon' \neq \varepsilon_j$ and $0 \leq h \leq \alpha_j - 1$, with the understanding that we will ignore parts of size 0.

Relations (1) and (2) are just like those of $\mathcal{C}$: while preserving the color, we add 1 to a part, or we add 1 to a part and split that part into two parts. Relation (3) handles the case where the color of the "1" we add differs from where we try to add it. Notice that it is immediate from these cover relations that $\mathcal{C}^{(m)}$ is a graded poset with level n consisting of all m-compositions of n. This property is analogous to property Y1 of Young's lattice. See Figure 1 for the first four levels of the 2-colored composition poset.

In this paper we will show that for any positive fixed m, the poset $\mathcal{C}^{(m)}$ possesses properties analogous to $\mathcal{C} = \mathcal{C}^{(1)}$ (and indeed to Young's lattice). In fact many of the arguments used in [**3**] generalize in a straightforward way. One important argument that doesn't generalize is the isomorphism shown between $\mathcal{C}$ and the subword order; see Remark 4.4. In section 2 we discuss colored permutations, their color-descent compositions, and chains in $\mathcal{C}^{(m)}$. In section 3 we present Poirier's colored quasisymmetric functions [**4**] and show that $\mathcal{C}^{(m)}$ gives a Pieri-type rule for multiplying a fundamental basis. We define a CL-labeling in section 4 and use this to calculate the Möbius function of lower intervals. Section 5 contains the proof that this labeling is a CL-labeling.

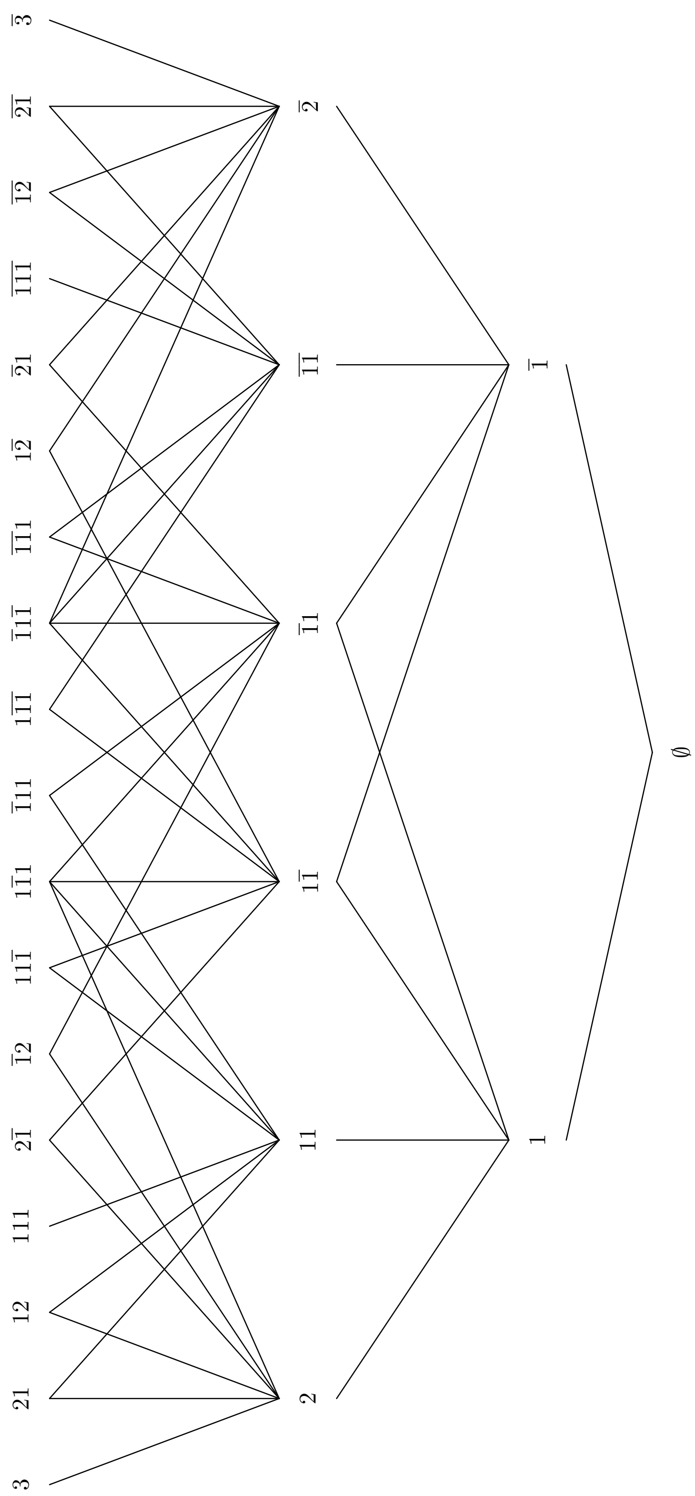


FIGURE 1. The first four levels of the 2-colored composition poset.

2. Colored permutations and descent sets

Compositions can be used to encode descent classes of ordinary permutations in the following way. Recall that a *descent* of a permutation $w \in \mathfrak{S}_n$ is a position i such that $w_i > w_{i+1}$, and that an *increasing run* (of length r) of a permutation w is a maximal subword of consecutive letters $w_{i+1}w_{i+2}\cdots w_{i+r}$ such that

$w_{i+1} < w_{i+2} < \cdots < w_{i+r}$. By maximality, we have that if $w_{i+1}w_{i+2}\cdots w_{i+r}$ is an increasing run, then i is a descent of w (if $i \neq 0$), and $i+r$ is a descent of w (if $i+r \neq n$). For any permutation $w \in \mathfrak{S}_n$ define the *descent composition*, $\mathrm{C}(w)$, to be the ordered tuple listing from left to right the lengths of the increasing runs of w. If $\mathrm{C}(w) = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, we can recover the descent set of w:

$$\mathrm{Des}(w) := \{i : w_i > w_{i+1}\} = \{\alpha_1, \alpha_1+\alpha_2, \ldots, \alpha_1+\alpha_2+\cdots+\alpha_{k-1}\}.$$

For example, the permutation $w = 345261$ has $\mathrm{C}(w) = (3,2,1)$ and $\mathrm{Des}(w) = \{3,5\}$. We now define colored permutations and colored descent compositions.

Loosely speaking, m-colored permutations are permutations where each of the elements permuted are given one of m "colors." If ω is any primitive m-th root of unity,

$$\omega 3 \ \omega 2 \ 1 \ \omega^3 4$$

is an example of a colored permutation. We can think of building colored permutations by taking an ordinary permutation and then arbitrarily assigning colors to the letters, so we see that there are $m^n n!$ m-colored permutations of $[n] := \{1, 2, \ldots, n\}$.

Strictly speaking, m-colored permutations are elements of the wreath product $C_m \wr \mathfrak{S}_n$, where $C_m = \{1, \omega, \ldots, \omega^{m-1}\}$ is the cyclic group of order m. We write an element $u = u_1u_2\cdots u_n \in C_m \wr \mathfrak{S}_n$ as a word in the alphabet

$$C_m \times [n] := \{1, 2, \ldots, n, \omega 1, \omega 2, \ldots, \omega n, \ldots, \omega^{m-1}1, \omega^{m-1}2, \ldots, \omega^{m-1}n\},$$

such that $|u| = |u_1||u_2|\cdots|u_n|$ is an ordinary permutation in $\mathfrak{S}_n$. We say $\varepsilon_i = u_i/|u_i|$ is the color of u_i.

For any $u \in C_m \wr \mathfrak{S}_n$, we can write $u = v_1v_2\cdots v_k$ so that each v_i is a word in which all the letters have the same color, ε_i', and no two consecutive colors are the same: $\varepsilon_i' \neq \varepsilon_{i+1}'$, $i = 1, 2, \ldots, k-1$. Then we define the *color composition* of u,

$$\mathrm{Col}(u) := (\varepsilon_1'\alpha_1', \varepsilon_2'\alpha_2', \ldots, \varepsilon_k'\alpha_k'),$$

where α_s' denotes the number of letters in v_s. Now suppose an m-colored permutation u has color composition $\mathrm{Col}(u) = (\varepsilon_1'\alpha_1', \varepsilon_2'\alpha_2', \ldots, \varepsilon_k'\alpha_k')$. Then the *colored descent composition*

$$\mathrm{C}^{(m)}(u) := (\varepsilon_1\alpha_1, \varepsilon_2\alpha_2, \ldots, \varepsilon_l\alpha_l),$$

is the refinement of $\mathrm{Col}(u)$ where we replace part $\varepsilon_i'\alpha_i'$ with $\varepsilon_i'\,\mathrm{C}(|v_i|)$, where C is the ordinary descent composition, and we view $|v_i|$ as an ordinary permutation of distinct letters.

More intuitively, the colored descent composition $\mathrm{C}^{(m)}(u)$ is the ordered tuple listing the lengths of increasing runs of u with constant color, where we record not only the length of such a run, but also its color. An example should cement the notion. If we have two colors (indicated with a bar), let

$$u = 1\bar{2}\bar{3}4\bar{8}\bar{5}76.$$

Then the color composition is $\mathrm{Col}(u) = (1, \bar{2}, 1, \bar{2}, 2)$, and

$$\mathrm{C}^{(m)}(u) = (1, \bar{2}, 1, \bar{1}, \bar{1}, 1, 1).$$

For any $\alpha \in \mathrm{Comp}^{(m)}(n)$, a *saturated chain* from $\emptyset$ to α is a sequence of compositions

$$\emptyset = \alpha^0 \prec \alpha^1 \prec \cdots \prec \alpha^n = \alpha,$$

where $\prec$ denotes a cover relation in $\mathcal{C}^{(m)}$, and therefore $\alpha^i \in \mathrm{Comp}^{(m)}(i)$. Now, given any $u \in C_m \wr \mathfrak{S}_n$, let $u[i]$ denote the restriction of u to letters in $C_m \times [i]$. For example, if $u = \bar{2}17\bar{6}\bar{3}\bar{4}58$, then $u[5] = \bar{2}1\bar{3}\bar{4}5$. We then define the sequence

$$\mathfrak{m}(u) := (\mathrm{C}^{(m)}(u[1]), \ldots, \mathrm{C}^{(m)}(u[n])),$$

so that $\mathrm{C}^{(m)}(u[i]) \in \mathrm{Comp}^{(m)}(i)$. Using the same example $u = \bar{2}17\bar{6}\bar{3}\bar{4}58$, we have

$$\mathfrak{m}(u) = (1, \bar{1}1, \bar{1}1\bar{1}, \bar{1}1\bar{2}, \bar{1}1\bar{2}1, \bar{1}1\bar{1}\bar{2}1, \bar{1}2\bar{1}\bar{2}1, \bar{1}2\bar{1}\bar{2}2).$$

The following theorem is the natural generalization of Theorem 2.1 of [**3**].

Theorem 2.1. *The map $\mathfrak{m}$ is a bijection from $C_m \wr \mathfrak{S}_n$ to saturated chains from $\emptyset$ to α, where α ranges over all colored compositions in* $\mathrm{Comp}^{(m)}(n)$.

This proof follows the same line of reasoning used by Björner and Stanley in proving Theorem 2.1 of [**3**].

PROOF. For any colored permutation $u \in C_m \wr \mathfrak{S}_n$ define, for all $0 \leq i \leq n$ and all $0 \leq j \leq m-1$,

$$u_{(i,j)} := u_1 \cdots u_i \, \omega^j (n+1) \, u_{i+1} \cdots u_n.$$

In other words, the $u_{(i,j)}$ are all those permutations w in $C_m \wr \mathfrak{S}_{n+1}$ such that $w[n] = u$. We will show that the compositions $\mathrm{C}^{(m)}(u_{(i,j)})$ are all distinct and moreover that they are precisely those compositions in $\mathrm{Comp}^{(m)}(n+1)$ that cover $\mathrm{C}^{(m)}(u)$.

Suppose $\mathrm{C}^{(m)}(u) = (\varepsilon_1\alpha_1, \ldots, \varepsilon_k\alpha_k)$, and let $b_s = \alpha_1 + \cdots + \alpha_s$, with the convention that $b_0 = 0$. For any fixed $j = 0, 1, \ldots, m-1$, we have two cases, corresponding to cover relations of type (1) or type (3):

$$\mathrm{C}^{(m)}(u_{(b_s,j)}) = \begin{cases} (\varepsilon_1\alpha_1, \ldots, \varepsilon_s(\alpha_s+1), \ldots, \varepsilon_k\alpha_k) & \text{if } \varepsilon_s = \omega^j, \\ (\varepsilon_1\alpha_1, \ldots, \varepsilon_s\alpha_s, \omega^j 1, \ldots, \varepsilon_k\alpha_k) & \text{otherwise.} \end{cases}$$

All these compositions, over $s = 0, \ldots, k$, $j = 0, \ldots, m-1$, are distinct and cover $\mathrm{C}^{(m)}(u)$. To consider the other cases, suppose i is not of the form $\alpha_1 + \cdots + \alpha_s$. Then it can be written as $i = \alpha_1 + \cdots + \alpha_s + h$, where $0 \leq s \leq k$ and $1 \leq h \leq \alpha_{s+1} - 1$ (if $s = 0$, then $i = h$). Again we have two cases, corresponding to cover relations of type (2) or type (3):

$$\mathrm{C}^{(m)}(u_{(i,j)}) = \begin{cases} (\varepsilon_1\alpha_1, \ldots, \varepsilon_{s+1}(1+h), \varepsilon_{s+1}(\alpha_{s+1}-h), \ldots, \varepsilon_k\alpha_k) & \text{if } \varepsilon_{s+1} = \omega^j, \\ (\varepsilon_1\alpha_1, \ldots, \varepsilon_{s+1}h, \omega^j 1, \varepsilon_{s+1}(\alpha_{s+1}-h), \ldots, \varepsilon_k\alpha_k) & \text{otherwise.} \end{cases}$$

These cases are again distinct and provide the remaining covers for $\mathrm{C}^{(m)}(u)$. $\square$

Theorem 2.1 yields several easy corollaries. The first is analogous to property Y2 of Young's lattice; the second corresponds to Y3.

COROLLARY 2.2. *The number of saturated chains from $\emptyset$ to α in $\mathcal{C}^{(m)}$ is equal to the number $f_n^{(m)}(\alpha)$ of m-colored permutations w with colored descent composition α.*

COROLLARY 2.3. *The total number of saturated chains from $\emptyset$ to rank n is equal to the number of m-colored permutations of* $[n]$,

$$\sum_{\alpha \in \mathrm{Comp}^{(m)}(n)} f_n^{(m)}(\alpha) = m^n n!.$$

COROLLARY 2.4. *The number of m-colored compositions $\beta \in \mathrm{Comp}^{(m)}(n+1)$ covering $\alpha \in \mathrm{Comp}^{(m)}(n)$ is $m(n+1)$.*

3. Colored quasisymmetric functions

One key use for compositions is as an indexing set for quasisymmetric functions. Similarly, there exist colored quasisymmetric functions (due to Poirier [**4**]) that use colored compositions as indices. Both these situations are analogous to how partitions index symmetric functions.

Recall ([**5**], ch. 7.19) that a quasisymmetric function is a formal series

$$Q(x_1, x_2, \ldots) \in \mathbb{Z}[[x_1, x_2, \ldots]]$$

of bounded degree such that for any composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, the coefficient of $x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_k}^{\alpha_k}$ with $i_1 < i_2 < \cdots < i_k$ is the same as the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k}$. One natural basis for the quasisymmetric functions homogeneous of degree n is given by the *fundamental quasisymmetric functions*, L_α, where α ranges over all of $\mathrm{Comp}(n)$. If $\alpha = (\alpha_1, \ldots, \alpha_k) \models n$, then define

$$L_\alpha := \sum x_{i_1} \cdots x_{i_n},$$

where the sum is taken over all $i_1 \leq i_2 \leq \cdots \leq i_n$ with $i_s < i_{s+1}$ if $s = \alpha_1 + \cdots + \alpha_r$ for some r. For example,

$$L_{21} = \sum_{i \leq j < k} x_i x_j x_k.$$

Colored quasisymmetric functions are simply a generalization of quasisymmetric functions to an alphabet with several colors for its letters. For fixed m, we consider formal series in the alphabet

$$X^{(m)} := \{x_{0,1}, x_{0,2}, \ldots, x_{1,1}, x_{1,2}, \ldots, x_{m-1,1}, x_{m-1,2}, \ldots\},$$

(so the first subscript corresponds to color) with the same quasisymmetric property. Namely, an m-colored quasisymmetric function $Q(X^{(m)})$ is a formal series of bounded degree such that for any m-colored composition $\alpha = (\omega^{j_1}\alpha_1, \ldots, \omega^{j_k}\alpha_k)$, the coefficient of $x_{j_1,i_1}^{\alpha_1} x_{j_2,i_2}^{\alpha_2} \cdots x_{j_k,i_k}^{\alpha_k}$ with $i_1 < i_2 < \cdots < i_k$ is the same as the coefficient of $x_{j_1,1}^{\alpha_1} x_{j_2,2}^{\alpha_2} \cdots x_{j_k,k}^{\alpha_k}$. Intuitively, the letters are colored the same as the parts of α. The *m-colored fundamental quasisymmetric functions* are defined as follows. First, if $s = \alpha_1 + \cdots + \alpha_r + h$, $1 \le h \le \alpha_{r+1}$, then define $j'_s = j_{r+1}$, the color of part α_{r+1}. Then,

$$L_\alpha^{(m)} := \sum x_{j'_1,i_1} \cdots x_{j'_n,i_n},$$

where the sum is taken over all $i_1 \le i_2 \le \cdots \le i_n$ with $i_s < i_{s+1}$ if both $j'_s \ge j'_{s+1}$ and $s = \alpha_1 + \cdots + \alpha_r$ for some r. For example,

$$L^{(2)}_{1\bar{2}\bar{1}} = \sum_{i \le j \le k < l} x_i y_j y_k y_l \quad \text{and} \quad L^{(3)}_{2\bar{\bar{1}}\bar{2}} = \sum_{i \le j \le k < l \le m} x_i x_j z_k y_l y_m.$$

As in the ordinary case, the $L_\alpha^{(m)}$, where α ranges over $\mathrm{Comp}^{(m)}(n)$, give a basis for the m-colored quasisymmetric functions homogeneous of degree n.

There is a nice formula for multiplying colored quasisymmetric functions in the fundamental basis. Let $u \in C_m \wr \mathfrak{S}_n$ and let v be an m-colored permutation of the set $\{n+1, n+2, \ldots, n+r\}$. Let $\alpha = \mathrm{C}^{(m)}(u)$ and $\beta = \mathrm{C}^{(m)}(v)$. Then we have

$$L_\alpha^{(m)} L_\beta^{(m)} = \sum_w L^{(m)}_{\mathrm{C}^{(m)}(w)},$$

where the sum is taken over all *shuffles* w of u and v, i.e., all colored permutations $w \in C_m \wr \mathfrak{S}_{n+r}$ such that $w[n] = u$ and w restricted to $\{n+1, n+2, \ldots, n+r\}$ is v.

If $r = 1$, then we see that the shuffles of u and $v = \omega^j(n+1)$ are precisely those permutations $u_{(i,j)}$ from the proof of Theorem 2.1. Applying the multiplication rule, and summing over all j, we have a Pieri-type rule analogous to property Y4 of Young's lattice.

Proposition 3.1. *We have:*

$$(L_1^{(m)} + L_{\omega 1}^{(m)} + \cdots + L^{(m)}_{\omega^{m-1}1}) L_\alpha^{(m)} = \sum_{\alpha \prec \beta} L_\beta^{(m)}.$$

As Björner and Stanley remark in the case of a single color, we could have used Proposition 3.1 to define the poset $\mathcal{C}^{(m)}$ in the first place. At the least, it is a good justification for the study of $\mathcal{C}^{(m)}$.

Repeated application of the proposition gives the formula

$$(L_1^{(m)} + L_{\omega 1}^{(m)} + \cdots + L^{(m)}_{\omega^{m-1}1})^n = \sum_{\alpha \in \mathrm{Comp}^{(m)}(n)} f_n^{(m)}(\alpha) L_\alpha^{(m)},$$

where $f_n^{(m)}$ is the number of m-colored permutations with colored descent composition α. This equation is equivalent to Corollary 2.2, and analogous to the following formula for Schur functions (see [**5**]) that corresponds to property Y2 of Young's lattice:

$$s_1^n = \sum_{\lambda \vdash n} f^\lambda s_\lambda,$$

where f^λ is the number of Young tableaux of shape λ.

4. Shellability and Möbius function

In this section we show that $\mathcal{C}^{(m)}$ is CL-shellable by giving an explicit dual CL-labeling. See [**2**] for an introduction to CL-shellable posets. We use a model of removing colored balls from urns to define our labeling on downward maximal chains. Given a colored composition of n, $\alpha = (\varepsilon_1\alpha_1, \ldots, \varepsilon_k\alpha_k)$, we picture k urns next to each other, labeled $U_1, U_2, \ldots, U_k$ from left to right. In urn U_i we start with α_i balls of color ε_i, for a total of n balls. Moving down along a maximal chain, we remove a ball from an urn for each covering relation, and possibly move some balls from one urn to another. There are three different types of moves, which we now describe. After some number of steps, suppose that U_i is a nonempty urn, U_h is the first nonempty urn on its left, and U_j is the first nonempty urn on its right. Let $\beta_i, \beta_h, \beta_j$ be the number of balls in the corresponding urns and let $\varepsilon_i, \varepsilon_h, \varepsilon_j$ be the colors of those balls. The three possible moves are:

(1) If $\beta_i \geq 2$, or if $\varepsilon_h \neq \varepsilon_i$, or if U_i is the first nonempty urn, then remove a ball from urn U_i.
(2) If $\beta_i = 1$ and $\varepsilon_h = \varepsilon_j \neq \varepsilon_i$, then remove the ball from U_i and place all the balls from U_h and U_j into U_i.
(3) If $\beta_i \geq 2$ and $\varepsilon_j = \varepsilon_i$, then move all balls from U_j to U_i and remove a ball from U_i.

After any number of moves, we may associate the distribution of colored balls in urns with an element of $\mathcal{C}^{(m)}$. The different urns represent the parts of the composition, the number of balls in an urn is the size of that part, and the color of the balls is the color of the part. Here we ignore parts of size 0. Notice that the color of a part is well defined, since none of the moves allows balls of different colors to be combined in a single urn. It is an easy exercise to check that each covering relation in $\mathcal{C}^{(m)}$ corresponds to one of these three possible moves for some urn, and furthermore that the urn and type of move are unique.

Let $[\emptyset, \alpha]$ be an interval in $\mathcal{C}^{(m)}$, $|\alpha| = n$. We will now define a labeling $\lambda(c) = (\lambda_1(c), \lambda_2(c), \ldots, \lambda_n(c))$ for a maximal chain

$$c = (\alpha = \alpha^0 \succ \alpha^1 \succ \cdots \succ \alpha^n = \emptyset).$$

Our set of labels is $\mathbb{N} \times \{1, 2, 3\}$, totally ordered with the lexicographic order. For each covering relation $\alpha^{r-1} \succ \alpha^r$ we have a unique urn and type of move that takes the distribution of balls in urns for α^{r-1} to the distribution for α^r. Suppose that move is of type t, and removes a ball from urn U_i. Then we define the label $\lambda_r(c) = (i, t)$.

Notice that with labels defined on maximal chains in lower intervals $[\emptyset, \alpha]$, there is an induced labeling defined on maximal chains in arbitrary intervals $[\beta, \alpha]$. As an example, consider the following two maximal chains in $[3, 22\bar{1}2]$:

$$c_0 = (22\bar{1}2 \succ 12\bar{1}2 \succ 2\bar{1}2 \succ 1\bar{1}2 \succ 3)$$
$$c = (22\bar{1}2 \succ 21\bar{1}2 \succ 2\bar{1}2 \succ 22 \succ 3)$$

They are labeled $\lambda(c_0) = ((1,1), (1,1), (2,1), (3,2))$ and $\lambda(c) = ((2,1), (1,3), (3,1), (1,3))$. Pictured as colored balls and urns, we have:

$$c_0 : \lfloor\bullet\bullet\rfloor\lfloor\bullet\bullet\rfloor\lfloor\circ\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\bullet\rfloor\lfloor\bullet\bullet\rfloor\lfloor\circ\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\,\rfloor\lfloor\bullet\bullet\rfloor\lfloor\circ\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\,\rfloor\lfloor\bullet\rfloor\lfloor\circ\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\,\rfloor\lfloor\,\rfloor\lfloor\bullet\bullet\bullet\rfloor\lfloor\,\rfloor$$
$$c : \lfloor\bullet\bullet\rfloor\lfloor\bullet\bullet\rfloor\lfloor\circ\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\bullet\bullet\rfloor\lfloor\bullet\rfloor\lfloor\circ\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\bullet\bullet\rfloor\lfloor\,\rfloor\lfloor\circ\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\bullet\bullet\rfloor\lfloor\,\rfloor\lfloor\,\rfloor\lfloor\bullet\bullet\rfloor \to \lfloor\bullet\bullet\bullet\rfloor\lfloor\,\rfloor\lfloor\,\rfloor\lfloor\,\rfloor.$$

In fact, the chain c_0 above is lexicographically minimal and has the only increasing label. Notice that if we start with all balls of the same color, moves of type (2) cannot occur and we recover the construction in the appendix of [**3**]. These labels agree with the labels defined there, with $(i,1) \mapsto i$ and $(i,3) \mapsto i'$.

By proving that this labeling is in fact a CL-labeling, we obtain our analog of property Y5 of Young's lattice. The proof is given in section 5.

THEOREM 4.1. *Intervals in $\mathcal{C}^{(m)}$ are dual CL-shellable and hence Cohen-Macaulay.*

Now we calculate the Möbius function of lower intervals. As always, we must have $\mu_{\mathcal{C}^{(m)}}(\emptyset, \emptyset) = 1$. For $\alpha \neq \emptyset$, we have the following.

PROPOSITION 4.2.

$$\mu_{\mathcal{C}^{(m)}}(\emptyset, \alpha) = \begin{cases} (-1)^{|\alpha|} & \textit{if } \alpha = (\varepsilon_1 1, \varepsilon_2 1, \ldots, \varepsilon_{|\alpha|} 1) \\ & \textit{for some colors } \varepsilon_1 \neq \varepsilon_2 \neq \cdots \neq \varepsilon_{|\alpha|}, \\ 0 & \textit{otherwise.} \end{cases}$$

PROOF. We make use of the combinatorial description of the Möbius function for a graded poset with a CL-labeling, given in [**2**]. That is, the Möbius function of an interval is -1 to the length of the interval, times the number of maximal chains with a strictly decreasing label.

Suppose that $\alpha = (\varepsilon_1 1, \varepsilon_2 1, \ldots, \varepsilon_{|\alpha|} 1)$, with $\varepsilon_1 \neq \varepsilon_2 \neq \cdots \neq \varepsilon_{|\alpha|}$. Then there is a unique chain with a strictly decreasing label, obtained by removing the balls from right to left using only type (1) moves. Therefore $\mu_{\mathcal{C}^{(m)}}(\emptyset, \alpha) = (-1)^{|\alpha|}$.

Now suppose that α has a part i of size 2 or greater. We want to show that there is no chain in $[\emptyset, \alpha]$ with a strictly decreasing label. Any chain that makes a type (1) move from the same urn twice will have a repeated label. The only way to remove the balls from urn i and possibly have a decreasing label is to remove at most one ball with a type (1) move, and then move all the balls to an urn on the left with a type (2) or (3) move. But in the new urn we have at least two balls, and the process repeats. At some point we must have an urn with at least two balls and no way to make a type (2) or (3) move. Then we must use two type (1) moves for the same urn, so the chain label cannot be strictly decreasing.

Finally, suppose that α has parts α_i and α_{i+1} of size 1 and the same color. The only legal way to remove the balls from the corresponding urns is to remove the left one first, creating an increase in the chain label. $\square$

Note that for $\alpha \models_m n$ with $\mu_{\mathcal{C}^{(m)}}(\emptyset, \alpha) \neq 0$, there are m choices for the color of the first part, and $m-1$ choices for the color of each succeeding part. Hence there are $m(m-1)^{n-1}$ compositions $\alpha \models_m n$ with $\mu_{\mathcal{C}^{(m)}}(\emptyset, \alpha) \neq 0$. For $m > 1$ an elementary calculation gives the following generating function.

$$\sum_{\alpha \in \mathcal{C}^{(m)}} \mu_{\mathcal{C}^{(m)}}(\emptyset, \alpha) t^{|\alpha|} = \frac{1+t}{1-(m-1)t}.$$

Define the following "truncated" poset,

$$\mathcal{C}_n^{(m)} := \widehat{1} \cup \bigcup_{1 \leq i \leq n} \mathrm{Comp}^{(m)}(i),$$

with the order relation as before except with a new maximal element $\widehat{1}$ that covers all the compositions in $\mathrm{Comp}^{(m)}(n)$.

Corollary 4.3. *The poset $\mathcal{C}_n^{(m)}$ is shellable, with Möbius function*

$$\mu(\emptyset, \widehat{1}) = (-1)^{n+1}(m-1)^n.$$

The proof of this corollary follows the argument of [**3**].

Proof. First we want to show that every m-colored composition $\alpha = (\varepsilon_1\alpha_1, \varepsilon_2\alpha_2, \ldots, \varepsilon_k\alpha_k)$ of at most n lies below the composition $\gamma^n \models_m mn$, defined as the concatenation of n copies of $\gamma = (1, \omega 1, \omega^2 1, \ldots, \omega^{m-1}1)$. To the i^{th} part of α we can associate α_i copies of γ. First, we use covering relations of type (2) (as originally described), α_i times to split the part into all parts of size 1 and color ε_i. Then we use covering relations of type (3) to fill in the remaining 1's of different colors. Therefore $\mathcal{C}_n^{(m)}$ is obtained via rank selection from the interval $[\emptyset, \gamma^n]$, so shellability follows by results of [**1**].

For the Möbius function:

$$\begin{aligned}
\mu(\emptyset, \widehat{1}) = -\sum_{|\alpha| \leq n} \mu(\emptyset, \alpha) &= -\left(1 + \sum_{k=1}^{n} (-1)^k m(m-1)^{k-1}\right) \\
&= -\left(1 - m\sum_{k=0}^{n-1} (1-m)^k\right) \\
&= -\left(1 - m\left(\frac{1-(1-m)^n}{1-(1-m)}\right)\right) \\
&= (-1)^{n+1}(m-1)^n.
\end{aligned}$$

$\square$

Remark 4.4. Björner and Stanley show that $\mathcal{C}^{(1)}$ is isomorphic to the subword order on 2 letters. This allows the transfer of many results, such as CL-shellability and the Möbius function of an arbitrary interval. Since there are $m(m+1)^{n-1}$ colored compositions of n, one might expect that $\mathcal{C}^{(m)}$ is isomorphic to the subword order on $m+1$ letters, with a restriction on the first or last letter of a word. However, this turns out to be false. For example, consider the colored composition $(1, \bar{1}, 1)$, which is present in $\mathcal{C}^{(m)}$ for all $m \geq 2$. It covers 4 colored compositions: $(2), (1, \bar{1}), (\bar{1}, 1), (1, 1)$, but a word of length 3 can cover at most 3 subwords. Therefore the results of this section do not follow directly from properties of the subword order.

5. Proof of CL-shellability

In this section we give the proof of Theorem 4.1. We first note that the labeling is a well defined chain labeling. That is, if two chains agree on their first k edges, then their first k labels agree. This is clear from the definition.

The labeling of maximal chains in $[\emptyset, \alpha]$ gives an induced labeling on rooted intervals $([\beta, \alpha], c)$, where c is a maximal chain in $[\emptyset, \beta]$. This induced labeling is of the same kind, so it suffices to check that the properties of a CL-labeling hold in an arbitrary interval $[\beta, \alpha]$.

First, we want to show that the chain with the lexicographically first label has a weakly increasing label. The lexicographically first label is well defined, since all the moves from a given distribution of balls in urns have distinct labels. Moreover, we can describe the lexicographically first chain, c_0, as follows. If

$$c_0 = (\alpha = \alpha^0 \succ \alpha^1 \succ \cdots \succ \alpha^k = \beta),$$

then at each step, to move down from α^{r-1} to α^r, we must remove a ball from an urn as far to the left as possible, such that the new composition is still in the interval $[\beta, \alpha]$. To prove that $\lambda_1(c_0) \leq \lambda_2(c_0) \leq \cdots \leq \lambda_k(c_0)$, we use the following lemma.

Lemma 5.1. *On an interval of length two, the chain with the lexicographically first label is weakly increasing.*

Proof. On an interval of length two, all chains correspond to removing two balls from urns, such that the starting and ending distributions are the same. For the chain c_0 with the lexicographically first label, the urns are consecutively chosen to be as far to the left as possible. Suppose that one ball is removed from urn U_i and the other ball is removed from urn U_j.

If $j > i + 1$, then there is a nonempty urn between U_i and U_j, and removing a ball from one of the urns does not affect the possibility of removing the other ball from its urn. Therefore it is clear that $\lambda(c_0)$ is weakly increasing.

Suppose that $j = i + 1$, so there is no urn between U_i and U_j. Removing a ball from U_i does not affect the possibility of removing a ball from U_{i+1}, unless $\varepsilon_i \neq \varepsilon_{i+1} = \varepsilon_{i-1}$ and $\alpha_i = 1$. In this case, the urns could have been chosen to be U_h and U_i, for an appropriate urn U_h with $h < i$, making a type (1) or type (3) move in U_h and then a type (2) move in urn U_i. However, this contradicts our assumption that U_i was chosen to be the leftmost possible. So if $i \neq j$, $\lambda(c_0)$ is weakly increasing.

The only remaining case is if $U_i = U_j$. If $\alpha_i > 2$, then we have the weakly increasing label $\lambda(c_0) = ((i, 1), (i, 1))$. Now suppose that $\alpha_i = 2$ and $\varepsilon_{i-1} \neq \varepsilon_i$. Then we also have $\lambda(c_0) = ((i, 1), (i, 1))$. If $\varepsilon_{i-1} = \varepsilon_i$, then c_0 is found by choosing an appropriate urn $U_h, h < i$ and making a type (1) or (3) move in urn U_h and then a type (1) move in urn U_i. Again, this contradicts our assumption that U_i was chosen to be the leftmost possible. Therefore if $i = j$, $\lambda(c_0)$ is weakly increasing. $\square$

Returning to the general case, for every r, the induced labeling of c_0 on the chain $\alpha^{r-1} \succ \alpha^r \succ \alpha^{r+1}$ is lexicographically first on the interval $[\alpha^{r+1}, \alpha^{r-1}]$. Then by Lemma 5.1, $\lambda_1(c_0) \leq \lambda_2(c_0) \leq \cdots \leq \lambda_k(c_0)$, i.e., $\lambda(c_0)$ is weakly increasing. Now it remains only to show no chain other than c_0 has a weakly increasing label.

If another chain results with the same distribution of balls into urns as the lexicographically first chain, including the locations of the empty urns, then the label on that chain must have a descent. To see this, consider the point where it deviates from the lexicographically first chain. It is leaving a ball behind in a lower numbered urn. At some later step it must remove a ball from that urn, which will create a descent.

Now we need to consider chains which result in the same distribution of balls into nonempty urns as the lexicographically first chain, but such that the empty urns are in different positions. Let c be such a chain. Since the lexicographically first chain removes balls from urns from left to right, the final distribution of balls into urns for the lexicographically first chain has its nonempty urns as far to the right as possible (though they need not be consecutive).

Let $U_1, U_2, \ldots$ be the urns labeled from left to right. Let $U_i(c)$ be the number of balls in urn U_i in the final distribution for the chain c. Let r be the largest number such that $U_r(c) > 0$ and $U_r(c_0) = 0$. There is such an r by our assumption on c. We must have a j such that: 1) $U_{r+i}(c) = 0$ for all $1 \leq i \leq j$, 2) $U_{r+i}(c_0) = 0$ for all $1 \leq i < j$, and 3) $U_{r+j}(c_0) \neq 0$. Note that the color and number of the balls in urn U_r for c is the same as the color and number of the balls in urn U_{r+j} for c_0.

If c has an increasing label, then the urns $U_{r+1}, \ldots, U_{r+j}$ must be emptied from left to right. Therefore at some point the urns $U_{r+1}, \ldots, U_{r+j-1}$ are empty and we need to remove the last ball from U_{r+j}. The only way to do this is to use a move of type (2), which has a label $(r, 2)$ and creates a descent.

References

[1] A. Björner, *Shellable and Cohen-Macaulay posets*, Trans. Amer. Math. Soc. **260** (1980), 159–183.

[2] A. Björner, A. Garsia and R. Stanley, *Cohen-Macaulay partially ordered sets* in *Ordered Sets*, (I. Rival, ed.), Reidel, Dordrecht/Boston, 1982, pp. 583–615.

[3] A. Björner and R. Stanley, *An analogue of Young's lattice for compositions*, arXiv: math.CO/0508043.

[4] S. Poirier, *Cycle type and descent set in wreath products*, Discrete Mathematics **180** (1998), 315–343.

[5] R. Stanley, *Enumerative Combinatorics, Volume II*, Cambridge University Press, 2001.

Department of Mathematics, Brandeis University, Waltham, MA, USA, 02454
E-mail address: bdrake@brandeis.edu
URL: http://people.brandeis.edu/~bdrake

Department of Mathematics, Brandeis University, Waltham, MA, USA, 02454
E-mail address: tkpeters@brandeis.edu
URL: http://people.brandeis.edu/~tkpeters

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The number of Z-convex polyominoes

Enrica Duchi, Simone Rinaldi, and Gilles Schaeffer

Abstract. In this paper we consider a restricted class of polyominoes that we call Z-convex polyominoes. Z-convex polyominoes are polyominoes such that any two pairs of cells can be connected by a monotone path making at most two turns (like the letter Z). In particular they are convex polyominoes, but they appear to resist standard decompositions. We propose a construction by "inflation" that allows us, through a quite tedious case analysis, to write a system of functional equations for their generating functions. Even though intermediate steps involve heavy computations, it turns out in the end that the generating function $P(t)$ of Z-convex polyominoes with respect to the semi-perimeter can be expressed as a simple rational function of t and the generating function of Catalan numbers, like the generating function of convex polyominoes.

Résumé. Dans cet article nous étudions une classe restreinte de polyominos que nous appelons Z-convexes : un polyomino est Z-convexe si deux cases quelconques peuvent toujours être jointes par un chemin monotone avec au plus deux virages. En particulier ces polyominos sont des polyominos convexes, mais ils ne se laissent pas compter par les méthodes usuelles (en tout cas pas par nous). Nous proposons une construction "gonflée" qui permet d'écrire un système d'équations fonctionnelles pour les séries gératrices associées au terme d'une longue analyse de cas. De manière inattendue, bien que les calculs intermédiaires soient assez touffus, la série génératrice $P(t)$ des polyominos Z-convexes selon le demi-périmètre, à l'instar de la série génératrice des polyominos convexes, s'exprime comme une fonction rationnelle en x et en la série génératrice des nombres de Catalan.

1. Introduction

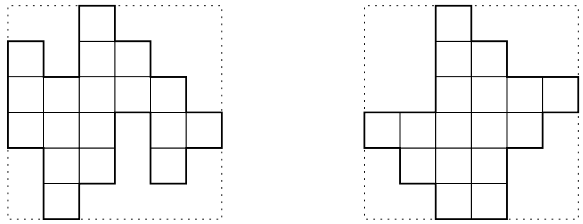
1.1. Convex polyominoes. In the plane $\mathbb{Z} \times \mathbb{Z}$ a *cell* is a unit square, and a *polyomino* is a finite connected union of cells having no cut point. Polyominoes are defined up to translations. A *column* (*row*) of a polyomino is the intersection between the polyomino and an infinite strip of cells lying on a vertical (horizontal) line. For the main definitions and results concerning polyominoes we refer to [**S**] and, for those who can read french, to [**BM**]. Invented by Golomb [**G2**] who coined the term *polyomino*, these combinatorial objects are related to many mathematical problems, such as tilings [**BN, G1**], or games [**Ga**] among many others. The enumeration problem for general polyominoes is difficult to solve and still open. The number a_n of polyominoes with n cells is known up to $n = 56$ [**JG**] and asymptotically, these numbers satisfy the relation $\lim_n (a_n)^{1/n} = \mu$, $3.96 < \mu < 4.64$, where the lower bound is a recent improvement of [**BMRR**].

In order to probe further, several subclasses of polyominoes have been introduced on which to hone enumeration techniques. One natural subclass is that of *convex* polyominoes. A polyomino is said to be *column-convex* (*row-convex*) when its intersection with any vertical (horizontal) line of cells in the square lattice is connected (see Fig. 1 (a)), and *convex* when it is both column and row-convex (see Fig. 1 (b)). The *area* of a polyomino is just the number of cells it contains, while its *semi-perimeter* is half the length of the boundary. Thus, in a convex polyomino the semi-perimeter is the sum of the numbers of its rows and columns. Moreover, any convex polyomino is contained in a rectangle in the square lattice which has the same semi-perimeter (called the *minimal bounding rectangle* of the polyomino).

2000 *Mathematics Subject Classification.* Primary 05A15; Secondary 82B41.

Key words and phrases. Enumeration, algebraic generating functions, recursive decomposition.

The authors acknowledge support from the french ANR under the SADA project.



(a) (b)

FIGURE 1. (a) a column-convex (but not convex) polyomino; (b) a convex polyomino.

The number f_n of convex polyominoes with semi-perimeter $n+2$ was obtained by Delest and Viennot, in [**DV**]:

$$f_{n+2} = (2n+11)4^n - 4(2n+1)\binom{2n}{n}, \quad n \geq 0; \quad f_0 = 1, \quad f_1 = 2.$$

In particular the generating function of convex polyominoes with respect to the semi-perimeter

$$F(t) = \sum_{n\geq 0} f_n t^{n+2} = t^2 + 2t^3 + 7t^4 + 28t^5 + 120t^6 + 528t^7 + O(t^8)$$

is an algebraic series which has a rational expression $R_0(t, d(t))$ in t and the Catalan generating function

$$c(t) = \frac{1-\sqrt{1-4t}}{2t} = 1 + t + 2t^2 + 5t^3 + 14t^4 + 42t^5 + 132t^6 + O(t^7).$$

More precisely, the generating function of convex polyominoes with respect to the numbers of columns (variable x) and rows (variable y) is

$$F(x,y) = \frac{8x^2y^2d(x,y)}{\Delta^2} + \frac{xy(1-x-xy-y)}{\Delta},$$

where $d(x,y)$ is the unique power series satisfying the relation $d = (x+d)(y+d)$,

$$d(x,y) = \frac{1}{2}(1 - x - y - \sqrt{\Delta}),$$

and,

$$\Delta = (1-x-y)^2 - 4xy = (1-x-y)^2(1 - \tfrac{4xy}{(1-x-y)^2}).$$

Observe that $d(t) := d(t,t)$ is just a shifted version of the Catalan generating function,

$$d(t) = t(c(t)-1) = \frac{1}{2}(1 - 2t - \sqrt{1-4t}) = t^2 + 2t^3 + 5t^4 + 14t^5 + 42t^6 + 132t^7 + O(t^8).$$

Incidentally, $d(x,y)$ is the generating function of parallelogram polyominoes with respect to the numbers of columns and rows.

1.2. Monotone paths and k-convexity. In [**CR03**] the authors observed that convex polyominoes have the property that every pair of cells is connected by a *monotone path*. More precisely, a *path* in a polyomino is a self-avoiding sequence of unitary steps of four types: north $N = (0,1)$, south $S = (0,-1)$, east $E = (1,0)$, and west $W = (-1,0)$. A path is *monotone* if it is made with steps of only two types. Given a path $w = u_1 \dots u_k$, with $u_i \in \{N, S, E, W\}$, each pair of steps $u_i u_{i+1}$ such that $u_i \neq u_{i+1}$, $0 < i < k$, is called a *change of direction*. These definitions are illustrated by Fig. 2, in which the non monotone path (a) has 6 changes of direction and the monotone path (b) has 4 changes of direction.

The authors of [**CR03**] further proposed a classification of convex polyominoes based on the number of changes of direction in the paths connecting any two cells of a polyomino. More precisely, a convex polyomino is k-convex if every pair of its cells can be connected by a monotone path with at most k changes of direction. In a convex polyomino of the first level of this classification, any two cells can be connected by a path with at most one change of direction: in view of the L-shape of these paths, 1-convex polyominoes

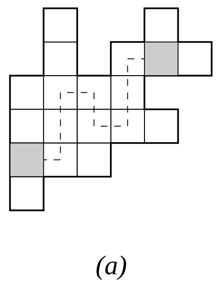


FIGURE 2. (a) a path between two highlighted cells in a polyomino; (b) a monotone path between the same cells, made only of north and east steps.

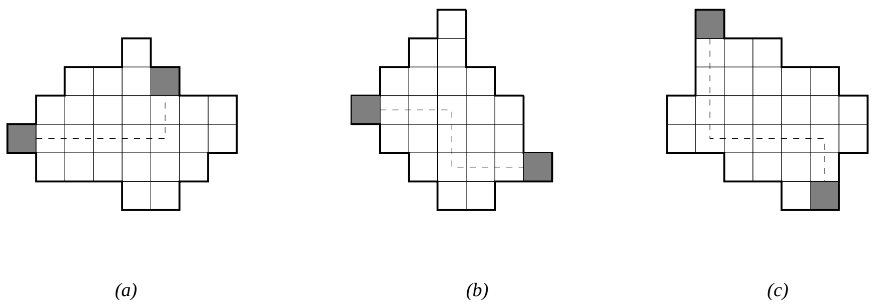


FIGURE 3. (a) a L-convex polyomino, and a monotone path with a single change of direction joining two of its cells; (b) a Z-convex but not L-convex polyomino: the two highlighted cells cannot be connected by a path with only one change of direction; (c) a centered polyomino (not L-convex).

are also called L-convex. The reader can easily check that in Fig. 3, the polyomino (a) is L-convex, while the polyominoes (b), (c) are not, but are 2-convex.

This class of polyominoes has been considered from several points of view: in [**CR05**] it is shown that the set of L-convex polyominoes is well-ordered with respect to the sub-picture order, in [**CFRR1**] the authors have investigated some tomographical aspects of this family, and have shown that L-convex polyominoes are uniquely determined by their horizontal and vertical projections. Finally, in [**CFRR2**] it is proved that the number g_n of L-convex polyominoes with semi-perimeter $n+2$ satisfies the recurrence relation:

$$g_n = 4g_{n-1} - 2g_{n-2}, \qquad n \geq 3,$$

with $g_0 = 1$, $g_1 = 2$, $g_2 = 7$. In other terms the generating function of L-convex polyominoes is rational:

$$\begin{aligned} G(t) &= \sum_{n\geq 0} g_n t^{n+2} = t^2 + 2t^3 + 7t^4 + 24t^5 + 82t^6 + 280t^7 + O(t^8) \\ &= \frac{1-2t+t^2}{1-4t+2t^2}. \end{aligned}$$

Indeed, in [**CFMRR**], the authors have provided an encoding of L-convex polyominoes by words of a regular language, and have furthermore studied the problem of enumerating L-convex polyominoes with respect to the area.

In view of the definition of L-convex polyominoes as 1-convex polyominoes, it is natural to investigate which of the previous properties remain true for some classes of k-convex polyominoes, with $k > 1$. Concerning enumeration in particular, one would like to know if the generating functions of k-convex polyominoes are rational, algebraic, or D-finite.

1.3. Z-convex polyominoes. In the present paper we deal with the family of *2-convex polyominoes*, which we rename Z-convex polyominoes in analogy with the L-convex notation. We shall prove the following results for the number p_n of Z-convex polyominoes with semi-perimeter $n+2$:

Theorem 1.1. *The generating function $P(t)$ of Z-convex polyominoes with respect to the semi-perimeter is*

$$\begin{aligned} P(t) &= \sum_{n\geq 0} p_n t^{n+2} = t^2+2t^3+7t^4+28t^5+116t^6+484t^8+O(t^8), \\ &= \frac{2t^4(1-2t)^2 d(t)}{(1-4t)^2(1-3t)(1-t)} + \frac{t^2(1-6t+10t^2-2t^3-t^4)}{(1-4t)(1-3t)(1-t)}, \end{aligned}$$

where

$$d(t) = \frac{1}{2}(1-2t-\sqrt{1-4t}).$$

More generally, the generating function $P(x,y)$ of Z-convex polyominoes with respect to the numbers of rows and columns is a rational power series $R(x,y,d(x,y))$ in x, y and the unique power series $d(x,y)$ solution of the equation $d=(x+d)(y+d)$,

$$d(x,y) = \frac{1}{2}(1-x-y-\sqrt{\Delta}),$$

where

$$\Delta = (1-x-y)^2-4xy = (1-x-y)^2(1-\frac{4xy}{(1-x-y)^2}).$$

More precisely,

$$\begin{aligned} P(x,y) &= \frac{2x^2y^2d(x,y)}{\Delta^2}\frac{(1-x-y)^2}{((1-x-y)^2-xy)} \\ &\quad + \frac{xy(1-x-y)^2(1-x-y-xy)-x^2y^2(1-x-y-3xy)}{\Delta\,((1-x-y)^2-xy)}. \end{aligned}$$

As conjectured by Marc Noy [**N**], the asymptotic number of Z-convex polyominoes with semi-perimeter $n+2$ grows like $n\cdot 4^n$ (more precisely, $p_n \sim \frac{n}{24}\cdot 4^n$, so that $f_n/p_n \to 3$), while the number of L-convex polyominoes grows only like $(2+\sqrt{2})^n$, and the number of centered polyominoes (see below) grows like 4^n.

The fact that the generating function ends up in the same algebraic extension as convex polyominoes looks surprising to us because we were unable to derive it using the standard approaches to convex polyomino enumeration (Temperley-like methods, wasp-waist decompositions, or inclusion/exclusion on walks). Instead, one interesting feature of our paper is a construction of polyominoes by "inflating" smaller ones along a hook. We believe that this approach could in principle allow for the enumeration of k-convex polyominoes in general.

The rest of the paper is organized as follows. The general strategy of decomposition is explained in Section 2. The different cases are listed and the corresponding relations for generating functions are derived in Section 3. Finally the resulting system of equations is solved in Section 4.

2. Classification and general strategy

In order to present our strategy for the decomposition, we need to distinguish between several types of Z-convex polyominoes.

2.1. Centered polyominoes. The first class we consider is the set $\mathcal{C}$ of *horizontally centered* (or simply *centered*) convex polyominoes. A convex polyomino is said to be centered if it contains at least one row touching both the left and the right side of its minimal bounding rectangle (see Fig. 3 (c)). Observe that centered polyominoes have a simple characterization in terms of monotone paths:

Lemma 2.1. *A convex polyomino is centered if and only if any pair of its cells can be connected by means of a path $S^{h_1}E^kS^{h_2}$ or $S^{h_1}W^kS^{h_2}$, with $h_1,h_2,k\geq 0$.*

In particular any L-convex polyomino is centered, and, more importantly for us, any centered polyomino is Z-convex, while the converse statements do not hold. Figure 3 (c) shows a centered polyomino which is not L-convex, and Figure 3 (b) a Z-convex polyomino which is not centered.

Centered convex polyominoes can also be described as being made of two stack polyominoes glued together at their bases. As we shall see in Section 3.1, this decomposition allows us to compute easily their generating function.

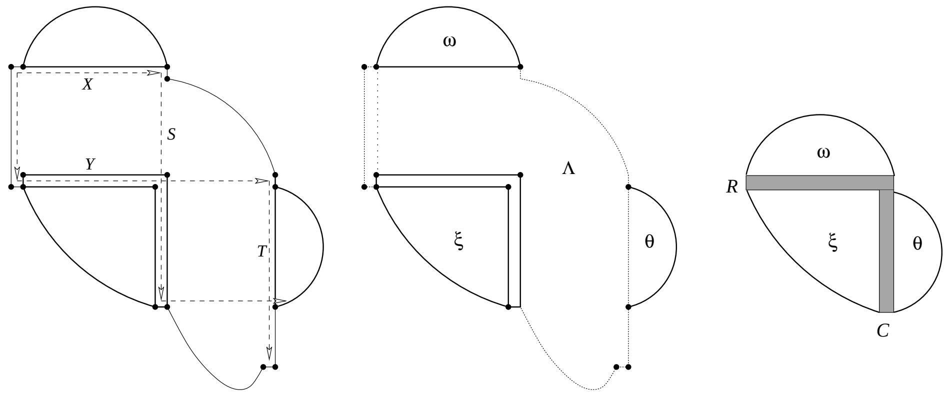


FIGURE 4. A non-centered polyomino of class $\mathcal{D}$ with its rows x, y, and columns s and t; its division into Regions ω, ξ, θ and Λ; its reduction.

2.2. Non centered polyominoes. Let us thus turn to non centered polyominoes. The starting point of our decomposition is that we wish to remove the leftmost column. By definition of Z-convexity, any two cells must be connected by a path of type $S^{h_1}E^kS^{h_2}$, $S^{h_1}W^kS^{h_2}$, $E^{h_1}N^kE^{h_2}$ or $E^{h_1}S^kE^{h_2}$, with $h_1, h_2, k \geq 0$. In particular, we are interested in the set of cells that require a path with two changes of direction to be reached from the cells of the leftmost column.

Let P be a non-centered convex polyomino, and let $c_1(P)$ (briefly, c_1) denote its leftmost column, and let us consider the following rows (as sketched in Fig. 4):

- The row X which contains the top cell of c_1.
- The row Y which contains the bottom cell of c_1.

Since the polyomino P is convex and non-centered, its rightmost column does not intersect any row between X and Y, hence it is placed entirely above X or below Y.

This remark leads to the following definitions:

- A non-centered convex polyomino is *ascending* if its rightmost column is above the row X. Let $\mathcal{U}$ denote the set of descending Z-convex polyominoes.
- A non-centered convex polyomino is *descending* if its rightmost column is below the row Y. Let $\mathcal{D}$ denote the set of ascending Z-convex polyominoes.

The whole set of Z-convex polyominoes is given by the union of the three disjoint sets $\mathcal{C}$, $\mathcal{D}$, and $\mathcal{U}$. Moreover, by symmetry, for any fixed size, $\mathcal{D}$ and $\mathcal{U}$ have the same number of elements, thus, we will only consider non-centered polyominoes of the class $\mathcal{D}$, as the one represented in Fig. 4.

A first property of polyominoes of class $\mathcal{D}$ is the following consequence of their convexity: the boundary path from the end of row X to the end of row Y is made only of south and east steps.

2.3. The strategy. Let us denote by S and T the columns starting from the rightmost cell of X and Y respectively, and running until they reach the bottom of the polyomino (see Fig. 4). The rows and columns X, Y, S and T allow us to separate the cells of any polyomino in class D into four connected sets, as illustrated by Figure 4:

(1) the set of cells strictly above X, called ω;
(2) the set of cells strictly on the right of T, called θ;
(3) the set of cells that are at the same time below Y and on the left of S, called ξ;
(4) the remaining set of cells, called Λ: these cells are either between X and Y, or between S and T (or both).
- In the previous definitions, the hook H starting horizontally with the left hand part of Y and continuing down with the bottom part of S is included in ξ. The other cells of the row X, Y and columns S and T are included in Λ.

The cells of θ require at least two turns to be reached with a monotone path from the cells of c_1. The Z-convexity thus induces a restriction on the position of the lowest cells of θ.

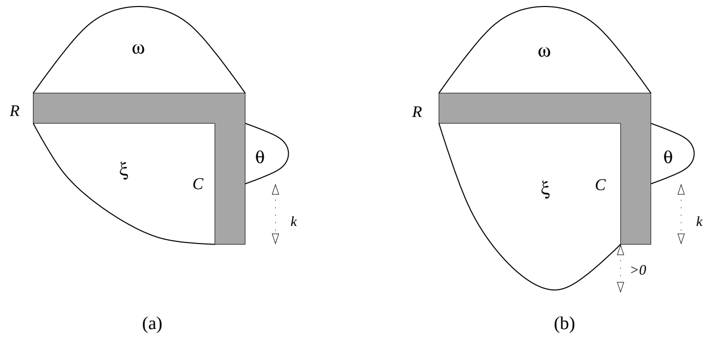


Figure 5. (a) a hooked polyomino with hook of type A and (b) one with hook of type B.

Property 2.1. *The region θ of a non-centered Z-convex polyomino contains no cell lower than the lowest cell of its column S.*

If a row between X and Y reaches the right side of the bounding box, the polyomino is centered:

Property 2.2. *The set θ of a non-centered convex polyomino is non empty.*

As already mentioned, we wish to decompose polyominoes of $\mathcal{D}$ by removing the leftmost column. For the decomposition to be one-to-one we then need to be able to replace a column to the left of a polyomino. But, as the reader can verify, if one takes a Z-convex polyomino and add a leftmost column, it is not so easy to grant *a priori* that Property 2.1 will be satisfied by the rows and columns X, Y, S, and T of the grown polyomino.

In order to circumvent this problem, our decomposition will consist in removing the whole region Λ together with the leftmost column. More precisely, given a descending polyomino P, let us define its *reduction* $\Phi(P)$ as the polyomino obtained as follows (see Figure 4):

- glue region ω to ξ, keeping the relative abscissa of cells between ω and ξ;
- glue region θ to $\omega \cup \xi$ by keeping the relative ordinates of cells between ξ and θ.

Since the hook H is kept in $\Phi(P)$, ω and ξ have at least one common column (as soon as ω is non empty) and ξ and θ have at least one common row, so that the reduction makes sense and it is a polyomino, in which we highlight the hook H. (The hook is highlighted in order to make easier the forthcoming description of the inverse construction.)

The following lemma explains our interest in this reduction.

Lemma 2.2. *A descending convex polyomino is Z-convex if and only if it satisfies Property 2.1 and its reduction $\Phi(P)$ is Z-convex.*

Proof. Assume first that P is Z-convex. Then Property 2.1 is satisfied and a monotone path connecting a cell x to a cell y of $\Phi(P)$ can easily be constructed from the monotone path connecting x and y in P: any section of the path in the deleted region Λ can be replaced by a simpler section in the hook.

Conversely assume that $\Phi(P)$ is Z-convex, that P satisfies Property 2.1 is satisfied, and let (x, y) be two cells of P. If x and y are not in Λ then there exists a monotone path in $\Phi(P)$ connecting these points, and there is no need to add a turn to extend this path into a monotone path in P. If x belongs to Λ, one easily construct the path in each case $y \in \omega$, $y \in \xi$ and, using Lemma 2.1, $u \in \theta$. □

To characterize the set of polyominoes that can occur in the image of $\mathcal{D}$ by Φ, let us define a *hooked polyomino* as a polyomino P of $\mathcal{C} \cup \mathcal{D}$ in which a hook is highlighted, in such a way that

- the hook is made of a top row (the *arm* of hook) starting in the leftmost column of P and traversing the polyomino, and a partial column (the *leg* of the hook) starting in the right most cell of the top row and including all cells below in this column,
- the region on the right hand side of the hook is non empty.

The hook is called a *hook of type A* if its bottom cell belongs to the lowest row of P, and a *hook of type B* otherwise (see Figure 5). The following lemma is an immediate consequence of the definition of hooked polyominoes.

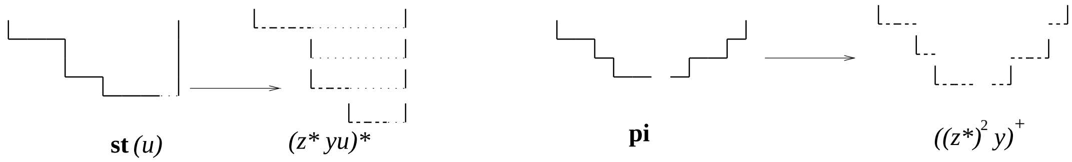


FIGURE 6. The decomposition of a staircase $\mathbf{st}(u) = (z^*yu)^*$ and of a non empty pile $\mathbf{pi} = ((z^*)^2y)^+$.

PROPERTY 2.3. *The reduction of a polyomino of $\mathcal{D}$ is a hooked polyomino.*

In view of Lemma 2.2, our strategy will consist in the description of the types of region Λ that can be added to a hooked polyomino so that the "inflated" polyomino satisfies Property 2.1.

2.4. Generating functions. We shall compute the generating function $P(x, y)$ of Z-convex polyominoes with respect to the number of columns, or *width* (variable x) and to the number of rows, or *height* (variable y). In order to do that we shall need generating functions of hooked polyominoes with respect to their width and their height, but also with respect to an auxiliary parameter k which will be marked by a variable u: given a hooked polyomino, the parameter k is a non negative integer indicating the difference of ordinate between the lowest cell of the leg of the hook and the lowest cell of the next column to the right (see Figure 5). This definition makes sense since the region on the right hand side of the hook is assumed non empty.

We shall more precisely use the generating functions

- $C_A(x, y, u)$ of hooked centered polyominoes with hook of type A,
- $C_B(x, y, u)$ of hooked centered polyominoes with hook of type B,
- $A(x, y, u) = \sum_k a_k(x, y)u^k$ of hooked non-centered polyominoes with hook of type A,
- $B(x, y, u) = \sum_k b_k(x, y)u^k$ of hooked non-centered polyominoes with hook of type B.

Most of the time we drop the variables x, y and use the shorthand notation $A(u) = A(x, y, u)$, $a_k = a_k(x, y)$, etc.

3. Decompositions

We shall need the following elementary notations and results, illustrated by Figure 6:

- The sequence notation for formal power series is $w^* = \frac{1}{1-w}$. The non-empty sequence notation is $w^+ = w \cdot w^* = \frac{w}{1-w}$.
- The generating function of possibly empty staircases with width marked by z and height marked by yu is $\mathbf{st}(u) = (z^*yu)^*$.
- The generating function of non empty piles of lines with width marked by z and height marked by y is $\mathbf{pi} = ((z^*)^2y)^+$.
- Given a generating function $F(u) = \sum_{n\geq 0} f_n u^n$ we define the series

$$F(u, v) = \sum_{n\geq 0} \sum_{i+j=n} f_n u^i v^j = \frac{uF(u)}{u - v} + \frac{vF(v)}{v - u},$$

  where the inner summation is on non negative i and j with $i + j = n$, and

$$F(u, v, w) = \sum_{n\geq 0} \sum_{i+j+k=n} f_n u^i v^j w^k = \frac{u^2F(u)}{(u - v)(u - w)} + \frac{v^2F(v)}{(v - u)(w - u)} + \frac{w^2F(w)}{(w - u)(w - v)},$$

  where the inner summation is on non negative i, j and k with $i + j + k = n$. The values at $u = v$ of the previous series can be obtained by continuity:

$$A(u, u, w) = \left(\frac{u^2A(u)}{u - w}\right)' + \frac{w^2A(w)}{(w - u)^2} = \frac{u^2A'(u)}{u - w} + \frac{u(u - 2w)A(u)}{(u - w)^2} + \frac{w^2A(w)}{(w - u)^2}.$$

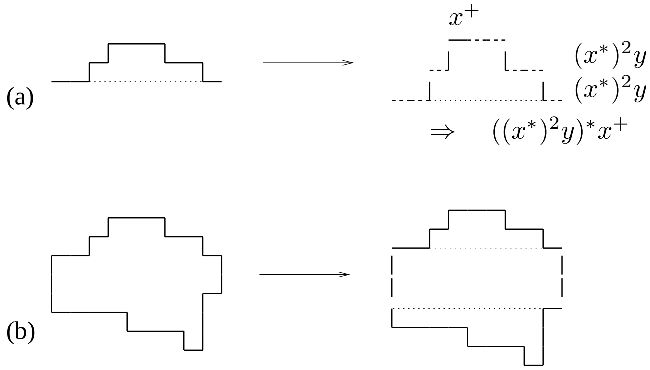


FIGURE 7. (a) The decomposition of a stack polyomino with baseline width marked by x and height marked by y. (b) The decomposition of a centered polyomino into a non-empty sequence of central rows and two stack polyominoes.

3.1. Centered polyominoes. Recall that a centered polyomino is a polyomino that contains at least one row touching both the left and the right hand side of its minimal bounding rectangle. We need to count polyominoes of the family $\mathcal{C}$ of centered polyominoes but also of the families $\mathcal{C}_A$ and $\mathcal{C}_B$ of hooked polyominoes with a hook of type A and B respectively.

Let $S(x, y)$ be the generating function of stack polyominoes with x marking the length of the baseline and y marking the height. In view of Figure 7(a),

$$S(x,y) = x^+ \cdot ((z^*)^2y)^* = \frac{x(1-x)}{(1-x)^2-y} = \left(\frac{1}{2}\frac{1}{1-\frac{x}{1-\sqrt{y}}} + \frac{1}{2}\frac{1}{1-\frac{x}{1+\sqrt{y}}} - 1\right).$$

Observe then that for any power series $F(x, y)$ the Hadamard product $S(x, y) \odot_x F(x, y)$ is equal to:

$$\frac{1}{2}\left(F\big(\frac{x}{1-\sqrt{y}}, y\big) + F\big(\frac{x}{1+\sqrt{y}}, y\big)\right) - F(0),$$

which is a rational function of x and y if $F(x)$ is.

In view of Figure 7(b), centered polyominoes are formed of a centered rectangle supporting 2 strictly smaller stacks polyominoes:

$$C(x,y) = y^+[S^>(x,y) \odot_x S^>(x,y)],$$

where $S^>(x, y)$ stands for the generating function of stack polyominoes with a first row strictly smaller than the baseline (so that the central rectangle is effectively given by the factor y^+). The series $S^>(x, y)$ is readily obtain by difference,

$$S^>(x,y) = S(x,y) - yS(x,y) = \frac{x(1-x)(1-y)}{(1-x)^2-y},$$

and computing the Hadamard product with the previous formula yields:

$$C(x,y) = \frac{xy(-y-xy+1-2x+x^2)(1-y)}{(1-x-y)(x^2-2xy-2x+y^2-2y+1)}.$$

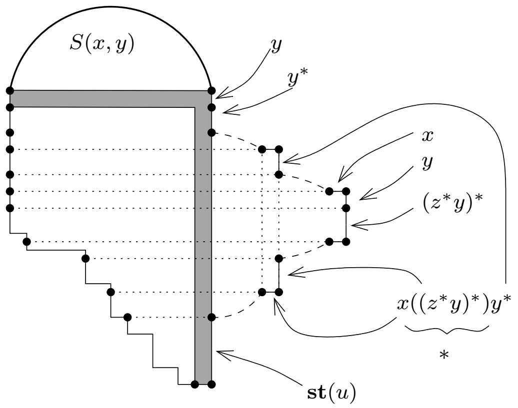
We shall also need centered polyominoes with a marked hook, of type A and B. As illustrated by Figure 8, the series for the first type is $C_A(u) = S(z, y) \odot_z F_A(zx, x, y, u)$ where

$$F_A(z,x,y,u) = y^2z^+ \cdot \mathbf{st}(u) \cdot (xy^*(yz^*)^*)^+.$$

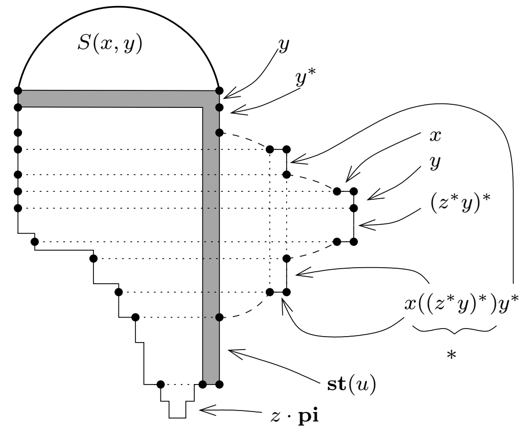
Indeed, with z marking columns on the left hand side of the leg of the hook, the Hadamard product accounts for gluying, along the arm of the hook, a staircase, with generating series $S(z, y)$, to the rest of the polyomino, with generating series $F_A(z, x, y, u)$: in this later series, a factor $xy(z^*y)^*$ corresponds to the central rectangle; each factor $x((z^*y)^*)y^*$ corresponds to a column on the right of the hook and to the lines having their rightmost cell in that column; the factor $\mathbf{st}(u)$ corresponds to the bottom staircase made of lines having their rightmost cell in the hook.

Similarly, the series for the second type is $C_B(u) = S(z, y) \odot_z F_B(zx, x, y, u)$ where

$$F_B(z,x,y,u) = F_A(z,x,y,u) \cdot z \cdot \mathbf{pi},$$



$F_A(u) = xy^2y^*(z^*y)^*((x(z^*y)^*)y^*)^* \cdot \mathbf{st}(u)$



$F_B(u) = xy^2y^*(z^*y)^*((x(z^*y)^*)y^*)^* \cdot \mathbf{st}(u) \cdot z \cdot \mathbf{pi}$

FIGURE 8. Construction of elements of the classes $C_A(u)$ and $C_B(u)$.

with the extra factor corresponding to cells lower than the leg of the hook.

3.2. Hooked polyominoes with hooks of type A. A hooked polyomino with hook ok type A can be a hooked centered polyomino (with gf $C_A(u)$ already computed) or can be obtained from its reduction which must be a hooked polyomino with a hook of type A (recall that type A means that the leg of the hook reaches the lowest row of the polyomino). Let us describe the different cases, with respect to the properties of the resulting inflated polyomino:

- The leg of the hook and the two columns S and T have same abscissa (Figure 9, left): let
$$A_1(u) = x(y^*)^2 \cdot A(u).$$
- The leg of the hook has same abscissa as the column T but not as S (Figure 9, middle): by definition of type A, the column S cannot be longer than the leg of the hook, and
$$A_2(u) = xy^*y^+ \cdot \mathbf{st}(u) \cdot z^+A(u).$$
The series $A_2(u)$ apparently does not takes into account the construction of the staircase starting on the righthand side of the column S and connecting it to the top-right angle of the hook. Instead each column between column S (excluded) and the leg of the hook (included) is marked by a factor z. However upon setting $z = xy^*$, each column marked by z gets a factor x and a factor y^* that accounts for the rows ending in that column. The generating function of polyominoes of this case is thus $A_2(u)\big|_{z=xy^*}$.

  In all forthcoming cases, we describe similarly generating functions of polyominoes without the staircase connecting X to the top-right corner of the hook. In other terms, in the following pages z is to be understood as a shorthand notation for xy^*.
- The leg of the hook has same abscissa as the column S but not as T (Figure 9, right): by definition of type A, the leg of the hook is at least as long as the column T, and
$$A_3(u) = xy^*y^+ \cdot z^+A(u, z^*).$$
As suggested by Figure 9, the factor $A(u, z^*)$ accounts for the fact that from a polyomino of type A with parameter k, one constructs a new polyomino with parameter i with $0 \le i \le k$.
- The abscissa of the leg of the hook is strictly between S and T and the cells marked by a factor u are strictly below the lowest cells of columns S and T (Figure 10, left):
$$A_4(u) = x(y^+)^2 \cdot \mathbf{st}(u) \cdot \mathbf{pi} \cdot (z^+)^2A(z^*).$$
- The abscissa of the leg of the hook is strictly between S and T and the cells marked by a factor u intersects the baseline of column S (Figure 10, right):
$$A_5(u) = x(y^+)^2 \cdot \mathbf{st}(u) \cdot (z^+)^2A(u, z^*).$$

The generating series of hooked polyominoes with a hook of type A is then

$$A(u) = C_A(u) + \sum_{i=1}^{5} A_i(u)\big|_{z=xy^*}.$$

3.3. Hooked polyominoes with hooks of type B. A hooked polyomino with hook of type B can be a hooked centered polyomino (with gf $C_B(u)$ already computed) or can be obtained by inflating its reduced polyomino. We start with those that are produced from a hooked polyomino with hook of type A, and we give again the different cases with respect to the properties of the obtained inflated polyomino:

- The leg of the hook has same abscissa as column T and it is strictly longer than column S: in order to produce a hook of type B, some other column between S and T must be even longer, and

$$B_1(u) = A_2(u) \cdot z \cdot \mathbf{pi} = xy^*y^+ \cdot \mathbf{pi} \cdot z \cdot \mathbf{st}(u) \cdot z^*yu \cdot z^+ A(u).$$

  In agreement with the relation $B_1(u) = A_2(u) \cdot z \cdot \mathbf{pi}$, polyominoes of B_1 (Figure 11, left) can be obtained from polyominoes of A_2 (Figure 9, middle) by adding a non-empty pile of rows just before the leg of the hook.
- The abscissa of the leg of the hook is strictly between S and T, and the cells marked by a factor u are strictly below the lowest cells of columns S and T:

$$B_2(u) = A_4(u) \cdot z \cdot \mathbf{pi} = x(y^+)^2 \cdot \mathbf{pi} \cdot z \cdot \mathbf{st}(u) \cdot \mathbf{pi} \cdot (z^+)^2 A(z^*).$$

  These polyominoes are obtained from the polyominoes of A_4 upon adding a non-empty pile of rows just before the leg of the hook.
- The abscissa of the leg of the hook is strictly between S and T, and the cells marked by a factor u intersects the baseline of column S.

$$B_3(u) = A_5(u) \cdot z \cdot \mathbf{pi} = x(y^+)^2 \cdot \mathbf{pi} \cdot z \cdot \mathbf{st}(u) \cdot (z^+)^2 A(u, z^*).$$

  These polyominoes are obtained from the polyominoes of A_5 upon adding a non-empty pile of rows just before the leg of the hook.
- The abscissa of the leg of the hook is strictly between S and T, and the cells marked by a factor u are strictly above the lowest cell of column S.

$$B_4(u) = x(y^+)^2 \cdot (1 + \mathbf{pi} \cdot z)(z^+)^2 (A(z^*, z^*, u) - A(z^*, u)).$$

  Observe that difference is due to the restriction $j \neq 0$, as illustrated by the Figure 11: the leg of the hook must end strictly above the lowest cell of columns S, so that one must have $j \geq 1$.

Now we present the cases produced from a hooked polyomino with hook of type B, again arranged according to the properties of the resulting polyomino. Observe that in these cases the column S is at least as long as the column T:

- The leg of the hook and the columns S and T have the same abscissa:

$$B_5(u) = x(y^*)^2 \cdot B(u)$$

- The leg of the hook has the same abscissa as the column S or the same abscissa as the column T (but not both):

$$B_6(u) = 2 \cdot xy^*y^+ \cdot z^* B(z^*, u).$$

- The abscissa of the leg of the hook is strictly between column S and T:

$$B_7(u) = x(y^+)^2 \cdot (z^+)^2 B(z^*, z^*, u).$$

The generating series of hooked polyominoes with a hook of type B is then

$$B(u) = C_B(u) + \sum_{i=1}^{7} B_i(u)\big|_{z=xy^*}.$$

3.4. Z-convex polyominoes. Again we start with polyominoes that are produced from a hooked polyomino with hook of type A:

- The columns S and T have the same abscissa:
$$P_1 = xy^* A(1).$$
- The columns S and T have distinct abscissa and the column S is strictly shorter than T:
$$P_2 = xy^+ \cdot \mathbf{pi} \cdot zA(1).$$
- The columns S and T have distinct abscissa and the column S is at least as long as T:
$$P_3 = xy^+ \cdot (1 + z \cdot \mathbf{pi})(z^* A(z^*) - A(1)).$$
The difference is due to the fact that at least one horizontal column must be inserted at the level of the rows that were marked by the factor u to ensure that the column T is not longer that S.

Next we present the polyominoes obtained from a hooked polyomino with hook of type B:

- The columns S and T have the same abscissa:
$$P_4 = xy^* B(1).$$
- The columns S and T have distinct abscissa and the lowest cell of S is below or at the same level as the lowest cell of T:
$$P_5 = xy^+ \cdot (z^* B(z^*) - B(1)).$$

Finally the generating function of Z-convex polyominoes is
$$P = C + \sum_{i=1}^{5} P_i\big|_{z=xy^*}.$$

4. Resolution

In view of the previous section, upon setting as announced $z = xy^*$, the system of equations defining the series P has the following form:
$$\begin{aligned} A(u) &= C_A(u) + a_1(u)A(u) + a_2(u)A(z^*) \\ B(u) &= C_B(u) + b_1(u)A(u) + b_2(u)A(z^*) + b_3(u)A'_u(z^*) + b_4(u)B(u) + b_5(u)B(z^*) + b_6(u)B'_u(z^*) \\ P &= C_0 + p_1 A(z^*) + p_2 A(1) + p_3 B(z^*) + p_4 B(1), \end{aligned}$$
where the $C_A(u)$, $C_B(u)$ and C_0 are the rational generating series of centered polyominoes computed in Section 3.1, the $a_i(u)$ and $b_i(u)$ are explicit rational functions of x, y and u, and the p_i are explicit rational functions of x and y.

The first step of the resolution is to apply the kernel method to the first equation, which involves only $A(u)$ and $A(z^*)$ as unknown. The kernel equation $1 - a_1(u) = 0$ contains a factor that can be written
$$u = 1 + (x - y)u + yu^2,$$
so that it clearly admits a power series root $c(x, y)$, which is a refinement of the Catalan generating function
$$c(t, t) = \frac{1 - \sqrt{1 - 4t}}{2t}.$$
Setting $u = c(x, y)$ in the first equation, the kernel is canceled and $A(z^*)$ is obtained as
$$A(z^*) = \frac{C_A(c)}{a_2(c)}.$$
Then, using again the first equation of the system we derive $A(u)$. Once $A(u)$ is known, $A'_u(z^*)$ can also be computed.

The second step consists in applying now the kernel method to the second equation of the system, which now has three unknowns $B(u)$, $B(z^*)$ and $B'_u(z^*)$. The kernel $1 - b_4(u)$ admits two roots R_1 and R_2 that are rational power series in $x^{1/2}$ and y:
$$R_1 = \frac{1 - y + x^{1/2}}{1 - y - x + (1 - x)x^{1/2}}, \quad \text{and} \quad R_2 = \frac{1 - y - x^{1/2}}{1 - y - x - (1 - x)x^{1/2}}.$$

Using these two roots we write two linear equations for $B(z^*)$ and $B'_u(z^*)$ and solve the system. The resulting series are rational series in x, y and $C(x, y)$ (in particular fractional powers of x cancel, as one could expect from the symmetry with respect to $\pm x^{1/2}$). Returning to the second equation of the system, we obtain $B(u)$ and finally, turning to the third equation, the generating function P of Z-convex polyominoes.

It should be remarked that our method leads to heavy computations in the intermediary steps, involving big rational expressions. The fact that things dramatically simplify when all pieces are put together in P calls for a simpler, more combinatorial, derivation. In particular, the expression are nicer in terms of the more symmetric parametrization $d(x, y) = y(c(x, y) - 1)$ satisfying $d = (x + d)(y + d)$.

Acknowledgments. The authors wish to thank Andrea Frosini and Marc Noy for inspiring discussions on the topic of this paper. Andrea Frosini also provided the first terms of the series by exhaustive generation, allowing us to double check our results.

References

[BMRR] G. Barequet, M. Moffie, A. Rib and G. Rote, Counting Polyominoes on Twisted Cylinders, Proceedings of *2005 European Conference on Combinatorics, Graph Theory and Applications (EuroComb '05)*, Ed. S. Felsner, Discrete Mathematics and Theoretical Computer Science Proceedings AE, pp. 369-374

[BN] D. Beauquier, M. Nivat, Tiling the plane with one tile, In *Proc. of the 6th Annual Symposium on Computational geometry (SGC'90)* (Berkeley, CA, 1990), ACM press (1990) 128–138.

[BM] M. Bousquet-Mélou, Habilitation à diriger les recherches, Université Bordeaux 1, 1996.

[CD] M. Chrobak, C. Durr, Reconstructing hv-Convex Polyominoes from Orthogonal Projections, *Information Processing Letters*, 69 283-289 (1999).

[CFRR1] G. Castiglione, A. Frosini, A. Restivo, S. Rinaldi, A tomographical characterization of L-convex polyominoes, Proceedings of *Discrete Geometry for Computer Imagery 12th International Conference, DGCI 2005*, Eds. E. Andres, G. Damiand, P. Lienhard, Lecture Notes in Computer Science, Poitiers 115-125.

[CFRR2] G. Castiglione, A. Frosini, A. Restivo, S. Rinaldi, Enumeration of L-convex polyominoes by rows and columns, to appear on *Theor. Comp. Sci.*.

[CFMRR] G. Castiglione, A. Frosini, E. Munarini, A. Restivo, S. Rinaldi, Enumeration of L-convex polyominoes II. Bijection and area, *Proceedings of Formal Power Series and Algebraic Combinatorics*, 20-26 June 2005, Taormina (Italy) # 49.

[CR03] G. Castiglione, A. Restivo, Reconstruction of L-convex Polyominoes, *Electronic Notes in Disc. Math.* 12 Elsevier Science (2003).

[CR05] G. Castiglione, A. Restivo, Ordering and Convex Polyominoes, *Machines, Computations, and Universality, 4th International Conference, MCU 2004*, M.Margenstern Eds., Saint Petersburg, Russia, Lecture Notes in Computer Science 3354 Springer (2005).

[DV] M. Delest, X. Viennot, Algebraic languages and polyominoes enumeration, *Theor. Comp. Sci.*, 34 169–206 (1984).

[Ga] M. Gardner, *Mathematical games*, Scientific American, (1958) Sept. 182–192, Nov. 136–142.

[G1] S. W. Golomb, *Polyominoes: Puzzles, Patterns, Problems, and Packings,* Princeton Academic Press, 1996.

[G2] S. W. Golomb, Checker boards and polyominoes, *Amer. Math. Monthly* **61**, n. 10 (1954) 675–682.

[JG] I. Jensen, A. J. Guttmann, Statistics of lattice animals (polyominoes) and polygons, *J. Phys. A:Math. Gen.*, 33, (2000), 257–263.

[N] M. Noy, Some comments on 2-convex polyominoes, *Personal communication.*

[Sl] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences.
Published electronically at `http://www.research.att.com/~njas/sequences/`

[S] R. P. Stanley, *Enumerative Combinatorics,* Vol.2, Cambridge University Press, Cambridge (1999).

Appendix

Figures 9, 10, 11, 12, 13, and 14 appear in the full version of the paper available as `arXiv:math.CO/0602124`.

Université Paris 7
E-mail address: `duchi@liafa.jussieu.fr`

Università di Siena
E-mail address: `rinaldi@unisi.it`

CNRS, École Polytechnique
E-mail address: `schaeffe@lix.polytechnique.fr`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A Spectral Approach to Pattern-Avoiding Permutations

Richard Ehrenborg, Sergey Kitaev, and Peter Perry

Abstract. We study the number of permutations in the symmetric group on n elements that avoid consecutive patterns S. We show that the spectrum of an associated integral operator on the space $L^2[0,1]^m$ determines the asymptotic behavior of such permutations. Moreover, using an operator version of the classical Frobenius-Perron theorem due to Kreĭn and Rutman, we prove asymptotic results for large classes of patterns S. This extends previously known results of Elizalde.

Résumé. Nous étudions le nombre de permutations dans le groupe symétrique sur n éléments qui évitent des motifs S consécutifs. Nous montrons que le spectre d'un opérateur intégral associé sur $L^2[0,1]^m$ détermine le comportement asymptotique de telles permutations. Utilisant de plus une version d'opérateur du théorme classique de Frobenius-Perron en raison de Kreĭn et Rutman, nous donnons des résultats asymptotiques pour les grandes classes de motifs S. Ceci étend résultats précédemment des connus de Elizalde.

1. Introduction

In this paper, we study integral operators of the form

$$T : L^2\left([0,1]^m\right) \longrightarrow L^2\left([0,1]^m\right) \tag{1.1}$$
$$f \longmapsto \int_0^1 \chi(t, x_1, \ldots, x_m) f(t, x_1, \ldots, x_{m-1})\, dt$$

and their applications to the theory of pattern avoidance in permutations. Here χ is a real-valued function on $[0,1]^{m+1}$ which takes the values 0 or 1 on each of the simplices in the standard triangulation of $[0,1]^{m+1}$, i.e., the partition

$$[0,1]^k = \bigcup_{\pi \in \mathfrak{S}_k} \Delta_\pi$$

where the simplex Δ_π is given by

$$\Delta_\pi = \left\{(x_1, \ldots, x_k) : x_{\pi^{-1}(1)} \leq x_{\pi^{-1}(2)} \leq \cdots \leq x_{\pi^{-1}(k)}\right\}$$

We will show how integral operators of this type arise naturally in counting pattern-avoiding permutations where the pattern has length $m+1$.

Recall that a *pattern* of length $m+1$ is an element $\sigma \in \mathfrak{S}_{m+1}$. A permutation $\pi \in \mathfrak{S}_n$, $n \geq m+1$, avoids the consecutive pattern σ if there is no integer j, $0 \leq j \leq n-m-1$, with the property that $\pi_{j+\sigma^{-1}(1)} < \pi_{j+\sigma^{-1}(2)} < \cdots < \pi_{j+\sigma^{-1}(m+1)}$. More generally, if S is a subset of $\mathfrak{S}_{m+1}$, we say that π avoids S if π avoids each $\sigma \in S$.

Fix a subset S of $\mathfrak{S}_{m+1}$ and, for $n \geq m+1$, let a_n denote the number of permutations $\pi \in \mathfrak{S}_n$ that avoid S. Let χ_S: $[0,1]^{m+1} \to \{0,1\}$ be given by

$$\chi_S(x_1, \ldots, x_{m+1}) = \begin{cases} 0 \text{ if } x_{\sigma^{-1}(1)} \leq x_{\sigma^{-1}(2)} \leq \cdots \leq x_{\sigma^{-1}(m+1)} \text{ for some } \sigma \in S; \\ 1 \text{ otherwise.} \end{cases} \tag{1.2}$$

2000 *Mathematics Subject Classification.* Primary 05A16, 05D40, 45C05, 45P05, 47N30; Secondary 05A05.
Key words and phrases. Consecutive pattern avoidance, integral operators, eigenvalues, eigenfunction expansion.
Ehrenborg supported in part by NSF Grant DMS-0200624.
Perry supported in part by NSF Grant DMS-0408419.

Let T_S be the integral operator on $L^2\left([0,1]^m\right)$ given by

$$(T_S f)(x_1,\dots,x_m) = \int_0^1 \chi_S(t,x_1,\dots,x_m) f(t,x_1,\dots,x_{m-1})\,dt. \tag{1.3}$$

THEOREM 1.1. *The formula*

$$\frac{a_n}{n!} = (1, T_S^{n-m} 1) \tag{1.4}$$

holds for any $n \geq m+1$, where 1 *denotes the constant function with value* 1 *and* $(\,\cdot\,,\,\cdot\,)$ *denotes the usual inner product on $L^2\left([0,1]^m\right)$. Moreover, we have the inequality*

$$\frac{a_n}{n!} \leq C_S \left(\frac{a_{2m}}{(2m)!}\right)^{n/m}. \tag{1.5}$$

The inequality (1.5) is not optimal.

It is natural to attempt a large-n asymptotic expansion of the right-hand side of (1.4) using the spectral theory of the operator T_S. Recall that, if A is a bounded operator, the *resolvent set* of the operator A is the set $\rho(A)$ of complex numbers with the property that $(A - zI)^{-1}$ exists as a bounded operator. The *spectrum* of A is the set $\sigma(A)$ is the complement of $\rho(A)$ in $\mathbb{C}$. The *spectral radius* of a bounded operator A is the quantity

$$r(A) = \limsup_{n\to\infty} \|A^n\|^{1/n}.$$

Note that $\sigma(A)$ is contained in the closed disc of radius $r(A)$ about 0 in $\mathbb{C}$ (see, for example, §VI.3 of [**12**]).

THEOREM 1.2. *Let T_S be an integral operator of the form (1.3) for some nonempty pattern S. Then $\sigma(T_S)$ is a discrete set with* 0 *as its only possible accumulation point. Moreover, $r(T_S) < 1$ strictly.*

The proof uses the fact that, although T_S is not compact, the operator T_S^k is compact–in fact Hilbert-Schmidt–for any $k \geq m$. We will show that the Hilbert-Schmidt norm of T_S^m is strictly less than 1, from which the statement about the spectral radius follows. We will also give examples of sets of patterns S for which $r(T_S) = 0$, and the ratio $a_n/n!$ converges to zero as $n \to \infty$.

Our main interest is in patterns for which $r(T_S) > 0$. With an additional condition on S, we can use spectral theory to obtain an asymptotic formula for a_n. Below (Theorem 1.5), we will give a sufficient condition on a pattern S so that the hypotheses of Theorem 1.3 hold. To state this condition, recall that an operator A on the space $L^2(X,\mu)$ of complex-valued measurable functions on the measure space (X,μ) is called *strongly positive* if for every $f \geq 0$ there is an integer n so that $(T^n f)(x) > 0$ for almost every x. As we show through examples below, there are patterns S for which T_S is not strongly positive.

THEOREM 1.3. *Suppose that T_S is an operator of the form (1.3) for some set of patterns S, and that T_S is strongly positive. Then T_S has a unique simple eigenvalue $\rho > 0$ with positive eigenfunction ϕ, and all other eigenvalues $\lambda \in \sigma(T_S)$ satisfy $|\lambda| < \rho$ strictly. Moreover, the adjoint operator T_S^* has ρ as its unique positive eigenvalue and a positive eigenfunction ψ of T_S^* with eigenvalue ρ.*

It is important to note that the strong positivity of T_S implies that T_S has nonzero spectral radius, *and* that the positive eigenvalue is the only eigenvalue on the circle $|z| = \rho$. The existence of such a "spectral gap" and the associated positive eigenfunctions follows from an operator version of the celebrated Perron-Frobenius Theorem (see, e.g., Gantmacher [**8**], vol. 2, §XIII.2) due to Kreĭn and Rutman (see Theorem 6.3 of [**10**]). Under the assumption of Theorem 1.3, let

$$r_2(T_S) = \sup_{\lambda\in\sigma(T_S),\lambda\neq\rho} |\lambda| \tag{1.6}$$

Using spectral theory, we obtain:

THEOREM 1.4. *Suppose that T_S is a strongly positive operator of the form (1.3). Let ρ be the largest eigenvalue of T_S with associated eigenfunction ϕ. Let ψ be the eigenfunction of the adjoint operator T_S^* with eigenvalue ρ. Finally, let r_2 be given by equation (1.6). Then we have*

$$\frac{a_n}{n!} = \rho^{n-m}\frac{(\psi,1)(1,\phi)}{(\psi,\phi)} + \mathcal{O}(r_2^{n-m}).$$

Here $(\,\cdot\,,\,\cdot\,)$ denotes the usual inner product on $L^2\left([0,1]^m\right)$. Note that the leading term in this expansion is strictly positive since ϕ and ψ are positive functions of T_S and T_S^*. Higher-order terms in the expansion can be computed if further eigenvalues and eigenfunctions of the operator T_S are known (see, for example, Section 3 in what follows); see Section 2.3 for a statement of the full expansion.

We can give a sufficient condition in combinatorial terms for a pattern S to have a spectral gap in the sense of Theorem 1.3. To do so we associate to a pattern S a directed graph, G_S, defined as follows. If $x \in \mathbb{Z}^m$ is a vector of positive integers define $\Pi(x)$ to be the permutation $\pi \in \mathfrak{S}_m$ with the property that $x_i < x_j$ if and only if $\pi(i) < \pi(j)$ for all $1 \leq i < j \leq m$. The vertices of G_S are the elements of $\mathfrak{S}_m$, and the edge $\sigma = (\sigma_1, \ldots, \sigma_{m+1}) \in \mathfrak{S}_{m+1} - S$ goes from the permutation $\Pi(\sigma_1, \ldots, \sigma_m)$ to the permutation $\Pi(\sigma_2, \ldots, \sigma_{m+1})$. The graph G_S is *strongly connected* if any point of G_S is connected to any other point of G_S by a directed path. A strongly connected graph is *aperiodic* if there exists a positive integer k and two vertices u and v such that there exists a directed path from u to v of any length greater than or equal to k. The condition that two such vertices exist is equivalent to the statement that between any two vertices in the graph, one can find directed paths of any length greater than or equal to k.

THEOREM 1.5. *Let $S \subset \mathfrak{S}_{m+1}$ and suppose that G_S is strongly connected and the two monotone permutations $12\cdots m+1$ and $m+1\cdots 21$ do not belong to the set S. Then T_S is strongly positive. Hence we conclude that there exist three positive constants ρ, r_2 and c such that $r_2 < \rho$ and*

$$\frac{a_n}{n!} = c\rho^{n-m} + \mathcal{O}(r_2^{n-m}).$$

EXAMPLE 1.6. Let S be the set $\{132, 231\}$. Hence, S-avoiding permutations are permutations without a peak, and there are 2^{n-1} such permutations in $\mathfrak{S}_n$. In this case, the operator T_S has no eigenvalues and our spectral methods do not apply. Also, observe that the graph G_S is not strongly connected, so Theorem 1.5 does not apply.

EXAMPLE 1.7. Let S be the set $\{123, 213, 231, 321\}$. The directed graph G_S is strongly connected, but not aperiodic. Again Theorem 1.5 does not apply. In fact, in this case, $a_n = 2$ for all $n \geq 2$.

In many cases of interest the leading term is explicitly computable. Using Theorem 1.4, we will prove the following asymptotic formulas.

THEOREM 1.8. *The number a_n of 123-avoiding permutations in $\mathfrak{S}_n$ obeys the asymptotic formula*

$$\frac{a_n}{n!} = \lambda_0^{n+1} \exp\left(\frac{1}{2\lambda_0}\right) + \mathcal{O}\left(\lambda_{-1}^n\right)$$

where

$$\lambda_0 = \frac{3\sqrt{3}}{2\pi}, \quad \lambda_{-1} = \frac{3\sqrt{3}}{4\pi}\ .$$

In this case, all of the eigenvalues of T_S are real and T_S has empty kernel. We can easily obtain higher-order terms in the expansion from the spectral methods used there since, in fact, all of the eigenvalues and eigenfunctions of the operator T_S and its adjoint can be computed explicitly: see Theorem 3.3.

We also have the following result for 213-avoiding permutations.

THEOREM 1.9. *The number b_n of 213-avoiding permutations in the symmetric group $\mathfrak{S}_n$ obeys the asymptotic formula*

$$b_n/n! = \exp\left(\frac{1}{2\lambda_0^2}\right) \cdot \lambda_0^{n+1} + \mathcal{O}\left(\left(\frac{1}{\sqrt{2}}\right)^{n-2}\right),$$

where $\lambda_0 = 0.7839769312\ldots$ is the unique real root to the equation

$$\operatorname{erf}\left(\frac{1}{\lambda\sqrt{2}}\right) = \sqrt{\frac{2}{\pi}}\ .$$

In this case, the other eigenvalues of T_S are not real and the kernel of T_S has infinite dimension.

We close our introduction by a brief overview on the subject of pattern avoidance in permutations (for more details we refer to [**2**]). The "classical" definition of a pattern is slightly different than one provided above. We say that a permutation π avoids a pattern σ if π does not contain a *subsequence* which is order-isomorphic to σ. The study of such patterns originated in theoretical computer science by Donald Knuth [**9**].

However, the first systematic study was done by Simon and Schmidt [**13**], who completely classified the avoidance of patterns of length three. Since then several hundred papers related to the field have been published.

One of the most important results in the subject is the proof by Marcus and Tardos [**11**] of the so-called Stanley-Wilf conjecture related to the asymptotic behavior of the number of permutations that avoid a given pattern. It states that for any pattern S there exists a constant c (depending on σ) such that the number of the permutations of length n that avoid S is less than c^n.

In this paper we also study asymptotic behavior of permutations avoiding patterns, but we consider *consecutive* patterns, occurrences of which correspond to (contiguous) factors, rather than subsequences, anywhere in permutations. Suppose $\alpha_n(S)$ is the number of permutations avoiding a consecutive pattern S. It is known [**5**] that $\lim_{n\to\infty} \sqrt[n]{\alpha_n(S)/n!}$ is a nonnegative constant. Moreover, in [**6**] asymptotics for the following consecutive patterns is given: 123, 132, 1342, 1234, and 1243. These results are obtained by representation of permutations as increasing binary trees, then using symbolic methods followed by solving certain linear differential equations with polynomial coefficients to get corresponding exponential generating functions, and, finally, using the following result:

THEOREM 1.10. *[See* [**7**, Chapter 4] *for a discussion] Let $A(z)$ be a meromorphic function on a domain of the complex plane including the origin, and let ρ be the unique pole of $A(z)$ such that $|\rho|$ is minimum. Then the following asymptotic estimate holds:*

$$[z^n]A(z) \sim \gamma \cdot \rho^{-n}$$

where γ is the residue of A in ρ.

In our paper we develop a general method (not involving generating functions) that gives detailed asymptotic expansions and allows for explicit computation of leading terms in many cases. As special cases of our results, we get a more detailed asymptotics for some of the results of Elizalde and Noy [**6**].

The outline of this paper is as follows. In § 2 we prove Theorems 1.1, 1.2, 1.3, and 1.4. We also note some symmetries of the operator T_S for certain patterns S, and consider the case of descent pattern avoidance. We use Theorem 1.4 to give the proof of Theorem 1.8 in Section 3 and the proof of Theorem 1.9 in Section 4.

2. The Operator T

LEMMA 2.1. *Let T be an operator of the form (1.1) with $0 \le \chi(x) \le 1$ for all $x \in [0,1]^m$. Then $\|T\| \le 1$ and T^m is compact.*

The adjoint operator of T is given by the expression

$$T^*(f) = \int_0^1 \chi(x_1, \ldots, x_m, u) f(x_2, \ldots, x_m, u)\, du.$$

2.1. Symmetries.

Let J and R be the following two involutions on the space $L^2([0,1]^m)$:

$$(Jf)(x_1, x_2, \ldots, x_m) = f(1-x_m, \ldots, 1-x_2, 1-x_1), \tag{2.1}$$

$$(Rf)(x_1, x_2, \ldots, x_m) = f(x_m, \ldots, x_2, x_1). \tag{2.2}$$

Observe that both J and R are self adjoint operators.

LEMMA 2.2. *Assume that χ has the symmetry*

$$\chi(x_1, x_2, \ldots, x_m, x_{m+1}) = \chi(1-x_{m+1}, 1-x_m, \ldots, 1-x_2, 1-x_1).$$

Then the adjoint of the associated operator T is given by $T^* = JTJ$. Moreover, if ϕ is an eigenfunction of the operator T with eigenvalue λ then $J\phi$ is an eigenfunction of the adjoint T^* with the eigenvalue λ.

Similarly to Lemma 2.2 we have the next lemma.

LEMMA 2.3. *Assume that χ has the symmetry*

$$\chi(x_1, x_2, \ldots, x_m, x_{m+1}) = \chi(x_{m+1}, x_m, \ldots, x_2, x_1).$$

Then we have that the adjoint of the associated operator T is given by $T^* = RTR$. Moreover, if ϕ is an eigenfunction of the operator T with eigenvalue λ then $R\phi$ is an eigenfunction of the adjoint T^* with the eigenvalue λ.

Finally, we have the following relation between T_S and T_S^*. For a permutation $\pi \in \mathfrak{S}_n$, let π^* be the reverse permutation, that is, if $\pi = (\pi_1, \pi_2, \dots, \pi_n)$ then $\pi^* = (\pi_n, \pi_{n-1}, \dots, \pi_1)$. Similarly, if $S \subset \mathfrak{S}_n$, then $S^* = \{\pi \in \mathfrak{S}_n : \pi^* \in S\}$.

Lemma 2.4. *The equality*

$$T_S^* = R T_{S^*} R$$

holds, where R is given by (2.2).

2.2. Connection with Pattern Avoidance. Here we show how operators of the form (1.1) arise naturally in the study of pattern-avoiding permutations, proving Theorem 1.1. We recall that the standard triangulation of the unit cube $[0,1]^n$ into n-simplices is in one-to-one correspondence with permutations $\sigma \in \mathfrak{S}_n$: a given σ corresponds to the simplex

$$\left\{(x_1, \dots, x_n) : x_{\sigma^{-1}(1)} \le x_{\sigma^{-1}(2)} \le \dots \le x_{\sigma^{-1}(n)}\right\}$$

which has Euclidean volume $(n!)^{-1}$.

Choose and fix a nonempty subset S of $\mathfrak{S}_{m+1}$ (the set of patterns to be avoided), and define χ_S and T_S respectively as in (1.2) and (1.3). For $n \ge m+1$, let

$$\chi_n(x_1, \dots, x_n) = \prod_{j=1}^{n-m} \chi_S(x_j, \dots, x_{m+j}) \tag{2.3}$$

Then $\chi_n(x)$ is 0 if x belongs to an n-simplex of $[0,1]^n$ corresponding to a permutation containing a forbidden pattern (starting at any j between 1 and $n-m$), and 1 otherwise. From this observation, the following lemma is immediate.

Lemma 2.5. *The formula*

$$a_n = n! \int_{[0,1]^n} \chi_n(x)\ dx$$

holds for any $n \ge m$.

Now define a sequence of functions $\{f_n\}_{n=m}^{\infty}$ on $[0,1]^m$ by the formulas

$$\begin{aligned} f_m(y_1, \dots, y_m) &= 1 \\ f_n(y_1, \dots, y_m) &= \int_{[0,1]^{n-m}} \chi_n(x_1, \dots, x_{n-m}, y_1, \dots, y_m)\ dx. \end{aligned}$$

Lemma 2.6. *For any $n \ge m$, the formula*

$$f_{n+1}(y_1, \dots, y_m) = (T_S f_n)(y_1, \dots, y_m)$$

holds.

We can also estimate the norm of $\|T_S^m\|$. The following estimate shows that $\|T_S^m\| < 1$ strictly, when S is non-empty.

Lemma 2.7. *The estimate*

$$\|T_S^m\| \le \left(\frac{a_{2m}}{(2m)!}\right)^{1/2}$$

holds.

Proof of Theorem 1.1. From Lemma 2.5 and the definition of f_n, it is easy to see that for any $n \ge m+1$,

$$\frac{a_n}{n!} = (1, f_n)$$

where the right-hand side is the inner product of the constant function 1 and the function f_n in $L^2\left([0,1]^m\right)$. From Lemma 2.6 it follows that $f_{m+n} = T_S^n f_m = T_S^n 1$ from which we conclude that for any $n \ge m+1$,

$$\frac{a_n}{n!} = \left(1, T_S^{n-m} 1\right).$$

It easily follows that

$$\frac{a_n}{n!} \le \left\|T_S^{n-m}\right\|$$

and if $n = km + r$ with $0 \le r \le m - 1$ we have by Lemma 2.7 that

$$\|T_S^{n-m}\| \le \left(\frac{a_{2m}}{(2m)!}\right)^{(k-1)/2} \|T\|^r$$

from which it follows that

$$\frac{a_n}{n!} \le C_S \left(\frac{a_{2m}}{(2m)!}\right)^{n/m}.$$

□

2.3. Spectral Theory: The Spectral Gap. In this subsection, we prove Theorem 1.2.

Suppose that T is a bounded operator on a Hilbert space $\mathcal{H}$ with the property that T^m is compact for some positive integer m. For a bounded operator A, let $\sigma(A)$ denote the spectrum of A, i.e., the set of all $\lambda \in \mathbb{C}$ for which $(A - \lambda I)^{-1}$ does not exist as a bounded operator on $\mathcal{H}$. Recall that the spectral mapping theorem (see Dunford and Schwarz [**3**], chapter VII, Theorem 11, p. 569) implies that if f is an analytic function and T is a bounded operator, then the spectrum of $f(T)$ is the image under f of $\sigma(T)$. Here $f(T)$ is defined by

$$f(T) = \frac{1}{2\pi i} \int_\gamma f(z)(T - zI)^{-1} dz$$

where γ is any contour surrounding $\sigma(T)$; is is easy to see that if $f(z) = z^m$, then this coincides with the usual definition of T^m. Since $\sigma(T^m)$ is at most a countable set with 0 as the only possible accumulation point, we immediately obtain:

Lemma 2.8. *Suppose that T is a bounded operator on a Hilbert space $\mathcal{H}$ and that T^m is compact for some positive integer m. Then the spectrum of T is at most countable and has zero as the only possible accumulation point.*

Proof of Theorem 1.2. All of the statements except the assertion that $r(T_S) < 1$ follow from Lemma 2.8. From Lemma 2.7 we have $\|T_S^m\| < 1$ The discreteness of the spectrum of T_S implies that $r(T_S) = \sup\{|\lambda| : \lambda \in \sigma(T_S)\}$. Since $\sigma(T_S^m) = \{\lambda^m : \lambda \in \sigma(T_S)\}$ it follows from this estimate that $\sigma(T_S)$ is contained in a closed disc of radius $(a_{2m}/(2m)!)^{1/(2m)} < 1$. □

To give the proof of Theorem 1.3, we note the following result which is a special case of Theorem 6.3 in Kreĭn and Rutman [**10**].

Theorem 2.9. *(see [**10**], Theorem 6.3) Let (X, μ) be a measure space and A be a compact operator on $L^2(X, \mu)$. Suppose that A is strongly positive. Then:*
(a) There is a unique strictly positive function $\phi \in L^2(X, \mu)$ and $\rho > 0$ with $A\phi = \rho\phi$ and $\|\phi\| = 1$,
(b) There is a unique nonnegative function $\psi \in L^2(X, \mu)$ with $A^*\psi = \rho\psi$ and $\|\psi\| = 1$, and
(c) If λ is any other eigenvalue of A, then $|\lambda| < \rho$ strictly.

Proof of Theorem 1.3. It follows from the hypothesis and Theorem 2.9(a) and (c) that the operator T_S^k has a positive eigenvalue α of maximum modulus with associated positive eigenfunction ϕ. Let ρ be the unique positive kth root of α. By the spectral mapping theorem, $\omega\rho$ is an eigenvalue of T_S for some kth root of unity $\omega = \exp(2\pi ij/k)$, $0 \le j \le k - 1$. From the spectral mapping theorem again, it follows that $\omega^n\rho^n$ is an eigenvalue of T_S^n for any positive integer n. Moreover, since $\omega^k\rho^k$ is an eigenvalue of maximum modulus for T_S^k, it follows from the spectral mapping theorem that $\omega^n\rho^n$ will be an eigenvalue of maximum modulus for T_S^n if $n \ge k$. But T_S^n is positivity improving for any such n, so ω^n is real for all $n \ge k$. Hence $\omega = 1$ and ρ is an eigenvalue of T_S. We may now identify ϕ as the unique positive eigenfunction of T_S whose real eigenvalue $\rho > 0$ has maximum modulus, and applying the spectral mapping theorem again we see that all other eigenvalues of T_S have modulus strictly less than ρ. The statements about T_S^* follow from Theorem 2.9(b) and (c) and a similar argument. □

To prove Theorem 1.5, we will need the following lemma. In what follows, Δ_π denotes the simplex in $[0,1]^n$ corresponding to $\pi \in \mathfrak{S}_n$.

Lemma 2.10. *Let $S \subset \mathfrak{S}_{m+1}$ and suppose that G_S is strongly connected and the two monotone permutations $12 \cdots m+1$ and $m+1 \cdots 21$ do not belong to the set S. Then there exist a positive integer k such that for any two permutations σ and π in $\mathfrak{S}_n$ and any function $f \in L^2\left([0,1]^m\right)$ such that $f|_{\Delta_\pi}$ is nonnegative and nonzero, the function $T^k f|_{\Delta_\sigma}$ is strictly positive.*

Now consider the adjoint operator T_S^*. Since $(T_S^*)^m$ is compact it follows that $\sigma(T_S^*)$ is a discrete set whose only accumulation point is 0. It is not difficult to see that $\sigma(T_S^*)\setminus\{0\}$ consists of those λ with $\overline{\lambda} \in \sigma(T_S)$. Indeed, if $\lambda \in \sigma(T)$ and $\lambda \neq 0$, then λ is an isolated singularity of T. Hence $\ker(T-\lambda I)$ is nonempty since $\ker(T-\lambda I)$ is the range of the projection given by the residue of $(T-zI)^{-1}$ at $z=\lambda$. Since any eigenvector of T with eigenvalue λ is also an eigenvector of T^m with eigenvalue λ^m and T^m is compact, it follows that $V_\lambda = \ker(T-\lambda I)$ has finite dimension N_λ for any $\lambda \neq 0$. A similar argument applies to T^*, and the identity

$$\left[(T-zI)^{-1}\right]^* = (T^* - \overline{z}I)^{-1}$$

shows that the finite-dimensional space $W_\lambda = \ker(T^* - \overline{\lambda}I)$ has the same dimension as V_λ. Recall that

$$P = \operatorname*{Res}_{z=\lambda} (T-zI)^{-1}$$

projects onto V_λ, so clearly P^* projects onto W_λ.

Now let $\{\varphi^i\}_{i=1}^{N_\lambda}$ be an orthogonal basis for V_λ. By the Riesz representation theorem, the functional $\psi \mapsto (\varphi^i, P\psi)$ is represented by a vector ψ_i so that

$$P = \sum_{i=1}^{N_\lambda} (\psi^i, \,\cdot\,)\varphi^i. \tag{2.4}$$

Since P^* is the projection onto $\ker\left(T^* - \overline{\lambda}I\right)$, the vectors ψ_i are eigenvectors of T^* with eigenvalue $\overline{\lambda}$. The condition that $P^2 = P$ implies that

$$\left(\psi^i, \varphi^j\right) = \delta_{ij} \tag{2.5}$$

These conditions suffice to determine the ψ^j given a choice of $\{\varphi^j\}$.

2.4. Spectral Theory: The Expansion Theorem. We now consider the spectral expansion of T^n, assuming now that $\sigma(T)$ is contained in the interior of the unit disc. From the analytic functional calculus we have

$$T^n = \frac{1}{2\pi i}\int_{|z|=1} (T-zI)^{-1}\ z^n\ dz$$

If we write $\sigma(T) = \{\lambda_k\}_{k=1}^\infty$ with $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq 0$ and let $r_k = |\lambda_k|$ we then have

$$T^n = \sum_{j=1}^k \lambda_j^n P_j + \mathcal{O}\left(r_{k+1}^n\right) \tag{2.6}$$

by shrinking the contour. Here P_j is the projection for $\lambda = \lambda_j$ and the remainder estimate depends on

$$\sup_{|z|=r} \left\|(T-zI)^{-1}\right\|$$

where $r > 0$ is chosen so that (i) all the eigenvalues $\{\lambda_j\}_{j=1}^k$ lie in the exterior of the disc of radius r and (ii) the circle $|z| = r$ contains no eigenvalues of T. This choice is possible since $\sigma(T)$ is discrete.

Note that, in case $\sigma(T) = \{0\}$, we do not obtain a meaningful formula–there must be at least one nonzero eigenvalue for the expansion to make sense.

Proof of Theorem 1.4. We take $T = T_S$ and note that, by hypothesis, the eigenvalue of T_S having greatest modulus is positive and simple. From (2.6), (2.4) and the simplicity of ρ we get

$$(1, T_S^k 1) = \rho^k(\psi, 1)(\varphi, 1) + \mathcal{O}\left(r_2^k\right)$$

provided $(\psi, \varphi) = 1$; here φ and ψ are respectively the eigenfunctions of T_S and T_S^* associated with eigenvalue ρ. The conclusion is immediate. □

From (2.6) and (2.4), one can refine the expansion as follows if other eigenvalues and eigenvectors are known. Ordering the eigenvalues as above we have for any integer N that

$$(1, T_S^k 1) = \sum_{j=1}^{N} c_j \lambda_j^k + \mathcal{O}(r_{N+1}^k) \tag{2.7}$$

where

$$\begin{aligned} c_j &= (1, P_j 1) \\ &= \sum_{m=1}^{N_j} (\psi_j^m, 1)(\varphi_j^m, 1) \end{aligned}$$

where $\{\varphi_j^m\}$ and $\{\psi_j^m\}$ are bases for the $\lambda = \lambda_j$ eigenspaces of T_S and T_S^*, respectively, so chosen that the normalization (2.5) holds.

2.5. Descent pattern avoidance. The descent set of a permutation π in the symmetric group on n elements is the subset of $\{1, \ldots, n-1\}$, given by $\{i : \pi_i > \pi_{i+1}\}$. An equivalent notion is the descent word, defined as follows. The descent word of the permutation π is the word $u = u_1 \cdots u_{n-1}$ where $u_i = a$ if $\pi_i < \pi_{i+1}$ and $u_i = b$ otherwise.

Let U be a collection of ab-words of length m. The permutation π avoids the set U if there is no consecutive subword of the descent word of π contained in the collection U.

Descent pattern avoidance is a special case of consecutive pattern avoidance. For instance, permutations avoiding the word aab is the permutations avoiding the set $S = \{1243, 1342, 2341\}$, since these three permutations are the permutations with descent word aab.

For an ab-word u of length $m-1$ define the descent polytope P_u to be the subset of the unit cube $[0,1]^m$ corresponding to all vectors with descent word u. That is,

$$P_u = \{(x_1, \ldots, x_m) \in [0,1]^m \ : \ x_i \leq x_{i+1} \text{ if } u_i = a \text{ and } x_i \geq x_{i+1} \text{ if } u_i = b\}.$$

Observe that the m-dimensional unit cube is the union of the 2^{m-1} descent polytopes P_u. Now the operator T corresponding to the descent pattern avoidance of the set U has the following form. For an ab-word u of length $m-2$ and $y \in \{a, b\}$ we have

$$\begin{aligned} T(f)|_{P_{uy}} &= \int_0^{x_1} \chi(auy) \cdot f(t, x_1, \ldots, x_{m-1})|_{P_{au}} \, dt \\ &\quad + \int_{x_1}^{1} \chi(buy) \cdot f(t, x_1, \ldots, x_{m-1})|_{P_{bu}} \, dt, \end{aligned} \tag{2.8}$$

where by abuse of notation we let $\chi(w) = 1$ if w does not belong to the set U and $\chi(w) = 0$ otherwise.

Proposition 2.11. *Let T be the operator associated with a descent pattern avoidance and let k be an integer such that $1 \leq k \leq m-1$. Let u be an ab-word of length $m-1$. Then the function $T^k(f)$ restricted to the descent polytope P_u only depends on the variables x_1 through x_{m-k}.*

Corollary 2.12. *Let T be the operator associated with a descent pattern avoidance and let ϕ be an eigenfunction associated with a non-zero eigenvalue λ. Let u be an ab-word of length $m-1$. Then the eigenfunction restricted to the descent polytope P_u only depends on the variable x_1.*

Let V be the subspace of $L^2([0,1]^m)$ consisting of all functions f that only depend on the variable x_1 when restricted to each of the descent polytopes P_u. Observe that the subspace V is invariant under the operator T. That is, the operator T restricts to the subspace V. Moreover the constant function 1 belongs to V. Hence to understand the behavior of $T^n(1)$ it is enough to study this restricted operator.

In order to describe the subspace V more explicitly define for an ab-word u of length $m-1$ the polynomial $f(u; x_1)$ as follows:

$$f(u; x_1) = \int_{(x_1, x_2, \ldots, x_m) \in P_u} 1 dx_2 \cdots dx_m.$$

This polynomial was first introduced and studied in [**4**], with a different indexing.

Let p be a vector $(p_u(x_1))_{u\in\{a,b\}^{m-1}}$. That is, the vector p consists of one-variable functions in the variable x_1 and is indexed by ab-words of length $m-1$. Consider the function f on $[0,1]^m$ defined by

$$f(x_1,\ldots,x_m)|_{P_u} = p_u(x_1)$$

for all ab-words u of length $m-1$. Observe that the function f belongs to $L^2([0,1]^m)$, and hence to the invariant subspace V, if and only if

$$\int_0^1 f(u;x_1)\cdot |p_u(x_1)|^2 \;\; dx_1 < \infty$$

for all ab-words u of length $m-1$. For two functions f and g in the subspace V, corresponding to the two vectors $(p_u(x_1))_{u\in\{a,b\}^{m-1}}$ and $(q_u(x_1))_{u\in\{a,b\}^{m-1}}$, the inner product is given by

$$(f,g) = \sum_{u\in\{a,b\}^{m-1}} \int_0^1 f(u;x_1)\cdot p_u(x_1)\cdot q_u(x_1)dx_1.$$

This discussion leads to the following structural result about the subspace V.

Proposition 2.13. *The invariant subspace V is isometrically isomorphic to the space $L^2\left([0,1]\right)^{2^{m-1}}$.*

3. 123-Avoiding Permutations

A 123-avoiding permutation is a permutation $\pi \in \mathfrak{S}_n$ with no index j so that $\pi_j < \pi_{j+1} < \pi_{j+2}$, where $1 \le j \le n-2$. We denote by a_n the number of 123-avoiding permutations in $\mathfrak{S}_n$. Thus, in the notation of the introduction S consists of the single permutation 123 and

$$\chi_S(x_1,x_2,x_3) = \begin{cases} 0 \text{ if } x_1 \le x_2 \le x_3; \\ 1 \text{ otherwise.} \end{cases} \tag{3.1}$$

We will obtain an asymptotic formula for a_n by computing the eigenvalues and eigenfunctions of the corresponding operator T_S and using the spectral expansions of Section 2.3. As we will see, in this case the operator T_S has real eigenvalues and a trivial kernel. This is related to the fact that the eigenvalue problem for T_S can be recast as an eigenvalue problem for a first-order system of differential equations.

3.1. Eigenfunctions and Eigenvectors. Since 123-avoiding permutations can be viewed as permutations with no double descents Corollary 2.12 allows us to recast then problem of finding eigenfunctions in two variables into finding two one-variable functions.

Proposition 3.1. *The eigenvalues λ_k of the operator T on $L^2([0,1]^2)$ are given by*

$$\lambda_k = \frac{\sqrt{3}}{2\pi\cdot\left(k+\frac{1}{3}\right)}, \tag{3.2}$$

where $k\in\mathbb{Z}$ and the associated eigenfunctions $\phi_k = \begin{cases} p_k(x) & \text{if } 0\le x\le y\le 1 \\ q_k(x) & \text{if } 0\le y\le x\le 1\end{cases}$ are given by

$$\phi_k = \exp\left(-\frac{x}{2\lambda}\right)\cdot \begin{cases} \cos\left(\frac{\pi}{6}+\frac{\sqrt{3}}{2}\cdot\frac{x}{\lambda}\right) & \text{if } 0\le x\le y\le 1, \\ \sin\left(\frac{\pi}{3}+\frac{\sqrt{3}}{2}\cdot\frac{x}{\lambda}\right) & \text{if } 0\le y\le x\le 1. \end{cases} \tag{3.3}$$

Note that the eigenvalues are ordered by

$$\lambda_0 > -\lambda_{-1} > \lambda_1 > -\lambda_{-2} > \lambda_2 > -\lambda_{-3} > \lambda_3 > \cdots > 0.$$

By applying the involution J we obtain the adjoint eigenfunction

$$\psi_k = \exp\left(\frac{y-1}{2\lambda}\right)\cdot \begin{cases} \cos\left(\frac{\pi}{6}+\frac{\sqrt{3}}{2}\cdot\frac{1-y}{\lambda}\right) & \text{if } 0\le x\le y\le 1, \\ \sin\left(\frac{\pi}{3}+\frac{\sqrt{3}}{2}\cdot\frac{1-y}{\lambda}\right) & \text{if } 0\le y\le x\le 1. \end{cases} \tag{3.4}$$

PROPOSITION 3.2. *For the eigenfunctions $\phi_k = \phi$ of T and $\psi_k = \psi$ of T^* with eigenvalue $\lambda_k = \lambda = \sqrt{3}/(2\pi(k+1/3))$,*

$$(1,\phi) = (1,\psi) = \frac{\sqrt{3}}{2}\lambda^2 \tag{3.5}$$

$$(\psi,\phi) = \frac{3}{4}(-1)^k \lambda \exp\left(-\frac{1}{2\lambda}\right) \tag{3.6}$$

In particular

$$\frac{(1,\phi)\,(1,\psi)}{(\phi,\psi)} = (-1)^k \lambda^3 \exp\left(\frac{1}{2\lambda}\right) \tag{3.7}$$

3.2. Asymptotics. The above computations show that all eigenvalues of T_S are simple and give explicit formulas. We thus obtain the following expansion for $a_n/n!$ as an immediate consequence of (2.7), Propositions 3.1, and 3.2.

THEOREM 3.3. *For any positive integer $n \geq 2$ and any positive integer K, the formula*

$$\frac{a_n}{n!} = \sum_{|k|\leq K} (-1)^k \lambda_k^{n+1} \exp\left(\frac{1}{2\lambda_k}\right) + \mathcal{O}\left(r_{K+1}^n\right)$$

holds, where λ_k is given by (3.2) and

$$r_k = \frac{\sqrt{3}}{2\pi \cdot \left(k - \frac{1}{3}\right)}$$

4. 213-Avoiding Permutations

A 213-avoiding permutation is a permutation $\pi \in \mathfrak{S}_n$ which contains no sequence of the form

$$\pi_{j+1} < \pi_j < \pi_{j+2}$$

for any j with $1 \leq j \leq n-2$. We denote the number of 213-avoiding permutations of $\mathfrak{S}_n$ by b_n. Thus, S consists of the single permutation (213) and

$$\chi_S(x_1,x_2,x_3) = \begin{cases} 0 & \text{if } x_2 \leq x_1 \leq x_3, \\ 1 & \text{otherwise.} \end{cases}$$

By symmetry, the study of 213-avoiding permutations is equivalent to 132-avoiding permutations, 231-avoiding permutations and 312-avoiding permutations. However the case of 213-avoiding permutations gives the most straightforward equations.

We will compute the eigenvalues and eigenfunctions of the operator T_S and obtain an asymptotic expansion for b_n using spectral methods. In this case, it turns out that T_S has a nontrivial kernel and its eigenvalues need not be real. However, its eigenvalue of largest modulus is real and isolated, as we will show, so that we can still obtain an asymptotic formula for b_n.

4.1. Eigenfunctions and Eigenvectors. In what follows, we will make use of the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt \tag{4.1}$$

which extends to an entire function on $\mathbb{C}$, and the function

$$q(x) = \exp\left(-\frac{x^2}{2\lambda^2}\right). \tag{4.2}$$

Let

$$f(x,y) = \begin{cases} p(x,y) & \text{if } 0 \leq x \leq y \leq 1, \\ q(x,y) & \text{if } 0 \leq y \leq x \leq 1. \end{cases}$$

Then

$$(Tf)(x,y) = \begin{cases} \int_0^x p(t,x)dt + \int_y^1 q(t,x)dt & \text{if } 0 \leq x \leq y \leq 1, \\ \int_0^x p(t,x)dt + \int_x^1 q(t,x)dt & \text{if } 0 \leq y \leq x \leq 1. \end{cases}$$

Now we characterize the nonzero eigenvalues and eigenfunctions.

Proposition 4.1. *The non-zero eigenvalues λ of the operator T satisfies the equation*

$$\operatorname{erf}\left(\frac{1}{\sqrt{2}\cdot\lambda}\right) = \frac{\sqrt{2}}{\sqrt{\pi}} \tag{4.3}$$

and the corresponding eigenfunctions are

$$\varphi(x,y) = \begin{cases} q(x) - \dfrac{1}{\lambda}\int_x^y q(t)\,dt & \text{if } x \le y, \\ q(x) & \text{if } x > y, \end{cases}$$

where $q(x)$ is given by (4.2).

The adjoint operator T^* is given by

$$T^*(f(x,y)) = \begin{cases} \int_0^y q(y,u)du + \int_y^1 p(y,u)du & \text{if } 0 \le x \le y \le 1, \\ \int_0^y q(y,u)du + \int_y^x p(y,u)du & \text{if } 0 \le y \le x \le 1. \end{cases}$$

Proposition 4.2. *For a non-zero eigenvalue λ of the operator T the corresponding eigenfunction of the adjoint operator T^* is*

$$\psi(x,y) = \begin{cases} p^*(y) & \text{if } 0 \le x \le y \le 1, \\ p^*(y) - \frac{1}{\lambda}\cdot\int_x^1 p^*(u)du & \text{if } 0 \le y \le x \le 1. \end{cases}$$

where

$$p^*(y) = -2\cdot y\cdot\exp\left(\frac{y^2}{2\lambda^2}\right) + 2\cdot\lambda + \sqrt{2\pi}\cdot y\cdot\exp\left(\frac{y^2}{2\lambda^2}\right)\cdot\operatorname{erf}\left(\frac{y}{\sqrt{2}\lambda}\right). \tag{4.4}$$

Proposition 4.3. *For a non-zero eigenvalue λ with eigenvector ϕ and adjoint eigenvector ψ, we have*

$$\begin{aligned} (1,\phi) &= \lambda^2, \\ (1,\psi) &= 2\cdot\lambda^3, \\ (\psi,\phi) &= 2\cdot\lambda^2\cdot\exp(-1/(2\lambda^2)). \end{aligned}$$

In particular,

$$\frac{(1,\phi)\cdot(1,\psi)}{(\psi,\phi)} = \lambda^3\cdot\exp(1/(2\lambda^2)).$$

4.2. Asymptotics. To obtain leading asymptotics for b_n, we need to compute the eigenvalue of greatest modulus of the operator T_S and show that all other eigenvalues of T have strictly smaller moduli. From the eigenvalue condition (4.3), it suffices to study the roots of the equation $\operatorname{erf}(z) = \sqrt{2/\pi}$.

Since the error function is an increasing function on the real axis, the equation $\operatorname{erf}(z) = \sqrt{2}/\sqrt{\pi}$ has a unique real root $z_0 = 0.9019484541\ldots$. Hence the eigenvalue equation (4.3) has the unique real root $\lambda_0 = 0.7839769312\ldots$. Since the error function is an odd function we know by the strong version of the little Picard theorem that the equation $\operatorname{erf}(z) = \sqrt{2}/\sqrt{\pi}$ has infinitely many roots. The location of these roots is the subject of the next result.

Proposition 4.4. *The equation* $\operatorname{erf}(z) = \sqrt{2}/\sqrt{\pi}$ *has exactly one root in the interior of the unit disc, namely the unique real root $z_0 = 0.9019484541\ldots$, and all other (infinitely many) roots lie in the complement of the closed unit disc.*

As a corollary we have:

Corollary 4.5. *The eigenvalue equation (4.3) has the unique real root*

$$\lambda_0 = 0.7839769312\ldots$$

outside the disc of radius $1/\sqrt{2}$ centered at the origin, and all other (infinitely many) roots lie inside this disc.

Combining Propositions 4.1 through 4.3 and Corollary 4.5 using Theorem 1.4 we obtain Theorem 1.9.

5. Concluding remarks

In the case of descent pattern avoidance, can one prove that T restricted to the invariant subspace V is compact? We have done so in the case of 123-avoiding permutations.

It is straightforward to design a Viennot "pyramid" to compute the number a_n of S-avoiding permutations. For the original Viennot triangle, see [**14, 15**]. Let the entry $a_n^{i_1,\ldots,i_m}$ of the pyramid be the number of permutations in the symmetric group on n elements, avoiding the set S and ending with the m entries $i_1, \ldots, i_m$. Then the entry $a_n^{i_1,\ldots,i_m}$ is a sum of entries of the form $a_{n-1}^{j,i_1,\ldots,i_{m-1}}$. This sum being a discrete analogue of the operator T. How far does this analogue between the discrete model and the continuous one go? Does the function $f_n = T^{n-m}(1)$ approximate the n-th level of the pyramid? More exactly, how well does the integer $a_n^{i_1,\ldots,i_m}$ compare with $n! \cdot f_n(i_1/n, \ldots, i_m/n)$?

The next four largest roots to the eigenvalue equation in the 213-avoiding permutation case are:

$$\lambda_1 = 0.2141426360\ldots \pm 0.2085807022\ldots \cdot i$$
$$\lambda_2 = -0.1677323922\ldots \pm 0.2418627350\ldots \cdot i$$

Knowing these roots enables us to give an explicit error estimate in Theorem 1.9.

In this paper our object is to understand consecutive pattern avoidance. Generalized pattern avoidance was introduced by Babson and Steingrímsson [**1**]. Is there an analytic approach to obtain asymptotics for these classes of permutations? Lastly, it would be daring to ask for an analytic proof of the former Stanley-Wilf conjecture, recently proved in [**11**].

References

[1] E. Babson and E. Steingrímsson, Generalized permutation patterns and a classification of the Mahonian statistics, *Sém. Lothar. Combin.* **44** (2000), Art. B44b, 18 pp.

[2] M. Bóna, *Combinatorics of Permutations*, Chapman and Hall/CRC Press, 2004.

[3] N. Dunford and J. T. Schwarz, *Linear Operators, Part I: General Theory.* New York: Interscience Publishers, 1967.

[4] R. Ehrenborg, M. Levin and M. Readdy, A probabilistic approach to the descent statistic, *J. Combin. Theory Ser. A* **98** (2002), 150–162.

[5] S. Elizalde, Asymptotic enumeration of permutations avoiding generalized patterns, *Adv. in Appl. Math.* **36** (2006), 138–155.

[6] S. Elizalde and M. Noy, Consecutive patterns in permutations, *Adv. in Appl. Math.* **30** (2003), 110–125.

[7] P. Flajolet and R. Sedgewick, Analytic Combinatorics, 1998, book in preparation. Individual chapters are available as INRIA, Research Reports 1888, 2026, 2376, 2956, 3162.

[8] F. R. Gantmacher, *The Theory of Matrices.* Vols. 1, 2. Translated by K. A. Hirsch. Chelsea Publishing Co., New York 1959.

[9] D. E. Knuth, *The art of computer programming.* Vol. 1 *Fundamental algorithms.* Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont, 1969.

[10] M. G. Kreĭn and M. A. Rutman, Linear operators leaving invariant a cone in a Banach space. *Uspehi Matem. Nauk (N.S.)* **3** (1948), no. 1 (23), 3–95. English translation in *Amer. Math. Soc. Translation* **1950** (1950), no. 26, 1–128.

[11] A. Marcus and G. Tardos, Excluded Permutation Matrices and the Stanley-Wilf Conjecture, *J. Combin. Th. Ser. A.* **107** (2004), 153–160.

[12] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, Volume I: Functional Analysis.* New York, London: Academic Press, 1972.

[13] R. Simion and F. Schmidt, Restricted permutations, *European J. Combin.* **6** (1985), no. 4, 383–406.

[14] G. Viennot, Permutations ayant une forme donnée, *Discrete Math.* **26** (1979), 279–284.

[15] G. Viennot, Équidistribution des permutations ayant une forme donnée selon les avances et coavances, *J. Combin. Th. Ser. A.* **31** (1981), 43–55.

Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA
E-mail address: jrge@ms.uky.edu
URL: http://www.ms.uky.edu/~jrge/

Department of Mathematics, University of Reykjavik, Ofanleiti 2, IS-103 Reykjavik, Iceland
E-mail address: sergey@ru.is
URL: http://www.ru.is/sergey/

Department of Mathematics, University of Kentucky, Lexington, KY 40506, USA
E-mail address: perry@ms.uky.edu
URL: http://www.ms.uky.edu/~perry/

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Statistics on Signed Permutations Groups (Extended Abstract)

Michael Fire

Abstract. A classical result of MacMahon shows that the length function and the major index are equidistributed over the symmetric groups. Through the years this result was generalized in various ways to signed permutation groups. In this paper we present several new generalizations, in particular, we study the effect of different linear orders on the letters $[-n, n]$ and generalize a classical result of Foata and Zeilberger.

Résumé. MacMahon a demontré que la fonction de longueur et l'indice majeur sont équi-distribué dans les groupes symétriques. Depuis, ce résultat a été generalisé aux groupes de permutations signées de plusieurs façons. Dans ce travail, nous présentons plusieurs généralisations, et en particulier, nous étudions l'effet d'imposer un ordre linéaire sur $[-n, n]$ et nous généralisons un résultat de Foata et Zeilberger.

1. Introduction

The *signed permutation groups*, also known as the Weyl groups of type B or as the hyperoctahedral groups, are fundamental objects in today's mathematics. A better understanding of these groups may help to advance research in many fields. One method of studying these groups is by using numerical statistics and finding their generating functions. This method was successfully applied in the case of the *symmetric groups*. MacMahon [**13**] considered four different statistics for a permutation π in the *symmetric group*: the number of *descents* ($des(\pi)$), the number of *excedances* ($exc(\pi)$), the *length* statistic ($\ell(\pi)$), and the *major index* ($maj(\pi)$). MacMahon showed that the *excedance* number is equidistributed with the *descent* number, and that the *length* is equidistributed with the *major index* over the *symmetric groups*.

We will discuss three types of statistics: *Eulerian* statistics, which are equidistributed with the *descent number*; *Mahonian* statistics, which are equidistributed with *length*; *Euler-Mahonian* pairs of statistics, which are equidistributed with the pair consisting of the *descent number* and the *major index*. Through the years many generalizations to MacMahon's results were found. In particular, Foata and Zeilberger found that the *Denert statistic* and the *excedance number* are Euler-Mahonian [**10**]. Recently, Adin and Roichman [**3**] generalized MacMahon's result on the *major index* to the *signed permutations groups*, by introducing a new Mahonian statistic, the *flag major index*. See also [**1**]. The associated signed Mahonian statistic was studied in [**2**]. In this extended abstract we will generalize the Foata-Zeilberger result to signed permutation groups, and will investigate the effect of different linear orders on the letters $[-n, n] \setminus \{0\}$ on the resulting generating functions.

The full background, proofs and extensions for *colored permutations groups* to this work can be found in [**8**].

2000 *Mathematics Subject Classification.* Primary 05A15.

Key words and phrases. algebraic combinatorics, permutations groups, signed permutations groups, permutations statistics, Mahonian statistics .

Partially supported by EC's IHRP Programme, within the Research Training Network "Algebraic Combinatorics in Europe", grant HPRN-CT-2001-00272.

2. Background

2.1. Statistics on the Symmetric Group. In this subsection we present the main definitions, notation, and theorems on the *symmetric groups* (i.e., the Weyl groups of type A), denoted S_n.

Definition 2.1. Let $\mathbf{N}$ the set of all the *natural numbers*, a *permutation* of order $n \in \mathbf{N}$ is a bijection $\pi : \{1,2,3,\ldots,n\} \rightarrow \{1,2,3,\ldots,n\}$.

Remark 2.2. Permutations are traditionally written in a two-line notation as:

$$\pi = \left(\begin{array}{ccccc} 1 & 2 & 3 & \ldots & n \\ \pi(1) & \pi(2) & \pi(3) & \ldots & \pi(n) \end{array} \right).$$

However for convenience we will use the shorter notation:

$$\pi = [\pi(1), \pi(2), \pi(3), \ldots, \pi(n)].$$

For example: $\pi = \left(\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 1 & 5 \end{array} \right)$ will be written as $\pi = [2,4,3,1,5]$.

Definition 2.3. The *symmetric group* of degree $n \in \mathbf{N}$ (denoted S_n) is the group consisting of all the permutations of order n, with composition as the group operation.

Definition 2.4. The Coxeter generators of S_n are $s_1, s_2, \ldots, s_{n-1}$ where $s_i := [1,2,\ldots,i+1,i,\ldots,n]$.

It is a well-known fact that the *symmetric group* is a *Coxeter group* with respect to the above generating set $\{s_i \mid 1 \leq i \leq n-1\}$. This fact gives rise to the following natural statistic of permutations in the symmetric group:

Definition 2.5. The *length* of a permutation $\pi \in S_n$ is defined to be:

$$\ell(\pi) := \min\{\ r \geq 0 \mid \pi = s_{i_1} \ldots s_{i_r} \text{ for some } i_1, \ldots, i_r \in [1,n]\ \}.$$

Here are other useful statistics on S_n that we are going to work with:

Definition 2.6. Let $\pi \in S_n$. Define the following:

(1) The *inversion number* of π:
$$inv(\pi) := |\{(i,j) \mid 1 \leq i < j \leq n,\ \pi(i) > \pi(j)\}|.$$
Note that $inv(\pi) = \ell(\pi)$.
(2) The *descent set* of π: $Des(\pi) := \{1 \leq i \leq n-1 \mid \pi(i) > \pi(i+1)\}$.
(3) The *decent number* of π: $des(\pi) = |Des(\pi)|$.
(4) The *major-index* of π: $maj(\pi) := \sum\limits_{i \in Des(\pi)} i$.
(5) The *sign* of π: $sign(\pi) := (-1)^{\ell(\pi)}$.
(6) The *excedance number* of π: $exc(\pi) := |\{1 \leq i \leq n \mid \pi(i) > i\}|$.

Example 2.7. Let $\pi = [2,3,1,5,4] \in S_5$. We can compute the above statistics on π, namely:

$$\begin{aligned} &inv(\pi) = \ell(\pi) = 3,\ Des(\pi) = \{2,4\},\ des(\pi) = 2,\ maj(\pi) = 6, \\ &sign(\pi) = (-1)^3 = -1, \text{ and } exc(\pi) = 3. \end{aligned}$$

Remark 2.8. Throughout the paper we use the following notations for a nonnegative integer n:

$$\begin{aligned} &[n]_q := \frac{1-q^n}{1-q},\ [n]_q! = [1]_q[2]_q \ldots [n]_q, \\ &[n]_{\pm q}! := [1]_q[2]_{-q}[3]_q[4]_{-q} \ldots [n]_{(-1)^{n-1}q}, \text{ and also} \\ &(a;q)_n := \begin{cases} 1, & \text{if } n = 0; \\ (1-a)(1-aq)\ldots(1-aq^{n-1}), & \text{otherwise.} \end{cases} \end{aligned}$$

MacMahon [**13**] was the first to find a connection between these statistics. He discovered that the *excedance* number is equidistributed with the *descent* number, and that the *inversion* number is equidistributed with the *major index*:

THEOREM 2.9. [**13**]

$$\sum_{\pi \in S_n} q^{inv(\pi)} = \sum_{\pi \in S_n} q^{maj(\pi)} = [1]_q[2]_q[3]_q \dots [n]_q = [n]_q!.$$

THEOREM 2.10. [**13**]

$$\sum_{\pi \in S_n} q^{exc(\pi)} = \sum_{\pi \in S_n} q^{des(\pi)}.$$

Gessel and Simion gave a similar factorial type product formula for the *signed Mahonian*:

THEOREM 2.11. [**14**, Cor. 2]

$$\sum_{\pi \in S_n} sign(\pi) q^{maj(\pi)} = [n]_{\pm q}!.$$

A bivariate generalization of MacMahon's Theorem 2.9 was achieved during the 1970's by Foata and Schützenberger :

THEOREM 2.12. [**9**]

$$\sum_{\pi \in S_n} q^{maj(\pi)} t^{des(\pi^{-1})} = \sum_{\pi \in S_n} q^{inv(\pi)} t^{des(\pi^{-1})}.$$

In the same article Foata and Schützenberger also proved another bivariate connection between the different statistics:

THEOREM 2.13. [**9**]

$$\sum_{\pi \in S_n} q^{maj(\pi^{-1})} t^{maj(\pi)} = \sum_{\pi \in S_n} q^{\ell(\pi)} t^{maj(\pi)}.$$

In 1990 during her research of the genus zeta function, Denert found a new statistic which was also Mahonian:

DEFINITION 2.14. [**6**] Let be $\pi \in S_n$, define the *Denert's statistic* to be:

$$\begin{array}{rcl} den(\pi) & := & |\{1 \le l < k \le n \mid \pi(k) < \pi(l) < k\}| \\ & + & |\{1 \le l < k \le n \mid \pi(l) < k < \pi(k)\}| \\ & + & |\{1 \le l < k \le n \mid k < \pi(k) < \pi(l)\}|. \end{array}$$

Later in the same year Foata and Zeilberger proved that the pair of statistics (exc, den) is equidistributed with the pair (des, maj):

THEOREM 2.15. [**10**]

$$\sum_{\pi \in S_n} q^{exc(\pi)} t^{den(\pi)} = \sum_{\pi \in S_n} q^{des(\pi)} t^{maj(\pi)}.$$

2.2. Signed Permutations Groups. In this subsection we present the main definitions, notation and theorems for the classical Weyl groups of type B, also known as the *hyperoctahedral groups* or the *signed permutations groups*, and denoted B_n.

DEFINITION 2.16. The *hyperoctahedral group* of order $n \in \mathbf{N}$ (denoted B_n) is the group consisting of all the bijections σ of the set $[-n, n] \backslash \{0\}$ onto itself such that $\sigma(-a) = -\sigma(a)$ for all $a \in [-n, n] \backslash \{0\}$, with composition as the group operation.

REMARK 2.17. There are different notations for a permutation $\sigma \in B_n$. We will use the notation $\sigma = [\sigma(1), ..., \sigma(n)]$.

We identify S_n as a subgroup of B_n, and B_n as a subgroup of S_{2n}. As in S_n we also have many different statistics; we will describe the main ones:

THEOREM 2.18. Let $\sigma \in B_n$, define the following statistics on σ:

(1) The *inversion number* of σ: $inv(\sigma) := |\{(i, j) \mid 1 \le i < j \le n,\ \sigma(i) > \sigma(j)\}|$.

(2) The *descent set* of σ:

$$Des(\sigma) := \{1 \le i \le n-1 \mid \sigma(i) > \sigma(i+1)\}.$$

(3) The *type A descent number* of σ: $des_A(\sigma) := |Des(\sigma)|$.
(4) The *type B descent number* of σ:

$$des_B(\sigma) := |\{0 \le i \le n-1 \mid \sigma(i) > \sigma(i+1)\}|, \text{ where here } \sigma(0) := 0.$$

(5) The *major index* of σ: $maj(\sigma) := \sum_{i \in Des(\sigma)} i$.
(6) The *negative set* of σ: $Neg(\sigma) := \{i \in [1, \dots, n] \mid \sigma(i) < 0\ \}$.
(7) The *negative number* of σ: $neg(\sigma) := |Neg(\sigma)|$.
(8) The *negative number sum* of σ: $nsum(\sigma) := -\sum_{i \in Neg(\sigma)} \sigma(i)$.

It is well known (see, e.g. [**5**, Proposition 8.1.3]) that B_n is a Coxeter group with respect to the generating set $\{s_0, s_1, \dots, s_{n-1}\}$, where s_i, $1 \le i \le n-1$, are defined as in S_n (see 2.4), and s_0 is defined as:

$$s_0 := [-1, 2, 3, \dots, n].$$

This gives rise to another natural statistic on B_n, the *length* statistic:

DEFINITION 2.19. For all $\sigma \in B_n$ the *length* of σ is:

$$\ell(\sigma) := \min\{r \ge 0 \mid \sigma = s_{i_1} s_{i_2} \dots s_{i_n} \text{ for some } i_1, \dots, i_r \in [0, n-1]\}.$$

There is a well-known direct combinatorial way to compute this statistic:

THEOREM 2.20. ([**5**, Propositions 8.1.1 and 8.1.2]) For all $\sigma \in B_n$ the length of σ can be computed as:

$$\ell(\sigma) = inv(\sigma) - \sum_{i \in Neg(\sigma)} \sigma(i).$$

Using the last definition we can define another natural statistic on B_n, the *sign* statistic:

DEFINITION 2.21. For all $\sigma \in B_n$ the *sign* of σ is:

$$sign(\sigma) := (-1)^{\ell(\sigma)}.$$

The generating function of *length* is also called the Poincaré polynomial and can be presented in the following manner:

THEOREM 2.22. [**12**, §3.15]

$$\sum_{\sigma \in B_n} q^{\ell(\sigma)} = [2]_q [4]_q \dots [2n]_q = \prod_{i=1}^{n} [2i]_q.$$

Recently, Adin and Roichman generalized MacMahon's result Theorem 2.9 to B_n, by introducing a new Mahonian statistic, the *flag major index*:

DEFINITION 2.23. [**3**] The *flag major index* of $\sigma \in B_n$ is defined as:

$$\text{flag-major}(\sigma) := 2maj(\sigma) + neg(\sigma),$$

where $maj(\sigma)$ is calculated with respect to the linear order

$$-1 < -2 < \dots < -n < 1 < 2 < \dots < n.$$

THEOREM 2.24. [**3**, §2]

$$\sum_{\sigma \in B_n} q^{\ell(\sigma)} = \sum_{\sigma \in B_n} q^{flag-major(\sigma)} = [2]_q [4]_q \dots [2n]_q.$$

REMARK 2.25. The previous result still holds if $maj(\sigma)$ is calculated with respect to the *natural order* $-n < -(n-1) < ... < -2 < -1 < 1 < 2 < ... < n-1 < n$, see also [**3**].

Adin, Brenti and Roichman introduced another statistic which was also Mahonian, the *nmaj* statistic:

DEFINITION 2.26. [**1**, §3.2] Let $\sigma \in B_n$ then the *negative major index* is defined as:

$$nmaj(\sigma) := maj(\sigma) - \sum_{i \in Neg(\sigma)} \sigma(i) = maj(\sigma) + nsum(\sigma).$$

THEOREM 2.27. [**1**]

$$\sum_{\sigma \in B_n} q^{\ell(\sigma)} = \sum_{\sigma \in B_n} q^{nmaj(\sigma)}.$$

In the same article [**1**] they also defined a new descent multiset and new descent statistics, and found a new Euler-Mahonian bivariate distribution for these statistics:

DEFINITION 2.28. [**1**, §3.1 and §4.2] Let $\sigma \in B_n$ define:

(1) The *negative descent multiset* of σ:

$$NDes(\sigma) := Des(\sigma) \bigcup \{-\sigma(i) \mid i \in Neg(\sigma)\},$$

where $\bigcup$ stands for multiset union.

(2) The *negative descent* statistic of σ: $ndes(\sigma) := |NDes(\sigma)|$.

(3) The *flag-descent* number of σ: $fdes(\sigma) := des_A(\sigma) + des_B(\sigma) = 2des_A(\sigma) + \varepsilon(\sigma)$, where

$$\varepsilon(\sigma) := \begin{cases} 1, & \text{if } \sigma(1) < 0; \\ 0, & \text{otherwise.} \end{cases}$$

THEOREM 2.29. [**1**, §4.3]

$$\sum_{\sigma \in B_n} t^{ndes(\sigma)} q^{nmaj(\sigma)} = \sum_{\sigma \in B_n} t^{fdes(\sigma)} q^{flag-major(\sigma)}.$$

In their article from 2005 Adin, Gessel, and Roichman gave a type B analogue to the Gessel-Simion Theorem(e.g. [**14**, Cor. 2]):

THEOREM 2.30. [**2**, §5.1]

$$\sum_{\sigma \in B_n} sign(\sigma) q^{flag-major(\sigma)} = [2]_{-q}[4]_q \ldots [2n]_{(-1)^n q}.$$

Where *flag major index* computed with respect to the linear order:

$$-1 < -2 < \ldots < -n < 1 < 2 < \ldots < n.$$

3. Main Results

3.1. Signed-Mahonian and Mahonian-Mahonian Statistics.

DEFINITION 3.1. A *linear order of length n, denoted K_n, is a bijection*

$$K_n : [-n, n] \backslash \{0\} \to [1, 2n].$$

We can calculate permutation statistics according to a linear order K_n, we use the following notation: $maj_{K_n}(\sigma)$, $des_{K_n}(\sigma)$, $flag - major_{K_n}(\sigma)$, $nmaj_{K_n}(\sigma)$ etc, to indicate that the corresponding statistic is calculated with respect to the linear order K_n. We also use the notation: $m >_{K_n} l$, to indicate, that according to the linear order K_n 'm' is larger than 'l', i.e. that $s = K_n(m)$, $r = K_n(l)$, and $s > r$.

EXAMPLE 3.2. Let K_n be a linear order and let $\sigma \in B_n$. Then:

$$maj_{K_n}(\sigma) := \sum_{\sigma(i) >_{K_n} \sigma(i+1)} i.$$

NOTE 3.3. Notice that for any linear order K_n, and for any $\sigma \in B_n$, $neg(\sigma) = neg_{K_n}(\sigma)$. This also applies to the *length* statistic, because it is defined with respect to the Coxeter generators, which do not depend on the choice of linear order.

The following proposition is a more general version of Remark 2.25:

Proposition 3.4. Let K_n be a linear order then:

$$\sum_{\sigma\in B_n} q^{flag-major(\sigma)} = \sum_{\sigma\in B_n} q^{flag-major_{K_n}(\sigma)}.$$

In the following theorems we give simple factorial-type product formulas for the generating function for the signed-Mahonian and Mahonian-Mahonian statistics over B_n.

Let be N the natural order, $N:\ -n < -(n-1) < \ldots < -1 < 1 < \ldots < n-1 < n$, then:

Theorem 3.5.

$$\sum_{\sigma\in B_n} sign(\sigma) q^{flag-major_N(\sigma)} = (q;-1)_n [n]_{\pm q^2}!.$$

The next theorem presents *signed-Mahonian* calculation using the new Mahonian statistic *nmaj*:

Theorem 3.6.

$$\sum_{\sigma\in B_n} sign(\sigma) q^{nmaj_N(\sigma)} = (q;-q)_n [n]_{\pm q}!.$$

Definition 3.7. Define the following set:

$$U_n := \{\tau \in B_n \mid \tau(1) < \tau(2) < \ldots < \tau(n-1) < \tau(n)\}.$$

There are several facts (see also [**1**],[**2**]) about the set U_n that can be directly concluded from the definition of U_n, namely: each $\sigma \in B_n$ has a unique representation as:

$$\sigma = \tau\pi\ (\tau \in U_n \text{ , and } \pi \in S_n).$$

Definition 3.8. Define the following subsets of U_n:

(1) $U_{n1} := \{\tau \in U_n \mid \tau(1) = -n\}$.
(2) $U_{n2} := \{\tau \in U_n \mid \tau(n) = n\}$.

Note 3.9. $U_n = U_{n1} \uplus U_{n2}$, where $\uplus$ stands for disjoint union.

We also define two bijections from U_{n-1} one onto U_{n1}, and one onto U_{n2}:

Definition 3.10. For $i \in 1, 2$, define $\varphi_{ni} : U_{n-1} \to U_{ni}$ by:

(1) $\varphi_{n1}(\tau)(i) = \begin{cases} -n, & \text{i=1;} \\ \tau(i-1), & 2 \le i \le n. \end{cases}$

(2) $\varphi_{n2}(\tau)(i) = \begin{cases} \tau(i), & 1 \le i \le n-1; \\ n, & i = n. \end{cases}$

Theorem 3.11.

$$\sum_{\sigma\in B_n} q^{flag-major_N(\sigma)} t^{nmaj_N(\sigma)} = \prod_{i=1}^{n} (1+qt^i)[n]_{q^2t}!.$$

Proof. (Sketch, more detailed proof can be found at [**8**]) We will prove this theorem by reducing the problem to U_n:

$$\begin{aligned}
\sum_{\sigma\in B_n} q^{flag-major_N(\sigma)} t^{nmaj_N(\sigma)} &= \sum_{\pi\in S_n,\ \tau\in U_n} q^{2maj(\pi)+neg(\tau)} t^{maj(\pi)+nsum(\tau)} \\
&= \sum_{\tau\in U_n} q^{neg(\tau)} t^{nsum(\tau)} \sum_{\pi\in S_n} q^{2maj(\pi)} t^{maj(\pi)} \\
&= \sum_{\tau\in U_n} q^{neg(\tau)} t^{nsum(\tau)} \sum_{\pi\in S_n} (q^2t)^{maj(\pi)}.
\end{aligned}$$

We know according to Theorem 2.9 that: $\sum_{\pi\in S_n} (q^2t)^{maj(\pi)} = [n]_{q^2t}!$, and by calculation we get:

$$\begin{aligned}
a_n = \sum_{\tau \in U_n} q^{neg(\tau)} t^{nsum(\tau)} &= \sum_{\tau \in U_{n1}} q^{neg(\tau)} t^{nsum(\tau)} + \sum_{\tau \in U_{n2}} q^{neg(\tau)} t^{nsum(\tau)} \\
&= \sum_{\tau' \in U_{n-1}} q^{neg(\varphi_{n1}(\tau'))} t^{nsum(\varphi_{n1}(\tau'))} \\
&+ \sum_{\tau' \in U_{n-1}} q^{neg(\varphi_{n2}(\tau'))} t^{nsum(\varphi_{n2}(\tau'))} \\
&= \sum_{\tau' \in U_{n-1}} q^{neg(\tau')+1} t^{nsum(\tau')+n} + \sum_{\tau' \in U_{n-1}} q^{neg(\tau')} t^{nsum(\tau')} \\
&= qt^n a_{n-1} + a_{n-1} = (1+qt^n) a_{n-1}.
\end{aligned}$$

We got the recurrence equation: $a_n = (1+qt^n)a_{n-1},\ a_1 = 1+qt$, and the solution to this equation is: $a_n = \prod_{i=1}^{n}(1+qt^i)$, and therefore; the general solution is:

$$\sum_{\sigma \in B_n} q^{flag-major_N(\sigma)} t^{nmaj_N(\sigma)} = [n]_{q^2 t}! \prod_{i=1}^{n} (1+qt^i)$$

□

NOTE 3.12. Notice that substituting $t = 1$ in Theorem 3.11, we get Theorem 2.24 and the equation: $[n]_{q^2}!(1+q)^n = \prod_{i=1}^{n}[2i]_q$.

We can also calculate the generating function of *length* and *flag major index* by using a similar method:

THEOREM 3.13.

$$\sum_{\sigma \in B_n} q^{flag-major_N(\sigma)} t^{\ell(\sigma)} = A_n(q^2, t) \prod_{i=1}^{n} (1+qt^i),$$

where $A_n(q,t) = \sum_{\pi \in S_n} q^{maj(\pi)} t^{\ell(\pi)} = \sum_{\pi \in S_n} q^{maj(\pi)} t^{inv(\pi)}$.

3.2. Flag-Excedance and Flag-Denerts Statistic. In this subsection we present the *flag-Denert's* statistic (denoted $fden$) and the *flag-excedance* (denoted $fexc$) statistic. We prove that the pair of statistics $(fden, fexc)$ are equidistributed with $(flag - major, fdes)$ over B_n and, therefore, the *flag-Denert* and *flag-excedance* statistics gives a type B generalization to the Foata-Zeilberger Theorem 2.15.

DEFINITION 3.14. Define the type b *excedance* number of $\sigma \in B_n$ to be:

$$exc_B(\sigma) := |\{\ 1 \le i \le n \mid i < |\sigma(i)|\ \}|.$$

DEFINITION 3.15. Define the *flag-excedance* of $\sigma \in B_n$ to be:

$$fexc(\sigma) := 2exc_B(\sigma) + \varepsilon(\sigma).$$

DEFINITION 3.16. Let n be a nonnegative integer. Define the following subset of B_n:

$$Color_2^n := \{\sigma \in B_n | \ \sigma(i) = \pm i,\ \forall i \in [1,n]\}.$$

NOTE 3.17. Notice that each $\sigma \in B_n$ has a unique representation as:

$$\sigma = \pi\tau,\ where\ \pi \in S_n, \tau \in Color_2^n.$$

DEFINITION 3.18. We define the friends order to be:

$$F : -1 < 1 < -2 < 2 < ... < -n < n.$$

We prove that the *flag-excedance* statistics is Eulerian:

THEOREM 3.19.

$$\sum_{\sigma \in B_n} q^{fexc(\sigma)} = \sum_{\sigma \in B_n} q^{fdes_F(\sigma)}.$$

We define the type B Denert's statistic (denoted den_B):

DEFINITION 3.20. Let $\sigma \in B_n$. Define the type B Denert's statistic to be:

$$\begin{aligned} den_B(\sigma) &= |\{1 \le l < k \le n \mid |\sigma(k)| < |\sigma(l)| < k\}| \\ &+ |\{1 \le l < k \le n \mid |\sigma(l)| < k < |\sigma(k)|\}| \\ &+ |\{1 \le l < k \le n \mid k < |\sigma(k)| < |\sigma(l)|\}|. \end{aligned}$$

REMARK 3.21. According to the definition of den_B we can see that:

$$den_B(\sigma) = den_B(\tau\pi) = den_B(\pi),\ \forall\sigma \in B_n,\ \tau \in Color_2^n,\ \pi \in S_n.$$

We define the *flag-Denert's* statistic (denoted $fden_B$), and prove that it is equidistributed with the *flag major index* over the *signed permutations groups*:

DEFINITION 3.22. Let $\sigma \in B_n$. Define the *flag-Denert's* statistic to be:

$$fden(\sigma) := 2den_B(\sigma) + neg(\sigma).$$

THEOREM 3.23.

$$\sum_{\sigma\in B_n} q^{fden(\sigma)} = \sum_{\sigma\in B_n} q^{flag-major_F(\sigma)}.$$

We prove that the pair of statistics (*fden*,*fexc*) is equidistributed with (*flag-major*,*fdes*).

THEOREM 3.24.

$$\sum_{\sigma\in B_n} q^{fden(\sigma)} t^{fexc(\sigma)} = \sum_{\sigma\in B_n} q^{flag-major_F(\sigma)} t^{fdes_F(\sigma)}.$$

PROOF. (Sketch, more detailed proof can be found at [**8**]) We use the Definitions 3.15, 3.22, [**8**, Lemma 6.4], and Theorem 2.15 and conclude the following equality:

$$\begin{aligned} \sum_{\sigma\in B_n} q^{fden(\sigma)} t^{fexc(\sigma)} &= \sum_{\sigma\in B_n} q^{2den_B(\sigma)+neg(\sigma)} t^{2exc_B(\sigma)+\varepsilon(\sigma)} \\ &= \sum_{\tau\in Color_2^n} q^{neg(\tau)} t^{\varepsilon(\tau)} \sum_{\pi\in S_n} q^{2den(\pi)} t^{2exc(\pi)} \\ &= \sum_{\tau\in Color_2^n} q^{neg(\tau)} t^{\varepsilon(\tau)} \sum_{\pi\in S_n} q^{2maj(\pi)} t^{2des(\pi)} \\ &= \sum_{\sigma\in B_n} q^{flag-major_F(\sigma)} t^{fdes_F(\sigma)} \end{aligned}$$

□

Remark. Extensions to wreath products and more results may be found in [**8**].

4. Acknowledgments

This paper is a part of the authors Master Thesis, which was written at Bar-Ilan University under the supervision of R. M. Adin and Y. Roichman.

References

[1] R. M. Adin, F. Brenti and Y. Roichman, *Descent numbers and major indices for the hyperoctahedral group*, Special issue in honor of Dominique Foata's 65th birthday (Philadelphia, PA 2000), Adv. Appl. Math. **27** (2001), 210–224.

[2] R. M. Adin, I. M. Gessel and Y. Roichman *Signed Mahonians*, J. Combin. Theory (ser. A) **109** (2005), 25–43

[3] R. M. Adin and Y. Roichman, *The flag major index and group actions on polynomial rings*, Europ. J. Combin. **22** (2001), 431–446.

[4] E. Bagno, *Euler-Mahonian parameters on colored permutations groups*, Seminaire Lotharingien de Combinatoire **51** (2004), Article B51f .

[5] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics, Springer-Verlag, to appear.

[6] M. Denert, *The genus zeta function of hereditary orders in central simple algebras over global fields*, Math. Comp. **54** (1990), 449-465.

[7] J. Désarmenien and D. Foata, *The signed Eulerian numbers*, Discrete Math. **99** (1992),49 -58.

[8] M. Fire, *Statistics on Wreath Products*, Master's Thesis, arXiv:math.CO/0409421.

[9] D. Foata and M. P. Schützenberger, *Major index and inversion number of permutations*, Math. Nachr. **83** (1978), 143–159.

[10] D. Foata and D. Zeilberger, *Denert's permutation statistic is indeed Euler-Mahonian*, Studies in Appl. Math. **83** (1990), 31-59.

[11] I. Gessel, *Generating Functions and Enumeration of Sequences*, Ph. D. Thesis, MIT, (1977).

[12] J. E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Studies in Advanced Mathematics, no. 29, Cambridge Univ. Press, Cambridge, 1990.

[13] P. A. MacMahon, *Combinatory Analysis I-II*, Cambridge Univ. Press, London/New-York, 1916. (Reprinted by Chelsea, New-York, 1960.)

[14] M. Wachs, *An involution for signed Eulerian numbers*, Discrete Math. **99** (1992), 59–62.

Department of Mathematics Bar Ilan University Ramat Gan 52900, Israel
E-mail address: mickyfi@gmail.com

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

New results on the combinatorial invariance of Kazhdan-Lusztig polynomials

Federico Incitti

ABSTRACT. We prove that the Kazhdan-Lusztig polynomials are combinatorial invariants for intervals up to length 8 in Coxeter groups of type **A** and up to length 6 in Coxeter groups of type **B** and **D**. As a consequence of our methods, we also obtain a complete classification, up to isomorphism, of Bruhat intervals of length 7 in type **A** and of length 5 in types **B** and **D**, which are not lattices.

RÉSUMÉ. On montre que les polynômes de Kazhdan-Lusztig sont invariants combinatoires pour les intervaux de longueur jusqu'à 8 pour les groupes de Coxeter de type **A** et de longueur jusqu'à 6 pour les groupes de Coxeter de type **B** et **D**. Comme conséquence de nos méthodes, on obtient aussi une classification complète, à isomorphisme près, des intervaux de Bruhat de longueur 7 pour le type **A** et de longueur 5 pour les types **B** et **D**, qui ne sont pas des réseaux.

1. Introduction

In [**12**] Kazhdan and Lusztig defined, for every Coxeter group W, a family of polynomials, indexed by pairs of elements of W, which have become known as the Kazhdan-Lusztig polynomials of W. They are related to the algebraic geometry and topology of Schubert varieties, and also play a crucial role in representation theory (see, e.g., [**7**, Chapter 7], [**1**, Chapter 5]). In order to prove the existence of these polynomials, Kazhdan and Lusztig used another family of polynomials which arise from the multiplicative structure of the Hecke algebra associated with W. These are known as the R-polynomials of W. Lusztig's and Dyer's combinatorial invariance conjecture states that the Kazhdan-Lusztig polynomial associated with a pair (x, y) supposedly only depends on the poset structure of the Bruhat interval $[x, y]$. The conjecture is equivalent to the same statement for the R-polynomials and it is known to hold for intervals up to length 4. In [**10**] we proved that the conjecture is true for intervals of length 5 and 6 in Coxeter groups of type **A**.

In this paper, we establish the conjecture for intervals of length 7 and 8 in Coxeter groups of type **A** and for those of length 5 and 6 in Coxeter groups of type **B** and **D**. We use the combinatorial descriptions of such groups in terms of (signed) permutations (see, e.g., [**1**, Chapter 8]). One of the main tools is an extension of the notion of diagram of a pair, introduced for the symmetric group by Kassel et al. in [**11**] and developed in [**9**], to the groups of signed permutations. The main idea behind the proof is that of determining certain subsets of pairs of (signed) permutations, which somehow "summarize" the behaviour of all the pairs. The combinatorial invariance is then proved by enumerating all the pairs in these sets, with the assistance of Maple computation, and for each of them determining the poset structure of the associated interval and computing the corresponding R-polynomial. As a consequence of our methods, we also obtain a complete classification, up to isomorphism, of Bruhat intervals of length 7 in type **A** and of length 5 in types **B** and **D**, which are not lattices (see [**2, 3, 6, 10**] for previous classification results).

2000 *Mathematics Subject Classification.* 05E15, 20F55.

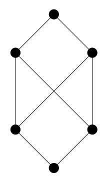
Key words and phrases. Classical Weyl group, Bruhat order, Kazhdan-Lusztig polynomial, R-polynomial, combinatorial invariance conjecture.

Research supported by the Swedish Institute of Stockholm and by the Istituto Nazionale di Alta Matematica of Rome.

2. Preliminaries

Let $\mathbf{N} = \{1, 2, \ldots\}$ and $\mathbf{Z}$ be the set of integers. For $n, m \in \mathbf{Z}$, with $n \leq m$, let $[n, m] = \{n, n+1, \ldots, m\}$. For $n \in \mathbf{N}$, let $[n] = [1, n]$, $[-n] = [-n, -1]$ and $[\pm n] = [n] \cup [-n]$. We refer to [**13**] for general poset theory. Given a poset P, we denote by $\lhd$ the covering relation. Given $x, y \in P$, with $x < y$, we set $[x, y] = \{z \in P : x \leq z \leq y\}$, and call it an *interval* of P. We denote by $-P$ the poset *dual* to P, that is, the poset having the same elements of P but the reverse order.

We refer to [**1**] for basic notions about Coxeter groups. Given a Coxeter group W, with set of generators S, the set of *reflections* of W is $T = \{wsw^{-1} : w \in W,\ s \in S\}$. Given $x \in W$, the *length* of x, denoted by $\ell(x)$, is the minimal k such that x is the product of k generators. The *Bruhat graph* of W, denoted by $BG(W)$ is the directed graph having W as vertex set and such that there is an edge $x \to y$ if and only if $y = xt$, with $t \in T$, and $\ell(x) < \ell(y)$. If this happens, we label the edge (x, y) by the reflection t and write $x \xrightarrow{t} y$. A *Bruhat path* is a (directed) path in the Bruhat graph of W. The *Bruhat order* of W is the partial order induced by $BG(W)$: given $x, y \in W$, $x \leq y$ in the Bruhat order if and only if there is a Bruhat path from x to y. Every Coxeter group W, partially ordered by the Bruhat order, is a graded poset with rank function given by the length. For $x, y \in W$, with $x < y$, we set $\ell(x, y) = \ell(y) - \ell(x)$ and call it the *length* of the pair (x, y). In [**9**] we introduced the *absolute length* of the pair (x, y), denoted by $a\ell(x, y)$, which is the (directed) distance from x to y in $BG(W)$. If $\ell(x, y) = 3$, then it is known that $x \xrightarrow{t} y$ if and only if the interval $[x, y]$ is isomorphic to the 2-crown, that is, the poset whose Hasse diagram is the following:



Finally, if W is finite then it has a maximum, denoted by w_0. The maps $x \mapsto x^{-1}$ and $x \mapsto w_0 x w_0$ are automorphisms of the Bruhat order, while the maps $x \mapsto x w_0$ and $x \mapsto w_0 x$ are antiautomorphisms.

We refer to [**1**, §5.2] for basic notions about reflection orderings, which are total orderings on the set T of reflections with certain properties. We only recall that, if W is finite and $s_1 s_2 \ldots s_m$ is a reduced decomposition of w_0, then a possible reflection ordering is $t_1 \prec t_2 \prec \cdots \prec t_m$, where $t_i = s_m \ldots s_{i+1} s_i s_{i+1} \ldots s_m$, for all $i \in [m]$. Moreover, all reflection orderings are obtained in this way (see [**1**, Exercise 5.20]).

We follow [**1**, Chapter 5] for the definition of R-polynomials and Kazhdan-Lusztig polynomials of W. There exists a unique family of polynomials $\{R_{x,y}(q)\}_{x,y \in W} \subseteq \mathbf{Z}[q]$ satisfying the following conditions:

(i) $R_{x,y}(q) = 0$, if $x \not\leq y$;
(ii) $R_{x,y}(q) = 1$, if $x = y$;
(iii) if $x < y$ and $s \in S$ is such that $ys \lhd y$ then

$$R_{x,y}(q) = \begin{cases} R_{xs,ys}(q), & \text{if} \quad xs \lhd x, \\ qR_{xs,ys}(q) + (q-1)R_{x,ys}(q), & \text{if} \quad xs \rhd x. \end{cases}$$

These are known as the *R-polynomials* of W. The existence of such a family is a consequence of the invertibility of certain basis elements of the Hecke algebra $\mathcal{H}$ of W and is proved in [**7**, §§7.4, 7.5]. Then, there exists a unique family of polynomials $\{P_{x,y}(q)\}_{x,y \in W} \subseteq \mathbf{Z}[q]$ satisfying the following conditions:

(i) $P_{x,y}(q) = 0$, if $x \not\leq y$;
(ii) $P_{x,y}(q) = 1$, if $x = y$;
(iii) if $x < y$ then $\deg(P_{x,y}(q)) < \ell(x, y)/2$ and

$$q^{\ell(x,y)}\, P_{x,y}\left(q^{-1}\right) - P_{x,y}(q) = \sum_{x < z \leq z} R_{x,z}(q)\, P_{z,y}(q).$$

These are known as the *Kazhdan-Lusztig polynomials* of W. The existence of such a family is proved in [**7**, §§7.9, 7.10, 7.11]. We also need the following property of the R-polynomials (see [**1**, Exercise 5.11]):

$$\sum_{x \leq z \leq y} (-1)^{\ell(x,z)} R_{x,z}(q) R_{z,y}(q) = 0. \tag{1}$$

Finally, there exists a unique family of polynomials $\{\widetilde{R}_{x,y}(q)\}_{x,y\in W} \in \mathbf{Z}_{\geq 0}[q]$ such that

$$R_{x,y}(q) = q^{\ell(x,y)/2}\widetilde{R}_{x,y}(q^{1/2} - q^{-1/2})$$

for all $x, y \in W$. These are known as the *$\widetilde{R}$-polynomials* of W and their coefficients have a nice combinatorial interpretation in terms of reflection orderings. Given $x, y \in W$, with $x < y$, we denote by $BP(x, y)$ the set of all Bruhat paths from x to y. The *length* of $\Delta = (x_0, x_1, \dots, x_k) \in BP(x, y)$, denoted by $|\Delta|$, is the number k of its edges. Let $\prec$ be a fixed reflection ordering on the set T of reflections. A path $\Delta = (x_0, x_1, \dots, x_k) \in BP(x, y)$, with

$$x_0 \xrightarrow{t_1} x_1 \xrightarrow{t_2} \cdots \xrightarrow{t_k} x_k,$$

is said to be *increasing* with respect to $\prec$ if $t_1 \prec t_2 \prec \cdots \prec t_k$. We denote by $BP^{\prec}(x, y)$ the set of all paths in $BP(x, y)$ which are increasing with respect to $\prec$. Then, we have the following (see [**1**, Theorem 5.3.4]):

$$\widetilde{R}_{x,y}(q) = \sum_{\Delta \in BP^{\prec}(x,y)} q^{|\Delta|}. \tag{2}$$

More precisely, set $\ell = \ell(x, y)$ and $a\ell = a\ell(x, y)$, the following holds (see [**4**] and [**9**, Corollary 2.6]):

$$\widetilde{R}_{x,y}(q) = q^{\ell} + c_{\ell-2}\, q^{\ell-2} + \cdots + c_{a\ell+2}\, q^{a\ell+2} + c_{a\ell}\, q^{a\ell}, \tag{3}$$

where $c_k = |\{\Delta \in BP^{\prec}(x, y) : |\Delta| = k\}| \geq 1$, for all $k \in [a\ell, \ell - 2]$, with $k \equiv \ell \pmod 2$. Finally, by results in [**3**] and [**5**], we have that the absolute length of a pair is a combinatorial invariant, that is, $a\ell(x, y)$ only depends on the poset structure of the interval $[x, y]$.

We now briefly recall some basic facts about Bruhat order in classical Weyl groups, that is, Coxeter groups of type $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{D}$, following [**1**, Chapter 8]. We denote by S_n the *symmetric group* over n elements. To denote a permutation $x \in S_n$ we use the *one-line notation*: we write $x = x_1 x_2 \dots x_n$ to mean that $x(i) = x_i$ for all $i \in [n]$. The symmetric group S_n is a Coxeter group of type $\mathbf{A}_{n-1}$, with generators given by the simple transpositions $(i, i + 1)$, for $i \in [n - 1]$. We recall that, given $x \in S_n$, a *free rise* of x is a pair $(i, j) \in \mathbf{N}^2$, with $i < j$ and $x(i) < x(j)$, such that there is no $k \in \mathbf{N}$, with $i < k < j$ and $x(i) < x(k) < x(j)$. Given $x, y \in S_n$, then $x \lhd y$ in the Bruhat order if and only if $y = x(i, j)$, where (i, j) is a free rise of x. Following [**8**], if this happen we write $y = ct_{(i,j)}(x)$ and $x = ict_{(i,j)}(y)$, where ct stands for *covering transformation* and ict for *inverse covering transformation*.

We denote by B_n the *hyperoctahedral group*, defined by

$$B_n = \{x : [\pm n] \to [\pm n] : x \text{ is a bijection}, \ x(-i) = -x(i) \text{ for all } i \in [n]\}.$$

and call its elements *signed permutations*. To denote a signed permutation $x \in B_n$ we use the *window notation*: we write $x = [x_1, x_2, \dots, x_n]$, to mean that $x(i) = x_i$ for all $i \in [n]$ (the images of the negative entries are then uniquely determined). We also denote x by the sequence $|x_1|\,|x_2| \dots |x_n|$, with the negative entries underlined. For example, $\underline{3}\,\underline{2}\,1$ denotes the signed permutation $[-3, -2, 1]$. As a set of generators for B_n, we take $S = \{s_0, s_1, \dots, s_{n-1}\}$, where $s_0 = (1, -1)$ and $s_i = (i, i + 1)(-i, -i - 1)$ for all $i \in [n - 1]$. The hyperoctahedral group B_n, with this set of generators, is a Coxeter group of type $\mathbf{B}_n$. Let $x \in B_n$. A rise (i, j) of x is *central* if $(0, 0) \in [i, j] \times [x(i), x(j)]$. A central rise (i, j) of x is *symmetric* if $j = -i$. Then, we have the following characterization of the covering relation in the Bruhat order of B_n (see [**8**, Theorem 5.5]). Let $x, y \in B_n$. Then $x \lhd y$ if and only if either (i) $y = x(i, j)(-i, -j)$, where (i, j) is a noncentral free rise of x, or (ii) $y = x(i, j)$, where (i, j) is a central symmetric free rise of x. In both cases we write $y = ct_{(i,j)}(x)$ and $x = ict_{(i,j)}(y)$. The maximum of B_n is $w_0 = \underline{1}\,\underline{2} \dots \underline{n}$.

We denote by D_n the *even-signed permutation group*, defined by

$$D_n = \{x \in B_n : neg(x) \text{ is even}\}.$$

Notation and terminology are inherited from the hyperoctahedral group. As a set of generators for D_n, we take $S = \{s_0, s_1, \dots, s_{n-1}\}$, where $s_0 = (1, -2)(-1, 2)$ and $s_i = (i, i + 1)(-i, -i - 1)$ for all $i \in [n - 1]$. The even-signed permutation group D_n, with this set of generators, is a Coxeter group of type $\mathbf{D}_n$. Let $x \in D_n$. A central rise (i, j) of x is *semi-free* if $\{k \in [i, j] : x(k) \in [x(i), x(j)]\} = \{i, -j, j\}$. Then, for $x, y \in D_n$, we have (see [**8**, Theorem 6.7]) $x \lhd y$ if and only if $y = x(i, j)(-i, -j)$, where (i, j) is (i) a noncentral free rise of x, or (ii) a central nonsymmetric free rise of x, or (iii) a central semi-free rise of x. In all cases we write $y = ct_{(i,j)}(x)$ and $x = ict_{(i,j)}(y)$. The maximum of D_n is $w_0 = \underline{1}\,\underline{2} \dots \underline{n}$ if n is even, $1\,\underline{2} \dots \underline{n}$ if n is odd.

3. Main tools

3.1. Diagram of a pair of (signed) permutations. Let $W \in \{S_n, B_n, D_n\}$. For convenience, we set $\langle n\rangle = [n]$ if $W = S_n$ and $\langle n\rangle = [\pm n]$ if $W \in \{B_n, D_n\}$. The *diagram* of a (signed) permutation $x \in W$ is the subset of $\mathbf{Z}^2$ defined by

$$\mathrm{Diag}(x) = \{(i, x(i)) : i \in \langle n\rangle\}.$$

For $x \in W$ and $(h,k) \in \langle n\rangle^2$, we set

$$x[h,k] = |\{i \in \langle n\rangle : i \le h,\ x(i) \ge k\}| \tag{4}$$

and given $x, y \in W$ and $(h,k) \in \langle n\rangle^2$, we set

$$(x,y)[h,k] = y(h,k) - x(h,k) \tag{5}$$

There are well-known characterizations of the Bruhat order in S_n and B_n (see [**1**, Theorems 2.1.5, 8.1.8]), which can be stated as follows: if $W \in \{S_n, B_n\}$ and $x, y \in W$ then

$$x \le y \quad \Leftrightarrow \quad (x,y)[h,k] \ge 0, \quad \text{for all } (h,k) \in \langle n\rangle^2.$$

See [**1**, Theorem 8.2.8] for a combinatorial characterization of the Bruhat order relation in D_n. Here we only recall that if $x, y \in D_n$ then only one implication is true:

$$x \le y \quad \Rightarrow \quad (x,y)[h,k] \ge 0, \quad \text{for all } (h,k) \in \langle n\rangle^2.$$

For our purposes, it is convenient to extend the definitions given in (4) and (5) to every $(h,k) \in \mathbf{R}^2$. We call the mapping $(h,k) \mapsto (x,y)[h,k]$ the *multiplicity mapping* of the pair (x,y). Then, the *diagram* of the pair (x,y) is the collection of: (i) the diagram of x, (ii) the diagram of y and (iii) the multiplicity mapping of (x,y). From the preceding considerations, if $x \le y$, then the values of this mapping are always non-negative. In this case, we pictorially represent the diagram of a pair (x,y) with the following convention: the diagrams of x and y are denoted by black and white dots, respectively, and the mapping $(h,k) \mapsto (x,y)[h,k]$ is represented by colouring the preimages of different positive integers with different levels of grey, with the rule that a lighter grey corresponds to a lower integer. Examples for the symmetric group can be found in [**9**]. In Figure 1, the diagram of (x,y), where $x = 2341$ and $y = 3\underline{4}2\underline{1} \in B_4$, is illustrated. Note that, although $x, y \in D_4$, we have $x \not\le y$ in D_4, since condition (ii) of [**1**, Theorem 8.2.8] fails for $(a,b) = (2,1)$. Figure 2 shows the diagram of (x,y), where $x = 1342$ and $y = 3\underline{4}1\underline{2} \in D_4$. Now, $x \le y$ in D_4.

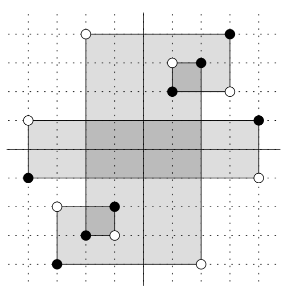


Figure 1: Diagram of a pair in B_n.

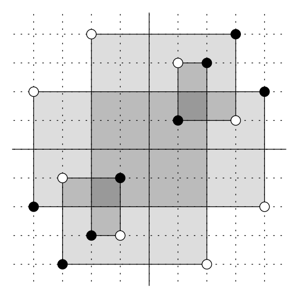


Figure 2: Diagram of a pair in D_n.

The *support* of (x,y) is

$$\Omega(x,y) = \{(h,k) \in \mathbf{R}^2 : (x,y)[h,k] > 0\}$$

and the *support index set* of (x,y) is

$$I_\Omega(x,y) = \{i \in \langle n\rangle : (i, x(i)) \in \overline{\Omega(x,y)}\},$$

where $\overline{\Omega(x,y)}$ denotes the (topological) closure of the set $\Omega(x,y)$. A pair $(x,y) \in W^2$, with $x < y$, is said to have *full support* if $I_\Omega(x,y) = \langle n\rangle$. For instance, both the pairs in Figures 1 and 2 have full support.

3.2. Computing $\widetilde{R}$-polynomials. In [**9**] we described an algorithm for computing $\widetilde{R}$-polynomials in the symmetric group. Following a similar strategy, $\widetilde{R}$-polynomials can be efficiently computed in the groups of signed permutations starting from equation (2), by choosing convenient reflection orderings.

We recall that, if we set $T_1 = \{(i,j)(-i,-j) : i \in [-n], j \in [\pm(-i-1)]\}$ and $T_2 = \{(i,-i) : i \in [-n]\}$, then, the set of reflections in D_n is T_1 (see, e.g., [**1**, Prop. 8.1.5]) and the set of reflections in B_n is $T_1 \cup T_2$ (see, e.g., [**1**, Prop. 8.2.5]). In both B_n and D_n we identify the reflection $(i,j)(-i,-j) \in T_1$, where $i \in [-n]$ and $j \in [\pm(-i-1)]$, with the pair (i,j). Then, we have the following.

Proposition 3.1. *A possible reflection ordering in D_n is the lexicographic order between pairs. And a possible reflection ordering in B_n is the same as in D_n, with the reflection $(i,-i)$ inserted between $(i,-1)$ and $(i,1)$, for all $i \in [-n,-2]$, and $(-1,1)$ inserted as the last one.*

Proof. They arise from appropriate choices of a reduce decomposition of the maximum element w_0. $\square$

For example, a reflection ordering in D_4 is

$$\begin{gathered}(-4,-3) \prec (-4,-2) \prec (-4,-1) \prec (-4,1) \prec (-4,2) \prec (-4,3) \prec \\ (-3,-2) \prec (-3,-1) \prec (-3,1) \prec (-3,2) \prec \\ (-2,-1) \prec (-2,1),\end{gathered}$$

and a reflection odering in B_4 is

$$\begin{gathered}(-4,-3) \prec (-4,-2) \prec (-4,-1) \prec (-4,4) \prec (-4,1) \prec (-4,2) \prec (-4,3) \prec \\ (-3,-2) \prec (-3,-1) \prec (-3,3) \prec (-3,1) \prec (-3,2) \prec \\ (-2,-1) \prec (-2,2) \prec (-2,1) \prec \\ (-1,1).\end{gathered}$$

3.3. Symmetries. Let W be any finite Coxeter group and let w_0 be its maximum. We define the following equivalence relations between pairs $(x,y) \in W^2$, with $x < y$:

$$\begin{aligned}(x_1,y_1) \sim^+ (x,y) \quad &\Leftrightarrow \quad (x_1,y_1) \in \{(x,y),(x^{-1},y^{-1}),(w_0xw_0,w_0yw_0),(w_0x^{-1}w_0,w_0y^{-1}w_0)\} \\ (x_1,y_1) \sim^- (x,y) \quad &\Leftrightarrow \quad (x_1,y_1) \in \{(yw_0,xw_0),(w_0y,w_0x),(y^{-1}w_0,x^{-1}w_0),(w_0y^{-1},w_0x^{-1})\} \\ (x_1,y_1) \sim (x,y) \quad &\Leftrightarrow \quad (x_1,y_1) \sim^+ (x,y) \quad \text{or} \quad (x_1,y_1) \sim^- (x,y)\end{aligned}$$

Then, it is known that

$$\begin{aligned}(x_1,y_1) \sim^+ (x,y) \quad &\Rightarrow \quad [x_1,y_1] \cong [x,y] \\ (x_1,y_1) \sim^- (x,y) \quad &\Rightarrow \quad [x_1,y_1] \cong -[x,y]\end{aligned}$$

Moreover (see, e.g., [**1**, Exercise 4.10]) we have

$$(x_1,y_1) \sim (x,y) \quad \Rightarrow \quad \widetilde{R}_{x_1,y_1}(q) = \widetilde{R}_{x,y}(q).$$

In classical Weyl groups, if $(x_1,y_1) \sim (x,y)$ then the diagram of (x_1,y_1) is obtained from that of (x,y) by a certain reflection, as described for the symmetric group in [**10**, Figure 2]. The only exception is the case $W = D_n$ and n odd when, for example, $xw_0 = [x(1), -x(2), \ldots, -x(n)]$. Then, in order to generate all possible intervals and $\widetilde{R}$-polynomials, we will consider diagrams up to these symmetries.

3.4. Odd signed permutation poset. In the remainder of the paper we will act on diagrams by "deleting" or "inserting" dots. In the groups of type **D**, this would not always be allowed, because of the restriction on the parity of the number of negative entries. In this subsection we present a way of bypassing this problem. We start with defining the *odd-signed permutation set*:

$$D_n^{\mathrm{odd}} = \{x \in B_n : neg(x) \text{ is odd}\}.$$

Although D_n^{odd} is not a group, we can still define on it the Bruhat order as in D_n, giving the same characterization of the covering relation (see the end of Section 2). More precisely, given $x, y \in D_n^{\mathrm{odd}}$, we say that $x \lhd y$ in the *Bruhat order* if and only if $y = x(i,j)(-i,-j)$, where (i,j) is (i) a noncentral free rise of x, or (ii) a central nonsymmetric free rise of x, or (iii) a central semifree rise of x. Then, we have the following.

Proposition 3.2. *The map $\varphi : D_n \to D_n^{\mathrm{odd}}$ defined by*

$$[x_1, x_2, \ldots, x_n] \stackrel{\varphi}{\longmapsto} [-x_1, x_2, \ldots, x_n]$$

is an isomorphism of posets.

By Proposition 3.2, whose proof is omitted, working with the posets D_n and D_n^{odd} is essentially the same thing. From now on, we will denote the even-signed permutation group by D_n^{even} and we will write $x, y \in D_n$ to mean either $x, y \in D_n^{\mathrm{even}}$ or $x, y \in D_n^{\mathrm{odd}}$, with the only requirement that $neg(x) \equiv neg(y) \pmod 2$.

3.5. Simplifications. Let $W \in \{S_n, B_n, D_n\}$. An *index set* is a subset $I \subseteq \langle n \rangle$, such that $I = -I$ if $W \in \{B_n, D_n\}$. Let $x \in W$ and I be an index set. We denote by $x\,|_I$ the (signed) permutation whose diagram is obtained from that of x, by considering only the dots corresponding to the indices in I, removing the others, and renumbering the remaining indices and values. We call $x\,|_I$ the *subpermutation* of x induced by I. We start with noting that all the information about the poset structure of $[x, y]$ and about the $\widetilde{R}$-polynomial associated is contained in the support of (x, y).

Proposition 3.3. *Let $x, y \in W$, with $x < y$. Set $x_\Omega = x\,\big|_{I_\Omega(x,y)}$ and $y_\Omega = y\,\big|_{I_\Omega(x,y)}$. Then*

(i) $[x, y] \cong [x_\Omega, y_\Omega]$;

(ii) $\widetilde{R}_{x_\Omega, y_\Omega}(q) = \widetilde{R}_{x,y}(q)$.

Proof. For the symmetric group, it has been proved in [**9**, Proposition 5.2] and [**10**, Proposition 3.1]. For the groups B_n and D_n, the characterization of the covering relation in terms of rises ensures that the interval $[x, y]$ reflects a process of "unmounting" the diagram of (x, y) similar to that described in [**9**] for the symmetric group and (i) follows. A similar consideration together with equation (2) implies (ii). $\square$

It is useful to introduce the following notion of Ω-*equivalence* between pairs:

$$(x', y') \sim_\Omega (x, y) \quad \Leftrightarrow \quad (x'_\Omega, y'_\Omega) = (x_\Omega, y_\Omega).$$

According to Proposition 3.3, the same interval (up to poset isomorphism) and the same $\widetilde{R}$-polinomial are associated with all the pairs in an Ω-equivalence class.

Now, let $x \in W$ and I be an index set. For $(h, k) \in \mathbf{R}^2$, we set

$$x[h, k]\,|_I = |\{i \in I : i \leq h,\, x(i) \geq k\}|,$$

Let $x, y \in W$ and I be an index set such that $x(I) = y(I)$. For $(h, k) \in \mathbf{R}^2$, we set

$$(x, y)[h, k]\,|_I = y(h, k)\,|_I - x(h, k)\,|_I \,.$$

Then, we set

$$\Omega(x, y)\,|_I = \{(h, k) \in \mathbf{R}^2 : (x, y)[h, k]\,|_I > 0\}$$

Definition 3.4. Let $x, y \in W$, with $x < y$. Let I_1 and I_2 be two index sets, with $I_\Omega(x, y) = I_1 \cup I_2$ and $I_1 \cap I_2 = \emptyset$, such that $x(I_1) = y(I_1)$ and $x(I_2) = y(I_2)$. Set $x_r = x\,|_{I_r}$, $y_r = y\,|_{I_r}$ and $\Omega_r = \Omega(x, y)\,|_{I_r}$, for $r = 1, 2$. Note that, necessarily, $x_1 < y_1$ and $x_2 < y_2$. We say that the pair (x, y) is *trivially decomposable* into the two pairs (x_1, y_1) and (x_2, y_2) if Ω_1 and Ω_2 are either disjoint or if they intersect in a region whose closure does not contain any of the dots of the diagrams of x and y.

For example, the pair $(x, y) \in B_4^2$, whose diagram is shown in Figure 1, is trivially decomposable into the two pairs $(123, 2\underline{3}1) \in B_3^2$ and $(1, \underline{1}) \in B_1^2$. We have the following general result.

Proposition 3.5. *Let $x, y \in W$, with $x < y$, be trivially decomposable into (x_1, y_1) and (x_2, y_2). Then*

(i) $[x, y] \cong [x_1, y_1] \times [x_2, y_2]$;

(ii) $\widetilde{R}_{x,y}(q) = \widetilde{R}_{x_1,y_1}(q) \cdot \widetilde{R}_{x_2,y_2}(q)$.

Proof. For S_n, it has been proved in [**9**, Proposition 2.16] and [**10**, Propositions 3.2, 3.4, 3.5]. For B_n and D_n the proof is similar, since under the hypotheses of the proposition, the process of "unmounting" the diagram of (x, y), that the interval $[x, y]$ reflects, is completely independent for Ω_1 and Ω_2 and (i) follows. A similar consideration together with equation (2) implies (ii). $\square$

3.6. Enlarging an interval. In this subsection we show how it is possible, given an interval $[x, y]$, to obtain all intervals of length $\ell(x, y) + 1$ containing $[x, y]$ as subinterval, in terms of the diagram of (x, y). We start with introducing a notion of "insertion" of a dot in a diagram.

Definition 3.6. Let $x \in S_n$ and $h, k \in [n + 1]$. The permutation obtained from x by *inserting the dot* (h, k), denoted by $x^{(h,k)}$, is the only permutation $\widehat{x} \in S_{n+1}$ satisfying (i) $\widehat{x}(h) = k$, (ii) $\widehat{x}\,\big|_{[n+1]\setminus\{h\}} = x$.

Similarly, for $x \in B_n$ (resp. D_n), $h \in [n + 1]$ and $k \in [\pm(n + 1)]$, the signed permutation obtained from x by *inserting the dot* (h, k), denoted by $x^{(h,k)}$, is the only permutation $\widehat{x} \in B_{n+1}$ (resp. D_{n+1} or D^{odd}_{n+1}, depending on whether $k > 0$ or $k < 0$) satisfying (i) $\widehat{x}(h) = k$ (thus $\widehat{x}(-h) = -k$), (ii) $\widehat{x}\,\big|_{[\pm(n+1)]\setminus\{h,-h\}} = x$.

If we consider the pairs (x, y) whose diagrams are shown in Figures 1 and 2, then the diagrams of the pairs $\big(x^{(3,-3)}, y^{(3,-3)}\big)$ are illustrated in Figures 3 and 4, respectively. Note that the signed permutations $x^{(3,-3)}, y^{(3,-3)}$ shown in Figure 4 belong to D^{odd}_5 and, according to the considerations following Proposition 3.2, we still write $\big(x^{(3,-3)}, y^{(3,-3)}\big) \in D_5$.

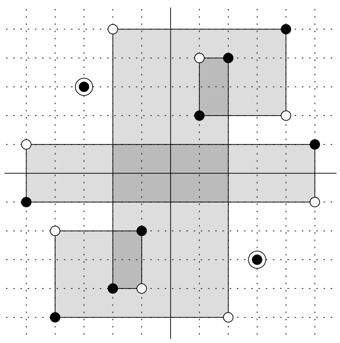


Figure 3: Inserting a dot in B_n.

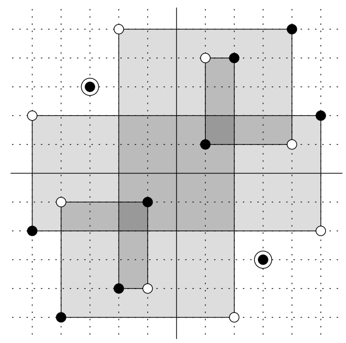


Figure 4: Inserting a dot in D_n.

In particular, we are interested in inserting dots out of the support, as it happens in the diagrams in Figures 3 and 4. In this case we obtain a pair which is in the same Ω-equivalence class as the originary pair. Then, we have the following result, which is an immediate consequence of Proposition 3.3.

Corollary 3.7. *Let $x, y \in W$, with $x < y$, and $(h, k) \in [n + 1] \times \langle n + 1\rangle$, with $(h, k) \notin \overline{\Omega(x^{(h,k)}, y^{(h,k)})}$.*

(i) $[x, y] \cong \big[x^{(h,k)}, y^{(h,k)}\big]$;

(ii) $\widetilde{R}_{x,y}(q) = \widetilde{R}_{x^{(h,k)},y^{(h,k)}}(q)$.

Let $x, y \in W$, with $x < y$. The intervals of length $\ell(x, y) + 1$ containing $[x, y]$ are exactly those of the form $[x, z]$ (if $y \neq w_0$), with $z = ct_{(i,j)}(y)$ and those of the form $[w, y]$ (if $x \neq id$), with $w = ict_{(i,j)}(x)$. In both cases we say that the new pair, (x, z) or (w, y), is *obtained* from (x, y) by

(i) an *external move*, if $\{i, j\} \subseteq \langle n\rangle \setminus I_\Omega(x, y)$;
(ii) an *internal move*, if $\{i, j\} \subseteq I_\Omega(x, y)$;
(iii) an *enlarging move*, if $|\{i, j\} \cap I_\Omega(x, y)| = 1$.

In case (iii), if $\{i, j\}\setminus I_\Omega(x, y) = \{h\}$, then we also say that the enlarging move *uses* the dot $(h, x(h))$. Also, if (x, y) is a pair with full support, (x', y') is any pair Ω-equivalent to (x, y) and (w, z) is obtained from (x', y') by one of the three kinds of moves described, then we say that (w, z) is *obtained* from (x, y) as well.

External moves can be easily managed by the following result.

Proposition 3.8. *Let $x, y \in W$, with $x < y$. Let (w, z) be obtained from (x, y) by an external move and suppose both (x, y) and (w, z) have full support. Then (w, z) is trivially decomposable into (x, y) and a pair (a, b), with $a \lhd b$. In particular* (i) $[w, z] = [x, y] \times \{0, 1\}$; (ii) $\widetilde{R}_{w,z}(q) = q\widetilde{R}_{x,y}(q)$.

We need one last definition.

Definition 3.9. Let $W \in \{S_n, B_n, D_n\}$ and $x, y \in W$, with $x < y$, be such that (x, y) has full support. The *enlarging set* of (x, y), denoted by $\mathrm{Enl}(x, y)$, is the union of all the pairs with full support obtained from (x, y) by internal moves and enlarging moves.

4. Main result

The combinatorial invariance of Kazhdan-Lusztig polynomials for intervals up to a certain length is equivalent to that of the R-polynomials (or their counterpart, the $\widetilde{R}$-polynomials) for the same intervals. We will prove our main result by showing that the $\widetilde{R}$-polynomials are combinatorial invariants. First of all, note that an interval $[x,y]$ does not contain a 2-crown if and only if $a\ell(x,y) = \ell(x,y)$ and, by equation (3), this happens if and only if $\widetilde{R}_{x,y}(q) = q^{\ell(x,y)}$. Thus, we only need to consider intervals containing 2-crowns.

Let $\mathcal{F}_A = \{S_n : n \geq 2\}$, $\mathcal{F}_B = \{B_n : n \geq 1\}$ and $\mathcal{F}_D = \{D_n : n \geq 2\}$.

Definition 4.1. Let $\mathcal{F} \in \{\mathcal{F}_A, \mathcal{F}_B, \mathcal{F}_D\}$. The *essential sets* of $\mathcal{F}$ are recursively defined by

$$ES_3(\mathcal{F}) = \{(x,y) \in W^2 : W \in \mathcal{F},\ [x,y] \text{ is a 2-crown and } (x,y) \text{ has full support}\}/\sim$$

and, for $k \geq 4$

$$ES_k(\mathcal{F}) = \Big[\bigcup \{Enl(x,y) : [(x,y)]_\sim \in ES_{k-1}(\mathcal{F})\} \Big] / \sim .$$

The sets $ES_3(\mathcal{F})$ can be easily determined and they are as follows:

$ES_3(\mathcal{F}_A) = \{(123, 321)\}$,
$ES_3(\mathcal{F}_B) = \{(123,321), (\underline{2}\,\underline{1}3, 3\underline{1}\,\underline{2}), (\underline{3}\,\underline{1}2, 2\underline{1}\,\underline{3}), (\underline{3}\,\underline{2}1, 1\underline{2}\,\underline{3}), (12, 1\underline{2}), (12, \underline{2}\,\underline{1})\}$,
$ES_3(\mathcal{F}_D) = \{(123,321), (\underline{2}\,\underline{1}3, 3\underline{1}\,\underline{2}), (\underline{3}\,\underline{1}2, 2\underline{1}\,\underline{3}), (\underline{3}\,\underline{2}1, 1\underline{2}\,\underline{3}), (123, \underline{3}2\,\underline{1}), (213, \underline{3}1\,\underline{2}), (312, \underline{2}1\,\underline{3})\}$,

where, for simplicity, we have identified every equivalence class with one of its elements.

theorem 4.2. *Let $\mathcal{F} \in \{\mathcal{F}_A, \mathcal{F}_B, \mathcal{F}_D\}$ and $k \geq 3$. The essential set $ES_k(\mathcal{F})$ contains, up to $\sim$, all possible pairs of length k, in Coxeter groups in $\mathcal{F}$, which have full support and are not trivially decomposable, whose corresponding interval $[x,y]$ contains a 2-crown.*

Proof. We proceed by induction on k. For $k=3$, the result is true by definition. Assume $k \geq 4$. It is easy to prove that all the pairs in $ES_k(\mathcal{F})$ have the required properties. Now, let (x,y) be a pair of length k which has full support and is not trivially decomposable, such that $[x,y]$ contains a 2-crown. We want to show that $[(x,y)]_\sim \in ES_k(\mathcal{F})$. As one can easily check, it is always possible to find an atom z (or a coatom w) of $[x,y]$ such that (z,y) (or (x,w)) is still not trivially decomposable and $[z,y]$ (or $[x,w]$) still contains a 2-crown. Let z be an atom of $[x,y]$ with this properties (the case of a coatom w is similar). Now, let $z_\Omega = z\,\big|_{I_\Omega(z,y)}$ and $y_\Omega = y\,\big|_{I_\Omega(z,y)}$. Then, (z_Ω, y_Ω) has length $k-1$, has full support, is not trivially decomposable and $[z_\Omega, y_\Omega]$ contains a 2-crown. By the induction hypotesis, this implies $[(z_\Omega, y_\Omega)]_\sim \in ES_{k-1}(\mathcal{F})$. Also note that (x,y) is necessarily obtained from (z_Ω, y_Ω) by either an internal move or an enlarging move. Thus, by definition, $(x,y) \in Enl(z_\Omega, y_\Omega)$ and $[(x,y)]_\sim \in ES_k(\mathcal{F})$. ☐

We can now state and prove the main result of this work.

theorem 4.3. *The Kazhdan-Lusztig polynomials are combinatorial invariants for intervals up to length* 6 *in Coxeter groups of type* **B** *and* **D** *and for intervals up to length* 8 *in Coxeter groups of type* **A**.

Proof. The combinatorial invariance is known to hold for intervals up to length 4 in all Coxeter groups and in [**10**] it has been established for intervals of length 5 and 6 in the symmetric group. Moreover, by equation (1), if the combinatorial invariance is true for intervals up to a given odd length ℓ, then it is also true for intervals of length $\ell+1$. So we only need to prove the combinatorial invariance of the $\widetilde{R}$-polynomials for intervals of length 5 in the groups of signed permutations and for those of length 7 in the symmetric group. As already observed, we only need to consider intervals containing 2-crowns. By Proposition 3.3, we may only consider pairs which have full support. Pairs which are trivially decomposable can be managed by Proposition 3.5. Then, by Theorem 4.2, we only need to consider the pairs in the sets $ES_k(\mathcal{F})$.

For the remainder of the proof, we need the assistance of Maple computation. In fact, the essential sets have been generated, according to Definition 4.1, by a Maple program. For the symmetric group it has been done up to length 7 and for the groups of signed permutations up to length 5. The cardinalities of the essential sets are as follows:

k	$\lvert ES_k(\mathcal{F}_A)\rvert$	$\lvert ES_k(\mathcal{F}_B)\rvert$	$\lvert ES_k(\mathcal{F}_D)\rvert$
3	1	6	7
4	4	209	158
5	47	9543	3942
6	913		
7	22400		

For each pair (x, y) in the essential sets $ES_7(\mathcal{F}_A)$, $ES_5(\mathcal{F}_B)$ and $ES_5(\mathcal{F}_D)$, the poset structure of the interval $[x, y]$ has been determined, and the corresponding $\widetilde{R}$-polynomial has been computed, by our own Maple programs, based on algorithms that use the characterizations of the Bruhat order, equation (2) and the reflection orderings mentioned in the previous section. Then, the pairs have been grouped in isomorphism classes, with the help of Stembridge's Maple package for posets [**14**], which includes a fast algorithm for isomorphism testing. Finally, the combinatorial invariance of the $\widetilde{R}$-polynomials for these pairs has been checked. The results of the computation are summarized in Tables 1, 2 and 3, described later.

Note that it may happen that a pair (x, y), which is not trivially decomposable, has a corresponding interval which is reducible as a poset (that is, direct product of smaller posets), say $[x, y] \cong [x, z] \times [z, y]$. Then, consistently with Proposition 3.5, it has to be proved that, whenever this happens, the factorization $\widetilde{R}_{x,y}(q) = \widetilde{R}_{x,z}(q) \cdot \widetilde{R}_{z,y}$ holds. This has also been checked by Maple computation. $\square$

In Tables 1, 2 and 3 (the last one in a short version) all isomorphism types of intervals associated with pairs in the essential sets $ES_5(\mathcal{F}_D)$, $ES_5(\mathcal{F}_B)$ and $ES_7(\mathcal{F}_A)$, respectively, are listed. They are grouped by the value of the $\widetilde{R}$-polynomial and, within each group, they are listed for lexicographically nondecreasing f-vector. For each isomorphism type a representative pair (x, y) is indicated. Self-dual intervals and reducible intervals are marked, and, for each group, the expression of the R-polynomial is also indicated.

Note that some of the reducible intervals associated with trivially decomposable pairs might not have been considered. Nevertheless, this is not the case, since, by Maple computation, it has also been checked that all possible intervals containing 2-crowns that are direct product of smaller intervals belong to one of the isomorphism classes listed in the tables. Moreover, by an unpublished result of Dyer, we have that a Bruhat interval is a lattice if and only if it does not contain a 2-crown. We can conclude that Tables 1, 2 and 3 contain a complete classification, up to isomorphism, of Bruhat intervals which are not lattices, for the respective lengths and types.

The diagrams of the representative pairs are finally depicted in Figures 5, 6 and 7.

type	(x, y)	f-vector	s.d.	red.	$\widetilde{R}_{x,y}(q)$	$R_{x,y}(q)$
1.	$(1\,2\,3,\ 1\,\underline{3}\,\underline{2})$	$(3, 5, 6, 4)$			$q^5 + 2q^3 + q$	$(q-1)(q^2-q+1)^2$
2.	$(1\,2\,3,\ \underline{2}\,1\,\underline{3})$	$(3, 5, 5, 3)$	$\surd$		$q^5 + 2q^3$	$(q-1)^3(q^2+1)$
3.	$(1\,2\,3\,4,\ \underline{2}\,\underline{4}\,1\,3)$	$(4, 7, 7, 4)$	$\surd$	$\surd$		
4.	$(1\,2\,3\,4,\ \underline{1}\,2\,\underline{4}\,3)$	$(4, 8, 9, 5)$		$\surd$		
5.	$(1\,2\,3\,4,\ \underline{1}\,4\,\underline{3}\,2)$	$(4, 9, 10, 5)$				
6.	$(2\,1\,3\,4,\ 4\,\underline{3}\,\underline{2}\,1)$	$(4, 10, 12, 6)$				
7.	$(\underline{1}\,\underline{2}\,3\,4,\ 3\,\underline{4}\,\underline{2}\,1)$	$(5, 10, 10, 5)$			$q^5 + q^3$	$(q-1)^3(q^2-q+1)$
8.	$(\underline{1}\,\underline{2}\,3\,4,\ 1\,\underline{4}\,\underline{3}\,2)$	$(5, 10, 11, 6)$				
9.	$(\underline{1}\,2\,\underline{3}\,4,\ 1\,2\,\underline{4}\,\underline{3})$	$(5, 11, 14, 8)$				
10.	$(\underline{1}\,\underline{2}\,3\,4,\ 1\,\underline{4}\,3\,\underline{2})$	$(5, 12, 13, 6)$				
11.	$(\underline{1}\,\underline{3}\,2\,4,\ 1\,\underline{4}\,2\,\underline{3})$	$(5, 12, 14, 7)$				
12.	$(\underline{1}\,4\,\underline{2}\,3,\ 1\,2\,\underline{4}\,\underline{3})$	$(7, 15, 16, 8)$				

TABLE 1. Isomorphism types of pairs in $ES_5(\mathcal{F}_D)$.

type	(x, y)	f-vector	s.d.	red.	$\widetilde{R}_{x,y}(q)$	$R_{x,y}(q)$
1.	$(1 2 3, \underline{3} 2 \underline{1})$	$(3, 5, 6, 4)$				
2.	$(\underline{1} 2 3, \underline{1} 2 \underline{3})$	$(3, 6, 7, 4)$			$q^5 + 2q^3 + q$	$(q-1)(q^2-q+1)^2$
3.	$(\underline{2} 1 3, \underline{2} 1 \underline{3})$	$(4, 7, 7, 4)$	$\surd$			
4.	$(\underline{3} 1 2, \underline{1} \underline{3} \underline{2})$	$(3, 4, 4, 3)$	$\surd$	$\surd$		
5.	$(1 2 3, 2 \underline{3} 1)$	$(3, 5, 5, 3)$	$\surd$			
6.	$(1 2 3, \underline{3} \underline{2} 1)$	$(3, 5, 6, 4)$				
7.	$(1 3 2, \underline{1} 2 \underline{3})$	$(3, 5, 6, 4)$			$q^5 + 2q^3$	$(q-1)^3(q^2+1)$
8.	$(\underline{1} 2 3, \underline{3} \underline{2} \underline{1})$	$(3, 6, 6, 3)$	$\surd$			
9.	$(\underline{1} 2 3, 2 \underline{3} \underline{1})$	$(3, 6, 7, 4)$				
10.	$(\underline{1} 3 2, 1 \underline{3} \underline{2})$	$(4, 7, 7, 4)$	$\surd$			
11.	$(\underline{2} 1 3, 1 \underline{3} \underline{2})$	$(4, 7, 7, 4)$				
12.	$(\underline{2} 1 3, 2 \underline{1} \underline{3})$	$(4, 7, 7, 4)$	$\surd$	$\surd$		
13.	$(1 2 3 4, 4 \underline{2} 1 3)$	$(4, 8, 9, 5)$		$\surd$		
14.	$(\underline{1} 2 3 4, \underline{3} 4 \underline{1} 2)$	$(4, 9, 10, 5)$				
15.	$(\underline{1} 2 4 3, \underline{4} 2 3 \underline{1})$	$(4, 10, 12, 6)$				
16.	$(2 1 3 4, \underline{4} \underline{2} 1 3)$	$(5, 10, 10, 5)$	$\surd$	$\surd$		
17.	$(2 1 3 4, 4 \underline{2} 3 1)$	$(5, 10, 10, 5)$				
18.	$(1 \underline{2} 3 4, 1 \underline{4} \underline{2} 3)$	$(5, 10, 10, 5)$		$\surd$		
19.	$(2 1 3 4, 4 \underline{3} 1 2)$	$(5, 10, 11, 6)$				
20.	$(1 \underline{3} 2 4, 1 \underline{4} \underline{3} 2)$	$(5, 10, 11, 6)$				
21.	$(3 1 2 4, 4 \underline{3} \underline{1} 2)$	$(5, 11, 12, 6)$				
22.	$(3 1 2 4, \underline{4} \underline{3} 1 2)$	$(5, 11, 12, 6)$				
23.	$(\underline{4} 2 3 \underline{1}, \underline{3} 2 \underline{1} \underline{4})$	$(5, 11, 12, 6)$				
24.	$(\underline{2} 1 3 4, \underline{3} 4 \underline{2} 1)$	$(5, 11, 12, 6)$				
25.	$(1 \underline{3} \underline{2} 4, 1 \underline{4} \underline{3} \underline{2})$	$(5, 11, 12, 6)$				
26.	$(1 \underline{2} 3 4, 4 \underline{3} \underline{2} 1)$	$(5, 11, 13, 7)$				
27.	$(\underline{1} 2 \underline{3} 4, 1 2 \underline{4} \underline{3})$	$(5, 11, 14, 8)$				
28.	$(\underline{1} \underline{2} 3 4, 1 \underline{4} 3 \underline{2})$	$(5, 12, 13, 6)$				
29.	$(\underline{4} \underline{1} 3 2, \underline{3} \underline{1} 2 \underline{4})$	$(5, 12, 13, 6)$			$q^5 + q^3$	$(q-1)^3(q^2-q+1)$
30.	$(\underline{2} 4 \underline{1} 3, \underline{2} 3 \underline{4} \underline{1})$	$(5, 12, 13, 6)$				
31.	$(2 4 \underline{1} 3, 2 3 \underline{4} \underline{1})$	$(5, 12, 14, 7)$				
32.	$(\underline{1} \underline{3} 2 4, 1 \underline{4} 2 \underline{3})$	$(5, 12, 14, 7)$				
33.	$(3 \underline{1} 2 4, 4 \underline{3} 2 \underline{1})$	$(6, 12, 12, 6)$	$\surd$			
34.	$(\underline{1} 3 2 4, 4 \underline{3} \underline{1} 2)$	$(6, 12, 12, 6)$				
35.	$(\underline{2} 4 1 3, 3 \underline{4} \underline{2} 1)$	$(6, 13, 13, 6)$	$\surd$			
36.	$(1 \underline{2} 4 3, 3 \underline{4} \underline{2} 1)$	$(6, 13, 13, 6)$				
37.	$(\underline{4} 3 2 1, \underline{3} 2 \underline{1} \underline{4})$	$(6, 13, 13, 6)$				
38.	$(2 \underline{3} \underline{1} 4, 2 \underline{4} \underline{3} \underline{1})$	$(6, 13, 13, 6)$				
39.	$(\underline{2} 1 4 3, \underline{4} 1 3 \underline{2})$	$(6, 13, 13, 6)$				
40.	$(3 \underline{2} 1 4, 3 \underline{4} \underline{2} 1)$	$(6, 13, 13, 6)$				
41.	$(4 3 \underline{1} 2, 1 2 \underline{4} \underline{3})$	$(6, 13, 14, 7)$				
42.	$(1 4 \underline{2} 3, 1 3 \underline{4} \underline{2})$	$(6, 13, 14, 7)$				
43.	$(3 \underline{1} 2 4, 4 \underline{3} \underline{2} 1)$	$(6, 13, 14, 7)$				
44.	$(3 2 1 4, 4 \underline{3} \underline{2} 1)$	$(6, 14, 15, 7)$				
45.	$(\underline{1} 4 \underline{3} 2, \underline{1} 2 \underline{4} \underline{3})$	$(6, 14, 15, 7)$				
46.	$(\underline{1} 4 \underline{2} 3, 1 2 \underline{4} \underline{3})$	$(7, 15, 16, 8)$				

Table 2. Isomoprhism types of pairs in $ES_5(\mathcal{F}_B)$.

types	$\widetilde{R}_{x,y}(q)$	$R_{x,y}(q)$
1–2	$q^7 + 3q^5 + 3q^3 + q$	$(q-1)(q^2-q+1)^3$
3–5	$q^7 + 3q^5 + 2q^3$	$(q-1)^3(q^2+1)(q^2-q+1)$
6–11	$q^7 + 3q^5 + q^3$	$(q-1)^3(q^4-q^3+q^2-q+1)$
12–57	$q^7 + 2q^5 + q^3$	$(q-1)^3(q^2-q+1)^2$
58–89	$q^7 + 2q^5$	$(q-1)^5(q^2+1)$
90–217	$q^7 + q^5$	$(q-1)^5(q^2-q+1)$

TABLE 3. Isomorphism types of pairs in $ES_7(\mathcal{F}_A)$.

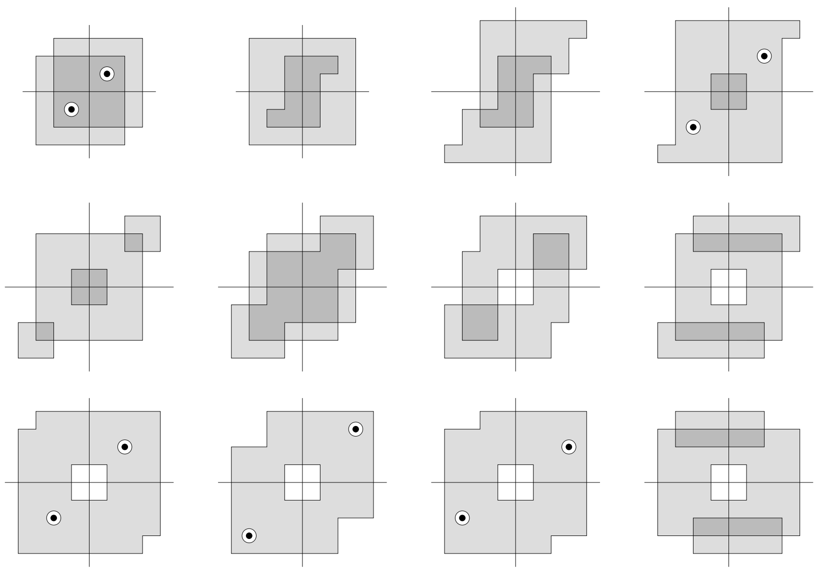


FIGURE 5. Representative diagrams of pairs in $ES_5(\mathcal{F}_D)$.

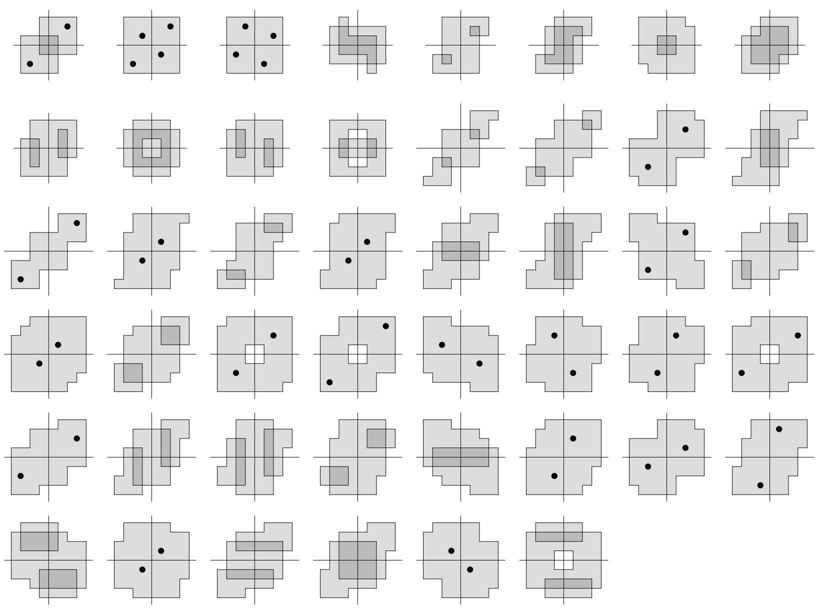


FIGURE 6. Representative diagrams of pairs in $ES_5(\mathcal{F}_B)$.

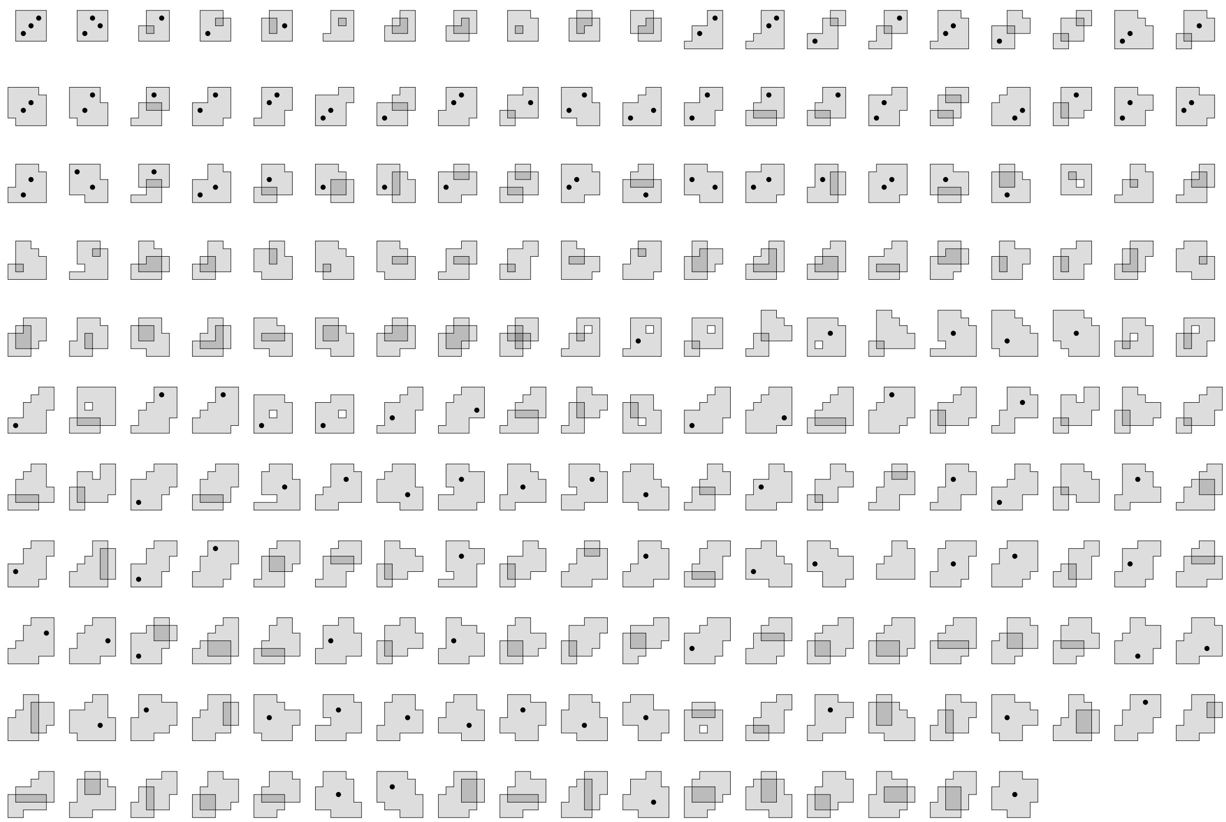


Figure 7. Representative diagrams of pairs in $ES_7(\mathcal{F}_A)$.

References

[1] A. Björner and F. Brenti, *Combinatorics of Coxeter Groups*, Graduate Texts in Mathematics, Vol. 231, Springer-Verlag, New York, 2005.

[2] A. Björner and M. Wachs, Bruhat order of Coxeter groups and shellabilty, *Adv. Math.* **43** (1982), 87–100.

[3] M. J. Dyer, On the "Bruhat graph" of a Coxeter system, *Compos. Math.* **78** (1991), 185–191.

[4] M. J. Dyer, Hecke algebras and shellings of Bruhat intervals, *Compos. Math.* **89** (1993), 91–115.

[5] M. J. Dyer, On minimal lengths of expressions of Coxeter group elements as products of reflections, *Proc. Amer. Math. Soc.* **129** (2001), no. 9, 2591–2595 (electronic).

[6] A. Hultman, Bruhat intervals of length 4 in Weyl groups, *J. Combin. Theory Ser. A* **102** (2003), 163–178.

[7] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Univ. Press, Cambridge, 1990.

[8] F. Incitti, Bruhat order on classical Weyl groups: minimal chains and covering relation, *European J. Combin.* **26** (2005), 729–753.

[9] F. Incitti, Permutation diagrams, fixed points and Kazhdan-Lusztig R-polynomials, to appear in *Ann. Combin.*

[10] F. Incitti, On the combinatorial invariance of Kazhdan-Lusztig polynomials, to appear in *J. Combin. Theory Ser. A.*

[11] C. Kassel, A. Lascoux and C. Reutenauer, The singular locus of a Schubert variety, *J. Algebra* **269** (2003), 74–108.

[12] D. Kazhdan and G. Lusztig, Representations of Coexeter groups and Hecke algebras, *Invent. Math.* **53** (1979), 165–184.

[13] R. P. Stanley, *Enumerative Combinatorics*, Vol. 1, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1986.

[14] J. R. Stembridge, Posets ver. 2.2, Maple package, 2002, see `http://www.math.lsa.umich.edu/~jrs/maple.html` .

Department of Mathematics, Royal Institute of Technology, 10044 Stockholm, Sweden

E-mail address: `incitti@kth.se`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The Partition Function of Andrews and Stanley and Al-Salam-Chihara Polynomials

Masao Ishikawa and Jiang Zeng

Abstract. For any partition λ let $\omega(\lambda)$ denote the four parameter weight

$$\omega(\lambda)=a^{\sum_{i\ge1}\lceil\lambda_{2i-1}/2\rceil}b^{\sum_{i\ge1}\lfloor\lambda_{2i-1}/2\rfloor}c^{\sum_{i\ge1}\lceil\lambda_{2i}/2\rceil}d^{\sum_{i\ge1}\lfloor\lambda_{2i}/2\rfloor},$$

and let $\ell(\lambda)$ be the length of λ. We show that the generating function $\sum\omega(\lambda)z^{\ell(\lambda)}$, where the sum runs over all ordinary (resp. strict) partitions with parts each $\le N$, can be expressed by the Al-Salam-Chihara polynomials. As a corollary we prove G.E. Andrews' result by specializing some parameters and C. Boulet's results when $N\to+\infty$. In the last section we study the weighted sum $\sum\omega(\lambda)z^{\ell(\lambda)}P_\lambda(x)$ where $P_\lambda(x)$ is Schur's P-function and the sum runs over all strict partitions.

Résumé.

Pour toute partition λ on définit $\omega(\lambda)$ comme la fonction poids de quatre paramètres

$$\omega(\lambda)=a^{\sum_{i\ge1}\lceil\lambda_{2i-1}/2\rceil}b^{\sum_{i\ge1}\lfloor\lambda_{2i-1}/2\rfloor}c^{\sum_{i\ge1}\lceil\lambda_{2i}/2\rceil}d^{\sum_{i\ge1}\lfloor\lambda_{2i}/2\rfloor},$$

et désigne $\ell(\lambda)$ la longueur de λ. On démontre que la fonction génératrice $\sum\omega(\lambda)z^{\ell(\lambda)}$, où la somme porte sur toutes les partitions ordinaires (resp. strictes) avec chaque part $\le N$, peut s'exprimer par les polynômes d'Al-Salam-Chihara. Comme corollaire on en déduit un résultat de G.E. Andrews en spécialisant certain parametres et ceux de C. Boulet quand $N\to+\infty$. Dans la dernière section on étudie la somme pondérée $\sum\omega(\lambda)z^{\ell(\lambda)}P_\lambda(x)$ où $P_\lambda(x)$ est la P-fonction de Schur et la somme porte sur toutes les partitions strictes.

1. Introduction

Let λ be an integer partition and λ' its conjugate. Let $\mathcal{O}(\lambda)$ denote the number of odd parts of λ and $|\lambda|$ the sum of its parts. R. Stanley ([**13**]) has shown that if $t(n)$ denotes the number of partitions λ of n for which $\mathcal{O}(\lambda)\equiv\mathcal{O}(\lambda')\pmod 4$, then

$$t(n)=\frac{1}{2}\left(p(n)+f(n)\right),$$

where $p(n)$ is the total number of partitions of n, and

$$\sum_{n=0}^{\infty}f(n)q^n=\prod_{i\ge1}\frac{(1+q^{2i-1})}{(1-q^{4i})(1+q^{4i-2})}.$$

In [**1**] G.E. Andrews has computed the generating function of ordinary partitions λ with parts each less than or equal to N, with respect to the weight $z^{\mathcal{O}(\lambda)}y^{\mathcal{O}(\lambda')}q^{|\lambda|}$. We should note that in [**12**] A. Sills has given a combinatorial proof of this result, and in [**14**] A. Yee has generalized this result to the generating function of ordinary partitions of parts$\le N$ and length $\le M$.

As a generalization of this weight, we consider the following four parameter weight. Let a, b, c and d be commuting indeterminates. Define the following weight functions $\omega(\lambda)$ on the set of all partitions,

$$\omega(\lambda)=a^{\sum_{i\ge1}\lceil\lambda_{2i-1}/2\rceil}b^{\sum_{i\ge1}\lfloor\lambda_{2i-1}/2\rfloor}c^{\sum_{i\ge1}\lceil\lambda_{2i}/2\rceil}d^{\sum_{i\ge1}\lfloor\lambda_{2i}/2\rfloor},\tag{1.1}$$

Key words and phrases. Partitions, symmetric functions, Al-Salam-Chihara polynomials, basic hypergeometric functions, Schur's Q-functions, Pfaffians, minor summation formula of Pfaffians.

where $\lceil x \rceil$ (resp. $\lfloor x \rfloor$) stands for the smallest (resp. largest) integer greater (resp. less) than or equal to x for a given real number x. For example, if $\lambda = (5, 4, 4, 1)$ then $\omega(\lambda)$ is the product of the entries in the following diagram for λ.

a	b	a	b	a
c	d	c	d	
a	b	a	b	
c				

In [**3**] C. Boulet has obtained results on the generating functions for the weights $\omega(\lambda)$ when λ runs over all ordinary partitions and when λ runs over all strict partitions

In this paper we consider a refinement of these results, i.e., the generating functions for the weights $\omega(\lambda)z^{\ell(\lambda)}$, where $\ell(\lambda)$ is the length of λ, when λ runs over all ordinary partitions with parts each $\leq N$ and when λ runs over all strict partitions with parts each $\leq N$, and show that they are related to the basic hypergeometric series, i.e. the Al-Salam-Chihara polynomials (see Theorem 3.4 and Theorem 4.3).

In the last section we show the weighted sum $\sum \omega(\mu)z^{\ell(\mu)}P_\mu(x)$ of Schur's P-functions $P_\mu(x)$ (when $z = 2$, this equals the weighted sum $\sum \omega(\mu)Q_\mu(x)$ of Schur's Q-functions $Q_\mu(x)$) can be expressed by a Pfaffian where μ runs over all strict partitions (with parts each $\leq N$).

2. Preliminaries

A q-shifted factorial is defined by

$$(a;q)_0 = 1, \qquad (a;q)_n = (1-a)(1-aq)\cdots(1-aq^{n-1}), \qquad n = 1, 2, \dots.$$

We also define $(a;q)_\infty = \prod_{k=0}^{\infty}(1-aq^k)$. Since products of q-shifted factorials occur very often, to simplify them we shall use the compact notations

$$\begin{aligned}(a_1, \dots, a_m; q)_n &= (a_1;q)_n \cdots (a_m;q)_n,\\ (a_1, \dots, a_m; q)_\infty &= (a_1;q)_\infty \cdots (a_m;q)_\infty.\end{aligned}$$

We define an ${}_{r+1}\phi_r$ *basic hypergeometric series* by

$${}_{r+1}\phi_r\left(\begin{matrix} a_1, a_2, \dots, a_{r+1} \\ b_1, \dots, b_r \end{matrix}; q, z\right) = \sum_{n=0}^{\infty} \frac{(a_1, a_2, \dots, a_{r+1}; q)_n}{(q, b_1, \dots, b_r; q)_n} z^n.$$

The *Al-Salam-Chihara polynomial* $Q_n(x) = Q_n(x;\alpha,\beta|q)$ is, by definition,

$$\begin{aligned} Q_n(x;\alpha,\beta|q) &= \frac{(\alpha\beta;q)_n}{\alpha^n}\, {}_3\phi_2\left(\begin{matrix} q^{-n}, \alpha u, \alpha u^{-1} \\ \alpha\beta, 0 \end{matrix}; q, q\right),\\ &= (\alpha u;q)_n u^{-n}\, {}_2\phi_1\left(\begin{matrix} q^{-n}, \beta u^{-1} \\ \alpha^{-1}q^{-n+1}u^{-1} \end{matrix}; q, \alpha^{-1}qu\right),\\ &= (\beta u^{-1};q)_n u^{n}\, {}_2\phi_1\left(\begin{matrix} q^{-n}, \alpha u \\ \beta^{-1}q^{-n+1}u \end{matrix}; q, \beta^{-1}qu^{-1}\right), \end{aligned}$$

where $x = \frac{u+u^{-1}}{2}$ (see [**6**] p.80). This is a specialization of the Askey-Wilson polynomials (see [**2**]), and satisfies the three-term recurrence relation

$$2xQ_n(x) = Q_{n+1}(x) + (\alpha+\beta)q^nQ_n(x) + (1-q^n)(1-\alpha\beta q^{n-1})Q_{n-1}(x), \tag{2.1}$$

with $Q_{-1}(x) = 0$, $Q_0(x) = 1$.

We also consider a more general recurrence relation:

$$2x\widetilde{Q}_n(x) = \widetilde{Q}_{n+1}(x) + (\alpha+\beta)tq^n\widetilde{Q}_n(x) + (1-tq^n)(1-t\alpha\beta q^{n-1})\widetilde{Q}_{n-1}(x), \tag{2.2}$$

which we call the *associated Al-Salam-Chihara recurrence relation.* Put

$$\widetilde{Q}_n^{(1)}(x) = u^{-n}\,(t\alpha u;q)_n\,{}_2\phi_1\left(\begin{matrix} t^{-1}q^{-n},\beta u^{-1} \\ t^{-1}\alpha^{-1}q^{-n+1}u^{-1}\end{matrix};\, q, \alpha^{-1}qu\right), \tag{2.3}$$

$$\widetilde{Q}_n^{(2)}(x) = u^{n}\,\frac{(tq;q)_n(t\alpha\beta;q)_n}{(t\beta uq;q)_n}\,{}_2\phi_1\left(\begin{matrix} tq^{n+1},\alpha^{-1}qu \\ t\beta q^{n+1}u\end{matrix};\, q, \alpha u\right), \tag{2.4}$$

where $x = \frac{u+u^{-1}}{2}$. In [**5**], Ismail and Rahman have presented two linearly independent solutions of the associated Askey-Wilson recurrence equation (see also [**4**]). By specializing the parameters, we conclude that $\widetilde{Q}_n^{(1)}(x)$ and $\widetilde{Q}_n^{(2)}(x)$ are two linearly independent solutions of the associated Al-Salam-Chihara equation (2.2) (see [**4**, p.203]). Here, we use this fact and omit the proof. The series (2.3) and (2.4) are convergent if we assume $|u| < 1$ and $|q| < |\alpha| < 1$ (see [**4**, p.204]).

Let

$$W_n = \widetilde{Q}_n^{(1)}(x)\widetilde{Q}_{n-1}^{(2)}(x) - \widetilde{Q}_{n-1}^{(1)}(x)\widetilde{Q}_n^{(2)}(x) \tag{2.5}$$

denote the Casorati determinant of the equation (2.2). Then we obtain

$$W_1 = \frac{u^{-1}(t\alpha u,\beta u;q)_\infty}{(\alpha u, t\beta uq;q)_\infty}. \tag{2.6}$$

In the following sections we need to find a polynomial solution of the recurrence equation (2.2) which satisfies a given initial condition, say $\widetilde{Q}_0(x) = \widetilde{Q}_0$ and $\widetilde{Q}_1(x) = \widetilde{Q}_1$. Since $\widetilde{Q}_n^{(1)}(x)$ and $\widetilde{Q}_n^{(2)}(x)$ are linearly independent solutions of (2.2), this $\widetilde{Q}_n(x)$ can be written as a linear combination of these functions, say

$$\widetilde{Q}_n(x) = C_1\,\widetilde{Q}_n^{(1)}(x) + C_2\,\widetilde{Q}_n^{(2)}(x).$$

If we substitute the initial condition $\widetilde{Q}_0(x) = \widetilde{Q}_0$ and $\widetilde{Q}_1(x) = \widetilde{Q}_1$ into this equation and solve the linear equation, then we conclude that

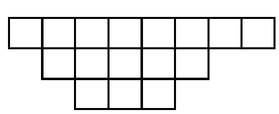
$$\begin{aligned}\widetilde{Q}_n(x) = \frac{u(\alpha u, t\beta uq;q)_\infty}{(t\alpha u,\beta u;q)_\infty}\Big[&\Big\{\widetilde{Q}_1\widetilde{Q}_0^{(2)}(x) - \widetilde{Q}_0\widetilde{Q}_1^{(2)}(x)\Big\}\widetilde{Q}_n^{(1)}(x) \\ &+\Big\{\widetilde{Q}_0\widetilde{Q}_1^{(1)}(x) - \widetilde{Q}_1\widetilde{Q}_0^{(1)}(x)\Big\}\widetilde{Q}_n^{(2)}(x)\Big]\end{aligned} \tag{2.7}$$

and

$$\lim_{n\to\infty} u^n\widetilde{Q}_n(x) = \frac{u(t\beta uq,\alpha u;q)_\infty}{(u^2;q)_\infty}\Big\{\widetilde{Q}_1\widetilde{Q}_0^{(2)}(x) - \widetilde{Q}_0\widetilde{Q}_1^{(2)}(x)\Big\}. \tag{2.8}$$

3. Strict Partitions

A partition μ is *strict* if all its parts are distinct. One represents the associated shifted diagram of μ as a diagram in which the ith row from the top has been shifted to the right by i places so that the first column becomes a diagonal. A strict partition can be written uniquely in the form $\mu = (\mu_1,\ldots,\mu_{2n})$ where n is an non-negative integer and $\mu_1 > \mu_2 > \cdots > \mu_{2n} \geq 0$. The *length* $\ell(\mu)$ is, by definition, the number of nonzero parts of μ. We define the weight function $\omega(\mu)$ exactly the same as in (1.1). For example, if $\mu = (8,5,3)$, then $\ell(\mu) = 3$, $\omega(\mu) = a^6b^5c^3d^2$ and its shifted diagram is as follows.



Let

$$\Psi_N = \Psi_N(a,b,c,d;z) = \sum \omega(\mu)z^{\ell(\mu)}, \tag{3.1}$$

where the sum is over all strict partitions μ such that each part of μ is less than or equal to N. For example, we have

$$\begin{aligned}
&\Psi_0 = 1,\\
&\Psi_1 = 1 + az,\\
&\Psi_2 = 1 + a(1+b)z + abcz^2,\\
&\Psi_3 = 1 + a(1+b+ab)z + abc(1+a+ad)z^2 + a^3bcdz^3.
\end{aligned}$$

In fact, the only strict partition such that $\ell(\mu) = 0$ is $\emptyset$, the strict partitions μ such that $\ell(\mu) = 1$ and $\mu_1 \le 3$ are the following three:

$$\begin{array}{|c|}\hline a \\ \hline\end{array} \qquad \begin{array}{|c|c|}\hline a & b \\ \hline\end{array} \qquad \begin{array}{|c|c|c|}\hline a & b & a \\ \hline\end{array},$$

the strict partitions μ such that $\ell(\mu) = 2$ and $\mu_1 \le 3$ are the following three:

$$\begin{array}{|c|c|}\hline a & b \\ \hline \multicolumn{1}{c|}{} & c \\ \cline{2-2}\end{array} \qquad \begin{array}{|c|c|c|}\hline a & b & a \\ \hline \multicolumn{1}{c|}{} & c & \multicolumn{1}{|c}{} \\ \cline{2-2}\end{array} \qquad \begin{array}{|c|c|c|}\hline a & b & a \\ \hline \multicolumn{1}{c|}{} & c & d \\ \cline{2-3}\end{array},$$

and the strict partition μ such that $\ell(\mu) = 3$ and $\mu_1 \le 3$ is the following one:

$$\begin{array}{|c|c|c|}\hline a & b & a \\ \hline \multicolumn{1}{c|}{} & c & d \\ \cline{2-3} \multicolumn{1}{c}{} & \multicolumn{1}{c|}{} & a \\ \cline{3-3}\end{array}.$$

The sum of the weights of these strict partitions is equal to Ψ_3. In this section we always assume $|a|, |b|, |c|, |d| < 1$. One of the main results of this section is that the *even index* terms and the *odd index* terms of Ψ_N respectively satisfy the associated Al-Salam-Chihara recurrence relation:

Theorem 3.1. Set $q = abcd$. Let $\Psi_N = \Psi_N(a, b, c, d; z)$ be as in (3.1) and put $X_N = \Psi_{2N}$ and $Y_N = \Psi_{2N+1}$. Then X_N and Y_N satisfy

$$\begin{aligned}
X_{N+1} = \left\{1 + ab + a(1+bc)z^2q^N\right\} X_N&\\
- ab(1 - z^2q^N)(1 - acz^2q^{N-1})X_{N-1}&, \qquad (3.2)\\
Y_{N+1} = \left\{1 + ab + abc(1+ad)z^2q^N\right\} Y_N&\\
- ab(1 - z^2q^N)(1 - acz^2q^N)Y_{N-1}&, \qquad (3.3)
\end{aligned}$$

where $X_0 = 1$, $Y_0 = 1 + az$, $X_1 = 1 + a(1+b)z + abcz^2$ and

$$Y_1 = 1 + a(1+b+ab)z + abc(1+a+ad)z^2 + a^3bcdz^3.$$

Especially, if we put $X'_N = (ab)^{-\frac{N}{2}}X_N$ and $Y'_N = (ab)^{-\frac{N}{2}}Y_N$, then X'_N and Y'_N satisfy

$$\begin{aligned}
\left\{(ab)^{\frac12} + (ab)^{-\frac12}\right\} X'_N = X'_{N+1} - a^{\frac12}b^{-\frac12}(1+bc)z^2q^N X'_N&\\
+ (1 - z^2q^N)(1 - acz^2q^{N-1})X'_{N-1}&, \qquad (3.4)\\
\left\{(ab)^{\frac12} + (ab)^{-\frac12}\right\} Y'_N = Y'_{N+1} - a^{\frac12}b^{\frac12}c(1+ad)z^2q^N Y'_N&\\
+ (1 - z^2q^N)(1 - a^2bc^2dz^2q^{N-1})Y'_{N-1}&, \qquad (3.5)
\end{aligned}$$

where $X'_0 = 1$, $Y'_0 = 1 + az$, $X'_1 = (ab)^{-\frac12} + a^{\frac12}b^{-\frac12}(1+b)z + (ab)^{\frac12}cz^2$ and

$$Y'_1 = (ab)^{-\frac12} + a^{\frac12}b^{-\frac12}(1+b+ab)z + a^{\frac12}b^{\frac12}c(1+a+ad)z^2 + a^{\frac52}b^{\frac12}cdz^3.$$

One concludes that, when $|a|, |b|, |c|, |d| < 1$, the solutions of (3.2) and (3.3) are expressed by the linear combinations of (2.3) and (2.4) as follows.

Theorem 3.2. Assume $|a|, |b|, |c|, |d| < 1$ and set $q = abcd$. Let $\Psi_N = \Psi_N(a, b, c, d; z)$ be as in (3.1).

(i) Put $X_N = \Psi_{2N}$. Then we have

$$X_N = \frac{(-az^2q, -abc; q)_\infty}{(-a, -abcz^2; q)_\infty}$$
$$\times \left\{ (s_0^X X_1 - s_1^X X_0)(-abcz^2; q)_N \, {}_2\phi_1 \left(\begin{matrix} q^{-N}z^{-2}, -b^{-1} \\ -(abc)^{-1}q^{-N+1}z^{-2} \end{matrix}; q, -c^{-1}q \right) \right.$$
$$\left. +(r_1^X X_0 - r_0^X X_1)(ab)^N \frac{(qz^2, acz^2; q)_N}{(-aqz^2; q)_N} \, {}_2\phi_1 \left(\begin{matrix} q^{N+1}z^2, -c^{-1}q \\ -aq^{N+1}z^2 \end{matrix}; q, -abc \right) \right\}, \tag{3.6}$$

where

$$r_0^X = {}_2\phi_1 \left(\begin{matrix} z^{-2}, -b^{-1} \\ -(abc)^{-1}z^{-2}q \end{matrix}; q, -c^{-1}q \right),$$
$$s_0^X = {}_2\phi_1 \left(\begin{matrix} z^2q, -c^{-1}q \\ -az^2q \end{matrix}; q, -abc \right),$$
$$r_1^X = (1 + abcz^2) \, {}_2\phi_1 \left(\begin{matrix} z^{-2}q^{-1}, -b^{-1} \\ -(abc)^{-1}z^{-2} \end{matrix}; q, -c^{-1}q \right),$$
$$s_1^X = \frac{ab(1 - z^2q)(1 - acz^2)}{1 + az^2q} \, {}_2\phi_1 \left(\begin{matrix} z^2q^2, -c^{-1}q \\ -az^2q^2 \end{matrix}; q, -abc \right).$$

(ii) Put $Y_N = \Psi_{2N+1}$. Then we have

$$Y_N = \frac{(-a^2bcdz^2q, -abc; q)_\infty}{(-a^2bcd, -abcz^2; q)_\infty}$$
$$\times \left\{ (s_0^Y Y_1 - s_1^Y Y_0)(-abcz^2; q)_N \, {}_2\phi_1 \left(\begin{matrix} q^{-N}z^{-2}, -acd \\ -(abc)^{-1}q^{-N+1}z^{-2} \end{matrix}; q, -c^{-1}q \right) \right.$$
$$\left. + (r_1^Y Y_0 - r_0^Y Y_1)(ab)^N \frac{(qz^2, a^2bc^2dz^2; q)_N}{(-a^2bcdqz^2; q)_N} \, {}_2\phi_1 \left(\begin{matrix} q^{N+1}z^2, -c^{-1}q \\ -a^2bcdq^{N+1}z^2 \end{matrix}; q, -abc \right) \right\}, \tag{3.7}$$

where

$$r_0^Y = {}_2\phi_1 \left(\begin{matrix} z^{-2}, -acd \\ (-abc)^{-1}qz^{-2} \end{matrix}; q, -c^{-1}q \right),$$
$$r_1^Y = (1 + abcz^2) \, {}_2\phi_1 \left(\begin{matrix} q^{-1}z^{-2}, -ac \\ -(abc)^{-1}z^{-2} \end{matrix}; q, -c^{-1}q \right),$$
$$s_0^Y = {}_2\phi_1 \left(\begin{matrix} z^2q, -c^{-1}q \\ -a^2bcdz^2q \end{matrix}; q, -abc \right),$$
$$s_1^Y = \frac{ab(1 - z^2q)(1 - a^2bc^2dz^2)}{1 + a^2bcdz^2q} \, {}_2\phi_1 \left(\begin{matrix} z^2q^2, -c^{-1}q \\ -a^2bcdz^2q^2 \end{matrix}; q, -abc \right).$$

If we take the limit $N \to \infty$ in (3.6) and (3.7), then by using (2.8), we obtain the following generalization of Boulet's result (see Corollary 3.6).

Corollary 3.3. Assume $|a|, |b|, |c|, |d| < 1$ and set $q = abcd$. Let s_i^X, s_i^Y, X_i, Y_i $(i = 0, 1)$ be as in the above theorem. Then we have

$$\sum_\mu \omega(\mu) z^{\ell(\mu)} = \frac{(-abc, -az^2q; q)_\infty}{(ab; q)_\infty} (s_0^X X_1 - s_1^X X_0)$$
$$= \frac{(-abc, -a^2bcdz^2q; q)_\infty}{(ab; q)_\infty} (s_0^Y Y_1 - s_1^Y Y_0), \tag{3.8}$$

where the sum runs over all strict partitions.

Especially, by substituting $z = 1$ into (3.6) and (3.7), we conclude that the solutions of the recurrence relations (3.4) and (3.5) with the above initial condition are exactly the Al-Salam-Chihara polynomials:

Theorem 3.4. Put $u = \sqrt{ab}$, $x = \frac{u+u^{-1}}{2}$ and $q = abcd$. Let $\Psi_N(a, b, c, d; z)$ be as in (3.1).

(i) The polynomial $\Psi_{2N}(a,b,c,d;1)$ is given by

$$\Psi_{2N}(a,b,c,d;1) = (ab)^{\frac{N}{2}} Q_N(x; -a^{\frac12}b^{\frac12}c, -a^{\frac12}b^{-\frac12}|q),$$

$$= (-a;q)_N \, {}_2\phi_1\left(\begin{matrix} q^{-N}, -c \\ -a^{-1}q^{-N+1} \end{matrix}; q, -bq\right). \tag{3.9}$$

(ii) The polynomial $\Psi_{2N+1}(a,b,c,d;1)$ is given by

$$\Psi_{2N+1}(a,b,c,d;1) = (1+a)(ab)^{\frac{N}{2}} Q_N(x; -a^{\frac12}b^{\frac12}c, -a^{\frac32}b^{\frac12}cd|q)$$

$$= (-a;q)_{N+1} \, {}_2\phi_1\left(\begin{matrix} q^{-N}, -c \\ -a^{-1}q^{-N} \end{matrix}; q, -b\right). \tag{3.10}$$

If we substitute $a = zyq$, $b = z^{-1}yq$, $c = zy^{-1}q$ and $d = z^{-1}y^{-1}q$ into Theorem 3.4, then we immediately obtain the following corollary, which is a strict version of Andrews' result.

Corollary 3.5.

$$\sum_{\substack{\mu \text{ strict partitions} \\ \mu_1 \le 2N}} z^{\mathcal{O}(\mu)} y^{\mathcal{O}(\mu')} q^{|\mu|} = \sum_{j=0}^{N} \begin{bmatrix} N \\ j \end{bmatrix}_{q^4} (-zyq;q^4)_j (-zy^{-1}q;q^4)_{N-j} (yq)^{2N-2j}, \tag{3.11}$$

and

$$\sum_{\substack{\mu \text{ strict partitions} \\ \mu_1 \le 2N+1}} z^{\mathcal{O}(\mu)} y^{\mathcal{O}(\mu')} q^{|\mu|} = \sum_{j=0}^{N} \begin{bmatrix} N \\ j \end{bmatrix}_{q^4} (-zyq;q^4)_{j+1} (-zy^{-1}q;q^4)_{N-j} (yq)^{2N-2j}, \tag{3.12}$$

where

$$\begin{bmatrix} N \\ j \end{bmatrix}_q = \begin{cases} \frac{(1-q^N)(1-q^{N-1})\cdots(1-q^{N-j+1})}{(1-q^j)(1-q^{j-1})\cdots(1-q)}, & \text{for } 0 \le j \le N, \\ 0, & \text{if } j < 0 \text{ and } j > N. \end{cases}$$

If we put $N \to \infty$ in Corollary 3.4, then we immediately obtain the following corollary (cf. Corollary 2 of [**3**]). We can also prove this corollary by setting $z \to 1$ in (3.8).

Corollary 3.6. (Boulet) Let $q = abcd$, then

$$\sum_{\mu} \omega(\mu) = \frac{(-a;q)_\infty (-abc;q)_\infty}{(ab;q)_\infty}, \tag{3.13}$$

where the sum runs over all strict partitions μ.

4. Ordinary Partitions

First we present a generalization of Andrews' result in [**1**]. Let us consider

$$\Phi_N = \Phi_N(a,b,c,d;z) = \sum_{\substack{\lambda \\ \lambda_1 \le N}} \omega(\lambda) z^{\ell(\lambda)}, \tag{4.1}$$

where the sum runs over all partitions λ such that each part of λ is less than or equal to N. For example, the first few terms can be computed directly as follows:

$$\Phi_0 = 1,$$

$$\Phi_1 = \frac{1+az}{1-acz^2},$$

$$\Phi_2 = \frac{1+a(1+b)z+abcz^2}{(1-acz^2)(1-qz^2)},$$

$$\Phi_3 = \frac{1+a(1+b+ab)z+abc(1+a+ad)z^2+a^3bcdz^3}{(1-z^2ac)(1-z^2q)(1-z^2acq)},$$

where $q = abcd$ as before. If one compares these with the first few terms of Ψ_n, one can easily guess the following theorem holds:

THEOREM 4.1. Let N be a non-negative integer, and let $\Phi_N = \Phi_N(a,b,c,d;z)$ be as in (4.1). Then we have

$$\Phi_N(a,b,c,d;z) = \frac{\Psi_N(a,b,c,d;z)}{(z^2q;q)_{\lfloor N/2\rfloor}(z^2ac;q)_{\lceil N/2\rceil}}, \tag{4.2}$$

where $\Psi_N = \Psi_N(a,b,c,d;z)$ is the generating function defined in (3.1). Note that Ψ_N is explicitly given in terms of basic hypergeometric functions in Theorem 3.2.

First of all, as an immediate corollary of Theorem 4.1 and Corollary 3.3, we obtain the following generalization of Boulet's result.

COROLLARY 4.2. Assume $|a|,|b|,|c|,|d| < 1$ and set $q = abcd$. Let s_i^X, s_i^Y, X_i, Y_i $(i = 0,1)$ be as in Theorem 3.2. Then we have

$$\sum_{\lambda} \omega(\lambda) z^{|\mu|} = \frac{(-abc, -az^2q;q)_\infty}{(ab, acz^2, z^2q;q)_\infty}(s_0^X X_1 - s_1^X X_0) \tag{4.3}$$

where the sum runs over all partitions λ.

Theorem 4.1 and Theorem 3.4 also give the following corollary:

COROLLARY 4.3. Put $x = \frac{(ab)^{\frac12}+(ab)^{-\frac12}}{2}$ and $q = abcd$. Let $\Phi_N = \Phi_N(a,b,c,d;z)$ be as in (4.1).

(i) The generating function $\Phi_{2N}(a,b,c,d;1)$ is given by

$$\begin{aligned}\Phi_{2N}(a,b,c,d;1) &= \frac{(ab)^{\frac{N}{2}} Q_N(x; -a^{\frac12}b^{\frac12}c, -a^{\frac12}b^{-\frac12}|q)}{(q;q)_N(ac;q)_N}\\ &= \frac{(-a;q)_N}{(q;q)_N(ac;q)_N}\, {}_2\phi_1\left(\begin{matrix} q^{-N}, -c \\ -a^{-1}q^{-N+1}\end{matrix}; q, -bq\right).\end{aligned} \tag{4.4}$$

(ii) The generating function $\Phi_{2N}(a,b,c,d;1)$ is given by

$$\begin{aligned}\Phi_{2N+1}(a,b,c,d;1) &= \frac{(1+a)(ab)^{\frac{N}{2}} Q_N(x; -a^{\frac12}b^{\frac12}c, -a^{\frac32}b^{\frac12}cd|q)}{(q;q)_N(ac;q)_{N+1}}\\ &= \frac{(-a;q)_{N+1}}{(q;q)_N(ac;q)_{N+1}}\, {}_2\phi_1\left(\begin{matrix} q^{-N}, -c \\ -a^{-1}q^{-N}\end{matrix}; q, -b\right).\end{aligned} \tag{4.5}$$

As before we immediately deduce the following corollary from Corollary 4.3. Let $S_N(n,r,s)$ denote the number of partitions π of n where each part of π is $\le N$, $\mathcal{O}(\pi) = r$, $\mathcal{O}(\pi') = s$. Then we have the result of Andrews [**1**, Theorem 1].

COROLLARY 4.4. (Andrews)

$$\sum_{n,r,s\ge0} S_{2N}(n,r,s)q^n z^r y^s = \frac{\sum_{j=0}^N \begin{bmatrix} N \\ j\end{bmatrix}_{q^4} (-zyq;q^4)_j(-zy^{-1}q;q^4)_{N-j}(yq)^{2N-2j}}{(q^4;q^4)_N(z^2q^4;q^4)_N}, \tag{4.6}$$

and

$$\sum_{n,r,s\ge0} S_{2N+1}(n,r,s)q^n z^r y^s = \frac{\sum_{j=0}^N \begin{bmatrix} N \\ j\end{bmatrix}_{q^4} (-zyq;q^4)_{j+1}(-zy^{-1}q;q^4)_{N-j}(yq)^{2N-2j}}{(q^4;q^4)_N(z^2q^4;q^4)_{N+1}}. \tag{4.7}$$

COROLLARY 4.5. (Boulet) Let $q = abcd$, then

$$\sum_{\lambda \text{ partitions}} \omega(\lambda) = \frac{(-a;q)_\infty(-abc;q)_\infty}{(q;q)_\infty(ab;q)_\infty(ac;q)_\infty}. \tag{4.8}$$

Here the sum runs over all partitions λ (cf. [**3**, Theorem 1]).

First we show the following recurrence equations hold.

Proposition 4.6. Let $\Phi_N = \Phi_N(a,b,c,d;z)$ be as before and $q = abcd$. Then the following recurrences hold for any positive integer N.

$$(1-z^2q^N)\Phi_{2N} = (1+b)\Phi_{2N-1} - b\,\Phi_{2N-2}, \tag{4.9}$$

$$(1-z^2acq^N)\Phi_{2N+1} = (1+a)\Phi_{2N} - a\Phi_{2N-1}. \tag{4.10}$$

5. A weighted sum of Schur's P-functions

We use the notation $X = X_n = (x_1,\dots,x_n)$ for the finite set of variables $x_1,\dots,x_n$. In [**8**], one of the authors used a Paffian expression of $\sum_\lambda \omega(\lambda)s_\lambda(X)$ to prove Stanley's open problem, where the sum runs over all partitions λ and $s_\lambda(X)$ stands for the Schur function with respect to a partition λ. The aim of this section is to give some determinantial formulas for the weighted sum $\sum \omega(\mu)z^{\ell(\mu)}P_\mu(x)$ where $P_\mu(x)$ is Schur's P-function.

Let A_n denote the skew-symmetric matrix

$$\left(\frac{x_i-x_j}{x_i+x_j}\right)_{1\le i,j\le n}$$

and for each strict partition $\mu = (\mu_1,\dots,\mu_l)$ of length $l \le n$, let Γ_μ denote the $n\times l$ matrix $\left(x_j^{\mu_i}\right)$. Let

$$A_\mu(x_1,\dots,x_n) = \begin{pmatrix} A_n & \Gamma_\mu J_l \\ -J_l {}^t\Gamma_\mu & O_l \end{pmatrix}$$

which is a skew-symmetric matrix of $(n+l)$ rows and columns. Define $\mathrm{Pf}_\mu(x_1,\dots,x_n)$ to be $\mathrm{Pf}\,A_\mu(x_1,\dots,x_n)$ if $n+l$ is even, and to be $\mathrm{Pf}\,A_\mu(x_1,\dots,x_n,0)$ if $n+l$ is odd. By Ex.13, p.267, [**11**], Schur's P-function $P_\mu(x_1,\dots,x_n)$ is defined to be

$$\frac{\mathrm{Pf}_\mu(x_1,\dots,x_n)}{\mathrm{Pf}_\emptyset(x_1,\dots,x_n)},$$

where it is well-known that $\mathrm{Pf}_\emptyset(x_1,\dots,x_n) = \prod_{1\le i<j\le n}\frac{x_i-x_j}{x_i+x_j}$. Meanwhile, by (8.7), p.253, [**11**], Schur's Q-function $Q_\mu(x_1,\dots,x_n)$ is defined to be $2^{\ell(\lambda)}P_\mu(x_1,\dots,x_n)$.

In this section, we consider a weighted sum of Schur's P-functions and Q-functions, i.e.

$$\xi_N(a,b,c,d;X_n) = \sum_{\substack{\mu\\ \mu_1\le N}} \omega(\mu)P_\mu(x_1,\dots,x_n),$$

$$\eta_N(a,b,c,d;X_n) = \sum_{\substack{\mu\\ \mu_1\le N}} \omega(\mu)Q_\mu(x_1,\dots,x_n),$$

where the sums run over all strict partitions μ such that each part of μ is less than or equal to N. More generally, we can unify these problems to finding the following sum:

$$\zeta_N(a,b,c,d;z;X_n) = \sum_{\substack{\mu\\ \mu_1\le N}} \omega(\mu)z^{\ell(\mu)}P_\mu(x_1,\dots,x_n), \tag{5.1}$$

where the sum runs over all strict partitions μ such that each part of μ is less than or equal to N. One of the main results of this section is that $\zeta_N(a,b,c,d;z;X_n)$ can be expressed by a Pfaffian. Further, let us put

$$\zeta(a,b,c,d;z;X_n) = \lim_{N\to\infty}\zeta_N(a,b,c,d;z;X_n) = \sum_\mu \omega(\mu)z^{\ell(\mu)}P_\mu(X_n), \tag{5.2}$$

where the sum runs over all strict partitions. We also write

$$\xi(a,b,c,d;X_n) = \zeta(a,b,c,d;1;X_n) = \sum_\mu \omega(\mu)P_\mu(X_n),$$

where the sum runs over all strict partitions. Then we have the following theorem:

Theorem 5.1. Let n be a positive integer. Then

$$\zeta(a,b,c,d;z;X_n) = \begin{cases} \mathrm{Pf}\,(\gamma_{ij})_{1\le i<j\le n}\,/\,\mathrm{Pf}_\emptyset(X_n) & \text{if } n \text{ is even,} \\ \mathrm{Pf}\,(\gamma_{ij})_{0\le i<j\le n}\,/\,\mathrm{Pf}_\emptyset(X_n) & \text{if } n \text{ is odd,} \end{cases} \tag{5.3}$$

where

$$\gamma_{ij} = \frac{x_i - x_j}{x_i + x_j} + u_{ij} z + v_{ij} z^2 \tag{5.4}$$

with

$$u_{ij} = \frac{a \det \begin{pmatrix} x_i + b x_i^2 & 1 - abx_i^2 \\ x_j + b x_j^2 & 1 - abx_j^2 \end{pmatrix}}{(1 - abx_i^2)(1 - abx_j^2)}, \tag{5.5}$$

$$v_{ij} = \frac{abcx_i x_j \det \begin{pmatrix} x_i + a x_i^2 & 1 - a(b+d)x_i^2 - abdx_i^3 \\ x_j + a x_j^2 & 1 - a(b+d)x_j^2 - abdx_j^3 \end{pmatrix}}{(1 - abx_i^2)(1 - abx_j^2)(1 - abcdx_i^2 x_j^2)}, \tag{5.6}$$

if $1 \le i, j \le n$, and

$$\gamma_{0j} = 1 + \frac{ax_j(1 + bx_j)}{1 - abx_j^2} z \tag{5.7}$$

if $1 \le j \le n$.

Especially, when $z = 1$, we have

$$\xi(a,b,c,d;X_n) = \begin{cases} \operatorname{Pf}\left(\widetilde{\gamma}_{ij}\right)_{1 \le i < j \le n} / \operatorname{Pf}_{\emptyset}(X_n) & \text{if } n \text{ is even}, \\ \operatorname{Pf}\left(\widetilde{\gamma}_{ij}\right)_{0 \le i < j \le n} / \operatorname{Pf}_{\emptyset}(X_n) & \text{if } n \text{ is odd}, \end{cases} \tag{5.8}$$

where

$$\widetilde{\gamma}_{ij} = \begin{cases} \frac{1 + ax_j}{1 - abx_j^2} & \text{if } i = 0, \\ \frac{x_i - x_j}{x_i + x_j} + \widetilde{v}_{ij} & \text{if } 1 \le i < j \le n, \end{cases} \textit{with} \tag{5.9}$$

$$\widetilde{v}_{ij} = \frac{a \det \begin{pmatrix} x_i + b x_i^2 & 1 - b(a+c)x_i^2 - abcx_i^3 \\ x_j + b x_j^2 & 1 - b(a+c)x_j^2 - abcx_j^3 \end{pmatrix}}{(1 - abx_i^2)(1 - abx_j^2)(1 - abcdx_i^2 x_j^2)}. \tag{5.10}$$

We can generalize this result in the following theorem (Theorem 5.2) using the generalized Vandermonde determinant used in [**9**]. Let n be an non-negative integer, and let $X = (x_1, \dots, x_{2n})$, $Y = (y_1, \dots, y_{2n})$, $A = (a_1, \dots, a_{2n})$ and $B = (b_1, \dots, b_{2n})$ be $2n$-tuples of variables. Let $V^n(X, Y, A)$ denote the $2n \times n$ matrix whose (i,j)th entry is $a_i x_i^{n-j} y_i^{j-1}$ for $1 \le i \le 2n$, $1 \le j \le n$, and let $U^n(X, Y; A, B)$ denote the $2n \times 2n$ matrix $\begin{pmatrix} V^n(X,Y,A) & V^n(X,Y,B) \end{pmatrix}$. For instance if $n = 2$ then $U^2(X, Y; A, B)$ is

$$\begin{pmatrix} a_1x_1 & a_1y_1 & b_1x_1 & b_1y_1 \\ a_2x_2 & a_2y_2 & b_2x_2 & b_2y_2 \\ a_3x_3 & a_3y_3 & b_3x_3 & b_3y_3 \\ a_4x_4 & a_4y_4 & b_4x_4 & b_4y_4 \end{pmatrix}.$$

Hereafter we use the following notation for n-tuples $X = (x_1, \cdots, x_n)$ and $Y = (y_1, \cdots, y_n)$ of variables:

$$X + Y = (x_1 + y_1, \dots, x_n + y_n), \quad X \cdot Y = (x_1 y_1, \dots, x_n y_n),$$

and, for integers k and l,

$$X^k = (x_1^k, \dots, x_n^k), \quad X^k Y^l = (x_1^k y_1^l, \dots, x_n^k y_n^l).$$

Let $\mathbf{1}$ denote the n-tuple $(1, \dots, 1)$. For any subset $I = \{i_1, \dots, i_r\} \in \binom{[n]}{r}$, let X_I denote the r-tuple $(x_{i_1}, \dots, x_{i_r})$.

Theorem 5.2. *Let $q = abcd$. If n is an even integer, then we have*

$$\xi(a,b,c,d;X_n) = \sum_{r=0}^{n/2} \sum_{I \in \binom{[n]}{2r}} \frac{(-1)^{|I| - \binom{r+1}{2}} a^r q^{\binom{r}{2}}}{\prod_{i \in I}(1 - abx_i^2)} \prod_{\substack{i,j \in I \\ i<j}} \frac{x_i + x_j}{(x_i - x_j)(1 - qx_i^2 x_j^2)}$$
$$\times \det U^r(X_I^2, \mathbf{1} + qX_I^4, X_I + bX_I^2, \mathbf{1} - b(a+c)X_I^2 - abcX_I^3). \tag{5.11}$$

If n is an odd integer, then we have

$$\xi(a,b,c,d;X_n)=\sum_{m=1}^{n}\frac{1+ax_m}{1-abx_m^2}\sum_{r=0}^{(n-1)/2}\sum_{I\in\binom{[n]\setminus\{m\}}{2r}}\frac{(-1)^{|I|-\binom{r+1}{2}}a^rq^{\binom{r}{2}}}{\prod_{i\in I}(1-abx_i^2)}\prod_{i\in I}\frac{x_m+x_i}{x_m-x_i}$$

$$\times\prod_{\substack{i,j\in I\\ i<j}}\frac{x_i+x_j}{(x_i-x_j)(1-qx_i^2x_j^2)}\cdot\det U^r(X_I^2,\mathbf{1}+qX_I^4,X_I+bX_I^2,\mathbf{1}-b(a+c)X_I^2-abcX_I^3). \tag{5.12}$$

THEOREM 5.3. Let $q=abcd$. If n is an even integer, then $\zeta(a,b,c,d;z;X_n)$ is equal to

$$\sum_{r=0}^{n/2}z^{2r}\sum_{I\in\binom{[n]}{2r}}\frac{(-1)^{|I|-\binom{r+1}{2}}(abc)^rq^{\binom{r}{2}}\prod_{i\in I}x_i}{\prod_{i\in I}(1-abx_i^2)}\prod_{\substack{i,j\in I\\ i<j}}\frac{x_i+x_j}{(x_i-x_j)(1-qx_i^2x_j^2)}$$

$$\times\det V^r(X_I^2,\mathbf{1}+qX_I^4,X_I+aX_I^2,\mathbf{1}-a(b+d)X_I^2-abdX_I^3)$$

$$+\sum_{r=0}^{n/2}z^{2r-1}\sum_{I\in\binom{[n]}{2r}}\sum_{\substack{k<l\\ k,l\in I}}\frac{(-1)^{|I|-\binom{r}{2}-1}a^rb^{r-1}c^{r-1}q^{\binom{r-1}{2}}\{1+b(x_k+x_l)+abx_kx_l\}\prod_{i\in I'}x_i}{\prod_{i\in I}(1-abx_i^2)}$$

$$\times\frac{\prod_{\substack{i,j\in I\\ i<j}}(x_i+x_j)\cdot\det V^{r-1}(X_{I'}^2,\mathbf{1}+qX_{I'}^4,X_{I'}+aX_{I'}^2,\mathbf{1}-a(b+d)X_{I'}^2-abdX_{I'}^3)}{\prod_{\substack{i,j\in I'\\ i<j}}(x_i-x_j)(1-qx_i^2x_j^2)}, \tag{5.13}$$

where $I'=I\setminus\{k,l\}$.

References

[1] G. E. Andrews, "On a partition function of Richard Stanley", *Electron. J. Combin.* **11(2)** (2004), #R1.

[2] G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Cambridge University Press, Cambridge, Second edition, 2004.

[3] C. Boulet, "A four parameter partition identity", `arXiv:math.CO/0308012`, to appear in Ramanujan J.

[4] D.P. Gupta and D.R. Masson, "Solutions to the associated q-Askey-Wilson polynomial recurrence relation", `arXiv:math.CA/9312210`.

[5] M.E.H. Ismail and M. Rahman "The associated Askey-Wilson polynomials", *Trans. Amer. Math. Soc.* **328** (1991), 201 – 237.

[6] 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).

[7] C. Krattenthaler, "Advanced determinant calculus", *Seminaire Lotharingien Combin.* **42** *("The Andrews Festschrift")* (1999), Article B42q, 67.

[8] M. Ishikawa, "Minor summation formula and a proof of Stanley's open problem", `arXiv:math.CO/0408204`.

[9] M. Ishikawa, S. Okada, H. Tagawa and J. Zeng, "Generalizations of Cauchy's determinant and Schur's Pfaffian", `arXiv:math.CO/0411280`, to appear in Adv. in Appl. Math.

[10] M. Ishikawa and M. Wakayama, "Minor summation formula of Pfaffians", *Linear and Multilinear Algebra* **39** (1995), 285-305.

[11] I.G. Macdonald, *Symmetric functions and Hall polynomials, 2nd Edition*, Oxford University Press, (1995).

[12] A. Sills, "A combinatorial proof of a partition identity of Andrews and Stanley", Journal of Mathematics and Mathematical Sciences 2004/47, 2495-2503.

[13] R. P. Stanley, "Some remarks on sign-balance and maj-balanced posets", Adv. in Appl. Math, 34(2005), 880–902.

[14] A. J. Yee, "On partition functions of Andrews and Stanley", J. Combin. Theory Ser.A **13** (2004), 135–139.

FACULTY OF EDUCATION, TOTTORI UNIVERSITY, KOYAMA, TOTTORI, JAPAN
E-mail address: `ishikawa@fed.tottori-u.ac.jp`

INSTITUT CAMILLE JORDAN, UNIVERSITÉ CLAUDE BERNARD LYON I, 43, BOULEVARD DU 11 NOVEMBRE 1918, 69622 VILLEURBANNE CEDEX, FRANCE
E-mail address: `zeng@math.univ-lyon1.fr`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Bijections of trees arising from Voiculescu's free probability theory

Artur Jeż and Piotr Śniady

Abstract. We present a bijective proof of the multidimensional generalizations of the Cauchy identity. Our bijection uses oriented planar trees equipped with some linear orders. The considered identities play an important role in the theory of operator algebras and our bijective prove can be used to prove multidimensional analogues of the arc-sine law in classical probability theory.

Résumé. Nous présentons une preuve bijective des généralisations multidimensionnelles de l'identité de Cauchy. Notre bijection emploie les arbres planaires orientés équipés de quelques ordres linéaires. Les identités considérés jouent un rôle important dans la théorie d'algèbres d'opérateur et notre bijection peut être employé pour prouver des analogues multidimensionnels de la loi d'arcsinus dans la théorie des probabilités classique.

1. Introduction

1.1. How to generalize the Cauchy identity? Cauchy identity states that for each nonnegative integer l

$$2^{2l} = \sum_{p+q=l} \binom{2p}{p}\binom{2q}{q}, \tag{1}$$

where the sum runs over nonnegative integers p, q. Cauchy identity and its bijective proof have important implications to the classical probability theory since they can be used to extract some information about random walks and arc-sine law [**Śni04**], it is therefore very tempting to look for some more identities which would share some resemblance to the Cauchy identity. Such identities could shed some light on the properties of the random walks in higher dimensions.

Guessing how the left-hand side of (1) could be generalized is not difficult and something like m^{ml} is a reasonable candidate. Unfortunately, it is by no means clear which sum should replace the right-hand side of (1). The strategy of writing down lots of wild and complicated sums with the hope of finding the right one by accident is predestined to fail. It is much more reasonable to find some combinatorial objects which are counted by the right-hand side of (1) and then to find a reasonable generalization of these objects.

For fixed integers $p, q \geq 0$ we consider the tree from Figure 1. Every edge of this tree is oriented and it is a good idea to regard these edges as one-way-only roads: if vertices x and y are connected by an edge and the arrow points from y to x then the travel from y to x is permitted but the travel from x to y is not allowed. This orientation defines a partial order $\prec$ on the set of the vertices: we say that $x \prec y$ if it is possible to travel from the vertex y to the vertex x by going through a number of edges (in order to remember this convention we suggest the Reader to think that $\prec$ is a simplified arrow $\leftarrow$). Let $<$ be a total order on the set of the vertices. We say that $<$ is compatible with the orientations of the edges if for all pairs of vertices x, y such that $x \prec y$ we also have $x < y$. It is very easy to see that for the tree from Figure 1 there are $\binom{2p}{p}\binom{2q}{q}$ total orders $<$ which are compatible with the orientations of the edges; this cardinality coincides with the

2000 *Mathematics Subject Classification.* Primary 05C05; Secondary 05A19, 05C30.

Key words and phrases. trees, bijections.

The second author thanks Kenneth Dykema for introducing him into the subject. Research supported by EU Research Training Network "QP-Applications", contract HPRN-CT-2002-00279 and European Commission Marie Curie Host Fellowship for the Transfer of Knowledge "Harmonic Analysis, Nonlinear Analysis and Probability" MTKD-CT-2004-013389. During the research the second author was a holder of a scholarship of European Post-Doctoral Institute for Mathematical Sciences.

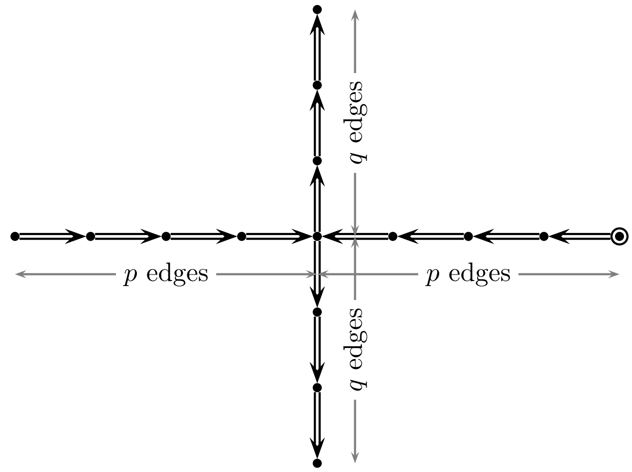


Figure 1. There are $\binom{2p}{p}\binom{2q}{q}$ total orders $<$ on the vertices of this oriented tree which are compatible with the orientation of the edges.

summand on the right-hand side of (1). It remains now to find some natural way of generating the trees of the form depicted on Figure 1 with the property $p+q=l$. We shall do it in the following.

1.2. Quotient graphs and quotient trees. We recall now the construction of Dykema and Haagerup [**DH04a**]. For integer $k \geq 1$ let G be an oriented k–gon graph with consecutive vertices $v_1, \dots, v_k$ and edges $e_1, \dots, e_k$ (edge e_i connects vertices v_i and v_{i+1}). The vertex v_1 is distinguished, see Figure 2. We encode the information about the orientations of the edges in a sequence $\epsilon(1), \dots, \epsilon(k)$ where $\epsilon(i) = +1$ if the arrow points from v_{i+1} to v_i and $\epsilon(i) = -1$ if the arrow points from v_i to v_{i+1}. The graph G is uniquely determined by the sequence ϵ and sometimes we will explicitly state this dependence by using the notation G_ϵ.

Let $\sigma = \big\{\{i_1, j_1\}, \dots, \{i_{k/2}, j_{k/2}\}\big\}$ be a pairing of the set $\{1, \dots, k\}$, i.e. pairs $\{i_m, j_m\}$ are disjoint and their union is equal to $\{1, \dots, k\}$. We say that σ is compatible with ϵ if

$$\epsilon(i) + \epsilon(j) = 0 \qquad \text{for every } \{i, j\} \in \sigma. \tag{2}$$

It is a good idea to think that σ is a pairing between the edges of G, see Figure 2. For each $\{i, j\} \in \sigma$ we identify (or, in other words, we glue together) the edges e_i and e_j in such a way that the vertex v_i is identified with v_{j+1} and vertex v_{i+1} is identified with v_j and we denote by T_σ the resulting quotient graph. Since each edge of T_σ origins from a pair of edges of G, we draw all edges of T_σ as double lines. The condition (2) implies that each edge of T_σ carries a natural orientation, inherited from each of the two edges of G it comes from, see Figure 3.

From the following on, we consider only the case when the quotient graph T_σ is a tree. One can show [**DH04a**] that the latter holds if and only if the pairing σ is non–crossing [**Kre72**]; in other words it is not possible that for some $p < q < r < s$ we have $\{p, r\}, \{q, s\} \in \sigma$. The name of the non–crossing pairings comes from their property that on their graphical depictions (such as Figure 2) the lines do not cross. Let the root R of the tree T_σ be the vertex corresponding to the distinguished vertex v_1 of the graph G.

1.3. How to generalize the Cauchy identity? (continued). Let us come back to the discussion from Section 1.1. We consider the polygon G_ϵ corresponding to

$$\epsilon = (\underbrace{-1}_{l \text{ times}}, \underbrace{+1}_{l \text{ times}}, \underbrace{-1}_{l \text{ times}}, \underbrace{+1}_{l \text{ times}}).$$

All possible non-crossing pairings σ which are compatible with ϵ are depicted on Figure 4 and it easy to see that the corresponding quotient tree T_σ has exactly the form depicted on Figure 1.

In this way we managed to find relatively natural combinatorial objects, the number of which is given by the right-hand side of the Cauchy identity (1). After some guesswork we end up with the following conjecture (please note that the usual Cauchy identity (1) corresponds to $m = 2$).

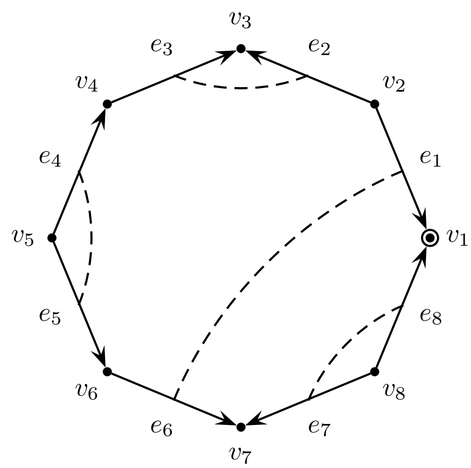


FIGURE 2. A graph G_ϵ corresponding to the sequence $\epsilon = (+1, -1, +1, +1, -1, -1, +1, -1)$. The dashed lines represent the pairing $\sigma = \big\{\{1, 6\}, \{2, 3\}, \{4, 5\}, \{7, 8\}\big\}$.

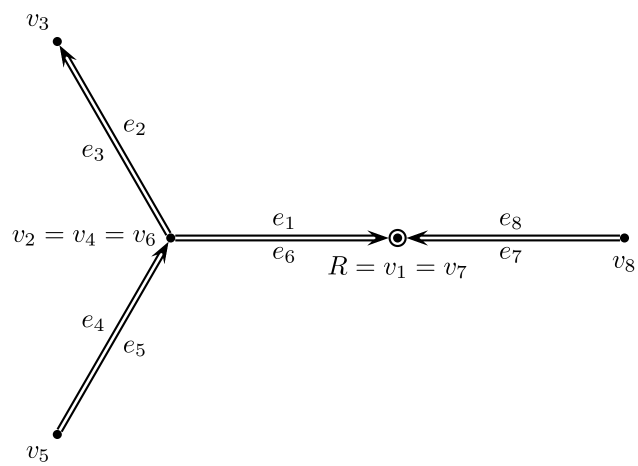


FIGURE 3. The quotient graph T_σ corresponding to the graph from Figure 2. The root R of the tree T_σ is encircled.

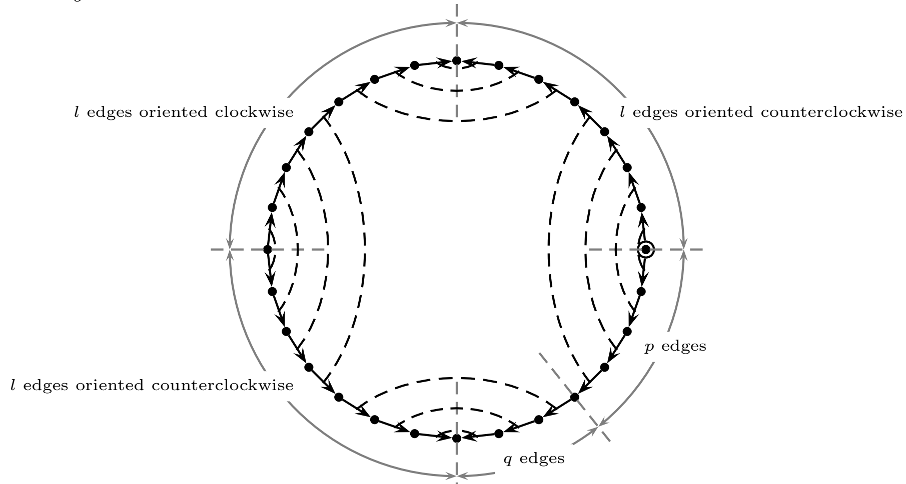


FIGURE 4. A graph T corresponding to sequence $\epsilon = (\underbrace{-1}_{l \text{ times}}, \underbrace{+1}_{l \text{ times}}, \underbrace{-1}_{l \text{ times}}, \underbrace{+1}_{l \text{ times}})$. The dashed lines denote a pairing σ for which the quotient graph T_σ is depicted on Figure 1.

THEOREM 1 (Generalized Cauchy identity). *For integers $l, m \geq 1$ there are exactly m^{ml} pairs $(\sigma, <)$, where σ is a non-crossing pairing compatible with*

$$\epsilon = (\underbrace{\underbrace{-1}_{l \text{ times}}, \underbrace{+1}_{l \text{ times}}, \underbrace{-1}_{l \text{ times}}, \underbrace{+1}_{l \text{ times}}, \dots}_{2m \text{ blocks, i.e. total of } 2ml \text{ elements}}) \tag{3}$$

and $<$ is a total order on the vertices of T_σ which is compatible with the orientations of the edges.

Above we provided only vague heuristical arguments why the above conjecture could be true. Surprisingly, as we shall see in the following, Theorem 1 is indeed true.

The formulation of Theorem 1 is combinatorial and therefore appears to be far from its motivation, the usual Cauchy identity (1), which is formulated algebraically, nevertheless for each fixed value of m one can enumerate all 'classes' of pairings compatible with (3) and for each class count the number of compatible orders $<$. To give to the Reader a flavor of the algebraic implications of Theorem 1, we present the case of $m = 3$ [**DY03**]:

$$3^{3l} = \sum_{p+q=l} \binom{3p}{p,p,p} \binom{3q}{q,q,q} + +3 \sum_{\substack{p+q+r=l-1 \\ r'+q'=r+q+1 \\ p''+r''=p+r+1}} \binom{2p+p''}{p,p,p''} \binom{2q+q'}{q,q,q'} \binom{r+r'+r''}{r,r',r''}. \tag{4}$$

and the case of $m = 4$ [**Śni03**]:

$$\begin{aligned} 4^{4k} = &\sum_{p+q=k} \binom{4p}{p,p,p,p} \binom{4q}{q,q,q,q} + 8 \sum_{\substack{p+q+r=k-1 \\ p'+q'=p+q+1 \\ p''+q''=p+q+1 \\ q'''+r'''=q+r+1}} \binom{2p+p'+p''}{p,p,p',p''} \binom{q+q'+q''+q'''}{q,q',q'',q'''} \binom{3r+r'''}{r,r,r,r'''} + \\ &+ 4 \sum_{\substack{p+q'+r'=k-1 \\ p+q''+r''=k-1 \\ p'''+q'''=p+q'+1 \\ p''''+q''''=p+q''+1}} \binom{2p+p'''+p''''}{p,p,p''',p''''} \binom{q'+q'''}{q',q'''} \binom{q''+q''''}{q'',q''''} \binom{2r''}{r'',r''} \binom{2r'}{r',r'} \binom{q'+q''+q'''+q''''+2r'+2r''+2}{q'+q'''+2r'+1, q''+q''''+2r''+1} + \\ &+ 8 \sum_{\substack{p+q+r+s=k-2 \\ q'+r'=q+r+s+2 \\ p''+r''=p+q+r+2}} \binom{2p}{p,p} \binom{q+q'}{q,q'} \binom{r+r''}{r,r''} \binom{2s}{s,s} \binom{3p+p''+2q+q'+2}{2p+q+q'+1, p+q+1, p''} \binom{2r+r'+r''+3s+2}{r+r''+2s+1, r+s+1, r'}. \end{aligned} \tag{5}$$

1.4. Bijective proof of generalized Cauchy identities. Theorem 1 was conjectured by Dykema and Haagerup [**DH04a**] and its first proof (analytic one) was given by the second-named author [**Śni03**]. Another analytic proof was given by Aagaard and Haagerup [**AH04**]. The main result of this article (which is a shortened and edited version of [**Śni04**]) is the first bijective proof of Theorem 1, formulated explicitly as the following theorem.

THEOREM 2 (The main result). *Let integers $l, m \geq 1$ be given. We set $L = lm + 1$ and*

$$\epsilon_i = (\underbrace{\underbrace{(-1)^{i-1}}_{l \text{ times}}, \underbrace{(-1)^{i}}_{l \text{ times}}}_{i \text{ times, i.e. a total of } 2li \text{ elements}}) \qquad \text{for } 1 \leq i \leq m.$$

Note that ϵ_m coincides (up to a possible sign change) with (3). The function described in this article provides a bijection between

(α) *the set of pairs $(\sigma, <)$, where σ is a pairing compatible with ϵ_m and $<$ is a total order on the vertices of T_σ which is compatible with the orientations of the edges;*

(β) *the set of tuples $(B_1, \dots, B_m)$, where $B_1, \dots, B_m$ are disjoint sets such that $B_1 \cup \cdots \cup B_m = \{1, 2, \dots, L\}$ and*

$$|B_1| + \cdots + |B_n| \leq ln$$

holds true for each $1 \leq n \leq m - 1$;

Alternatively, set (β) can be described as

(γ) *the set of sequences* $(a_1, \dots, a_L)$ *such that* $a_1, \dots, a_L \in \{1, \dots, m\}$ *and for each* $1 \le n \le m-1$ *at most* ln *elements of the sequence* (a_i) *belong to the set* $\{1, \dots, n\}$*;*

where the bijection between sets (β) *and* (γ) *is given by* $B_j = \{k : a_k = j\}$.

From the Raney lemma [**Ran60**] it follows that the set (β) has m^{ml} elements [**Śni03**] hence Theorem 1 indeed follows from Theorem 2.

2. Quotient trees

In the following we shall discuss some aspects of the quotient trees which were not included in Section 1.2. Sometimes, with a very small abuse of notation, we will denote by the same symbol T_σ the set of the vertices of the tree T_σ.

2.1. Structure of a planar tree. Order $\lhd$. For a non–crossing pairing σ we can describe the process of creating the quotient graph as follows: we think that the edges of the graph G are sticks of equal lengths with flexible connections at the vertices. Graph G is lying on a flat surface in such a way that the edges do not cross. For each pair $\{i, j\} \in \sigma$ we glue together edges e_i and e_j by bending the joints in such a way that the sticks should not cross. In this way T_σ has a structure of a planar tree, i.e. for each vertex we can order the adjacent edges up to a cyclic shift (just like points on a circle). We shall provide an alternative description of this planar structure in the following.

Let us visit the vertices of G in the usual cyclic order $v_1, v_2, \dots, v_k, v_1$ by going along the edges $e_1, \dots, e_k$; by passing to the quotient graph T_σ we obtain a journey on the graph T_σ which starts and ends in the root R. The structure of the planar tree defined above can be described as follows: if we travel on the graphical representation of T_σ by touching the edges by our left hand, we obtain the same journey. For each vertex of T_σ we mark the time we visit it for the first time; comparison of these times gives us a total order $\lhd$, called preorder [**Sta99**], on the vertices of T_σ. For example, in the case of the tree from Figure 3 we have $v_1 \lhd v_2 \lhd v_3 \lhd v_5 \lhd v_8$.

2.2. Catalan sequences. We say that $\epsilon = \big(\epsilon(1), \dots, \epsilon(k)\big)$ is a Catalan sequence if $\epsilon(1), \dots, \epsilon(k) \in \{-1, +1\}$, $\epsilon(1) + \cdots + \epsilon(k) = 0$ and all partial sums are non-negative: $\epsilon(1) + \cdots + \epsilon(l) \ge 0$ for all $1 \le l \le k$. We say that ϵ is anti-Catalan if $-\epsilon$ is Catalan.

LEMMA 3. *For a Catalan sequence ϵ there exists a unique compatible pairing σ with the property that $R \preceq v$ for every vertex $v \in T_\sigma$. For an anti-Catalan sequence ϵ there exists a unique compatible pairing σ with the property that $R \succeq v$ for every vertex $v \in T_\sigma$.*

3. Proof of almost the main result

3.1. Statement of the result. The following result will be crucial for the bijective proof of generalized Cauchy identities in Section 4.

THEOREM 4. *Let $\epsilon = \big(\epsilon(1), \dots, \epsilon(k)\big)$ be a Catalan sequence. The function described in this section provides a bijection between*

(A) *the set of pairs $(\sigma, <)$, where σ is a pairing compatible with ϵ and $<$ is a total order on the vertices of T_σ compatible with the orientation of the edges;*

(B) *the set of pairs $(\sigma, <)$, where σ is a pairing compatible with ϵ and $<$ is a total order on the vertices of T_σ with the following two properties:*
- *on the set $\{x \in T_\sigma : x \succeq R\}$ the orders $<$ and $\lhd$ coincide;*
- *for all pairs of vertices $v, w \in T_\sigma$ such that $R \not\preceq v$ and $R \not\preceq w$ we have*

$$v \prec w \implies v < w.$$

PROOF. In this article we will present only the bijection without presenting its inverse and without any proofs which can be found in [**Śni04**].

Our bijection will be given by repeating the following procedure: if the pair $(\sigma, <)$ is as in (B) then we our algorithm finishes. Otherwise, let D be the maximal element (with respect to the order $<$) such that $D \succ R$ and such that on the subtree $U = \{x : x \succeq R \text{ and } x < D\}$ the orders $<$ and $\lhd$ coincide. The vertex D is a leaf of the tree U which is not maximal in U (with respect to the order $\lhd$); otherwise this would contradict the maximality of D. We start in D a walk on the graph U with the first step going

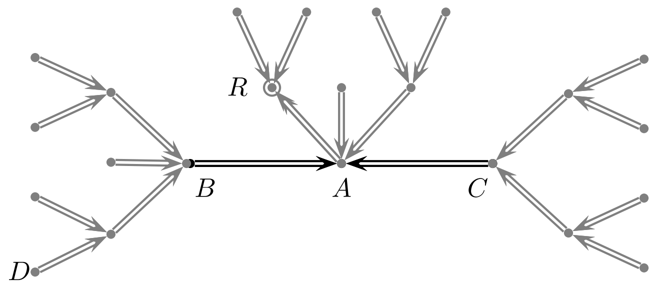


Figure 5. The case $D \neq B$. The order of the vertices is given by $R \leq A < B < C < D$. Note that only the edges belonging to the subtree U are displayed.

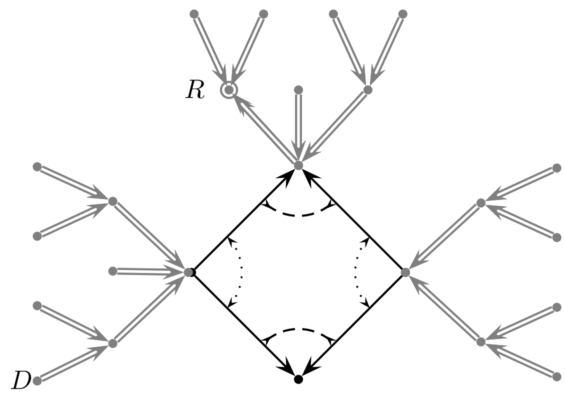


Figure 6. The tree from Figure 5 after ungluing the edges BA and CA.

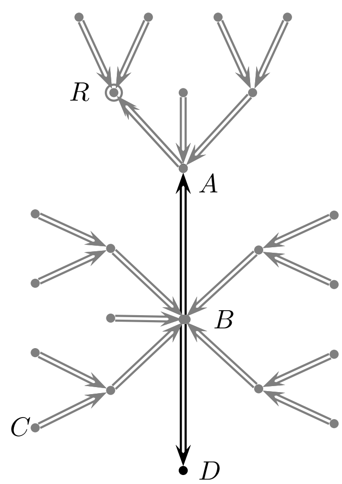


Figure 7. The tree from Figure 5 after regluing the edges BA and CA in a different way. Please notice the change of the labels of the vertices A, B, C, D.

towards the root R, always touching the edges by our left hand (as we did in Section 2.1) and we denote by $w_0 = D, w_1, w_2, \dots$ the consecutive vertices we visit on our journey. Let n be the smallest number for which the arrow on the edge connecting w_n and w_{n+1} points from w_{n+1} towards w_n; we denote $B = w_{n-1}$, $A = w_n$, $C = w_{n+1}$.

Let us consider the case when $B \neq D$, cf. Figure 5. Each of the edges BA and CA of the quotient graph T_σ was created by gluing a pair of edges of the graph G; let us unglue these four edges of G, cf. Figure 6 and let us glue these four edges in pairs in a different way, cf Figure 7. In this way we obtain a quotient

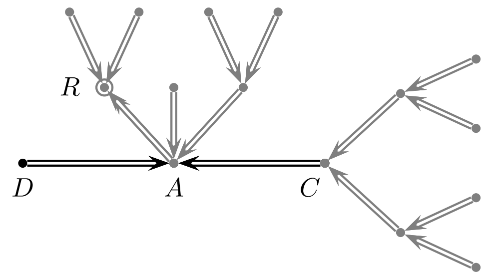


FIGURE 8. The case $D = B$. The order of vertices is given by $R \leq A < C < D$.

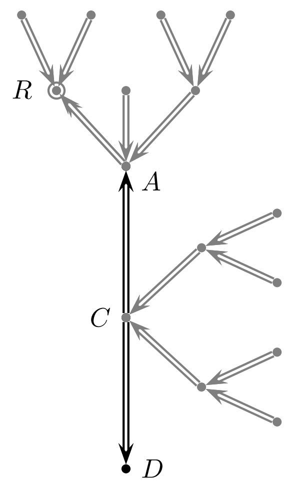


FIGURE 9. The tree from Figure 8 after regluing the edges DA and CA in a different way. Please notice the change of the labels of the vertices A, C, D.

graph $T_{\sigma'}$, where σ' is a pairing of edges obtained from σ by changing connections between certain four edges. Figure 5 and Figure 7 show an identification between the vertices of T_σ and $T_{\sigma'}$; please note that this identification is nontrivial only on the vertices A, B, C, D. We define the order $<$ on $T_{\sigma'}$ to be the inherited order $<$ from T_σ under the above identification of the vertices.

We consider now the case when $B = D$, cf. Figure 8. Similarly as above, we unglue and reglue in a different way edges DA and CA and thus we obtain a tree $T_{\sigma'}$ depicted on Figure 9. Figure 8 and Figure 9 show the identification between the vertices of T_σ and the vertices of $T_{\sigma'}$ and we define the order $<$ on $T_{\sigma'}$ to be the inherited order $<$ from T_σ.

After a finite number of steps the above procedure will eventually stop. □

REMARK 5. For each pair $(\sigma, <)$ from the set (A) and the corresponding pair $(\sigma', <)$ from the set (B) there is a canonical unique bijection j mapping the vertices of T_σ onto the vertices of $T_{\sigma'}$ with the property that for all $v, w \in T_\sigma$ the condition $v < w$ holds if and only if $j(v) < j(w)$. In fact this identification is very easy to see since the bijection from Theorem 4 is a composition of a number of elementary operations. Each such operation is either a replacement of Figure 5 by Figure 7 or replacement of Figure 8 by Figure 9 and for each such a replacement the corresponding identification preserves the labels of the vertices.

4. Proof of the main result

PROOF. We shall construct now the main result of the article: the bijection announced in Theorem 2. In this article we will present only the bijection without presenting its inverse and without any proofs which can be found in [**Śni04**].

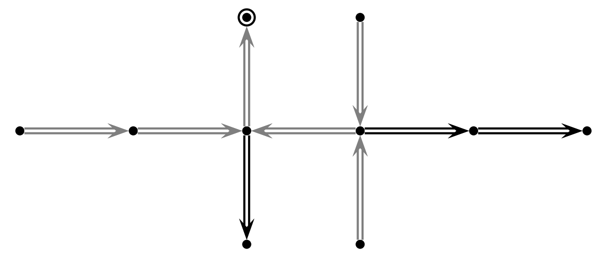


Figure 10. Example of a tree $T_{\tilde{\sigma}}$. A subtree $\{v : R \preceq v\}$ was marked in gray.

Firstly, observe that the order $<$ on the vertices of the tree T_σ can be alternatively described by labeling the vertices by the numbers from the set $\{1, 2, \dots, L\}$ in such a way that each number appears exactly once and the order of the labels coincides with the order $<$ on the vertices.

Our algorithm consists of $m-1$ steps; in the first step the variable i takes the value $i := m$ and after each step its value decreases by one. At the beginning of each step we start with a tree T_σ, where σ is a pairing compatible with ϵ_i such that some of the vertices are labeled by the numbers from the set $\{1, 2, \dots, L\}$ and some vertices might be unlabeled (in the first step $i = m$ there are no unlabeled vertices) and in this step we will construct the set B_i.

Let us consider the case when i is odd. We define a total order $<$ on the vertices of T_σ as follows: for a pair of vertices v, w which carry some labels we set $v < w$ if and only if the label of v is smaller than the label of w; if v has no label and w has a label then $v < w$; if both v and w have no labels then $v < w$ if and only if $v \lhd w$. In this way $(\sigma, <)$ is as prescribed in point (A) of Theorem 4.

Let $(\tilde{\sigma}, <)$ denote the corresponding element of the point (B). We consider the canonical identification of the vertices of the tree T_σ with the vertices of the tree $T_{\tilde{\sigma}}$, as described in Remark 5; in this way some of the vertices of the tree $T_{\tilde{\sigma}}$ are labeled by the numbers from the set $\{1, \dots, L\}$. We consider a subtree $U = \{x \in T_{\tilde{\sigma}} : R \preceq x\}$. We define B_i to be the set of the labels on the vertices of U and we remove all labels from the vertices of U.

Each edge of $T_{\tilde{\sigma}}$ consists of two edges of the graph G; let us unglue all the edges belonging to the tree U. We denote by T' the resulting graph, cf Figure 10 and Figure 11. The sequence ϵ_{i-1} can be obtained from the sequence ϵ_i by removal of the first l and the last l elements therefore the polygonal graph $G_{\epsilon_{i-1}}$ can be obtained from the graph G_{ϵ_i} by removing two groups (of l edges each) surrounding the distinguished vertex R from both sides; clearly these $2l$ edges must be among the unglued ones in the graph T'. We denote by T'' the graph obtained from T' by the removal of these $2l$ edges, cf Figure 12.

Please note that T'' can be obtained from the polygonal graph $G_{\epsilon_{i-1}}$ by gluing some pairs of edges hence it can be viewed as a certain polygonal graph $G_{\epsilon'}$ with a number of trees attached to it. The sequence ϵ' can be obtained from ϵ_{i-1} by a removal of a number of blocks of consecutive elements, provided the sum of elements of each block is equal to zero. Since ϵ_{i-1} is anti-Catalan, ϵ' is anti-Catalan as well. We denote by $T_{\sigma'}$ the tree resulting from T'' by gluing the edges constituting $G_{\epsilon'}$ by the pairing given by Lemma 3 applied to ϵ'; please note that in this way we defined implicitly the pairing σ' compatible with ϵ_{i-1}, cf Figure 13. Thus, the description of the step of the algorithm in the case when i is odd is finished.

To cover the case when i is even we can simply reverse the orientations on all edges (which corresponds to a change of signs in the sequence ϵ_i) and consider the opposite order on the set $\{1, \dots, L\}$; since sequence $-\epsilon_i$ is Catalan and $-\epsilon_{i-1}$ is anti-Catalan we reduced the situation to the case considered previously.

Our algorithm takes a particularly simple form for $i = 1$; we simply set B_1 to be the set of the labels of the tree T_σ and the algorithm stops. □

5. Combinatorial calculus: how to convert an analytic proof into a bijection?

The bijection presented in this article might look artificial and it is by no means clear how the authors invented it. It turns out that there is a very systematic way of constructing this bijection given by careful analysis of the analytic proof of generalized Cauchy identities given by the second-named author [**Śni03**]. In this analytic proof we associated to oriented trees certain polynomials and we proved that these polynomials fulfill recursion relation analogous to the one fulfilled by Abel polynomials. It turns out that if we replace the

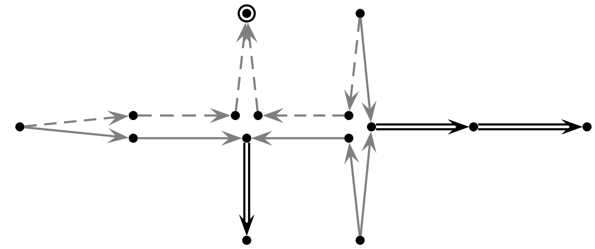


FIGURE 11. The graph T' obtained for $T_{\sigma'}$ depicted on Figure 10.

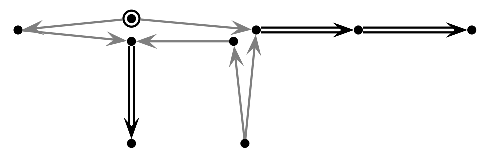


FIGURE 12. The graph T'' is obtained from T' depicted on Figure 11 by removal of the dashed edges. The graph T'' can be regarded as a certain polygonal graph $G_{\epsilon'}$ with a number of trees attached to it.

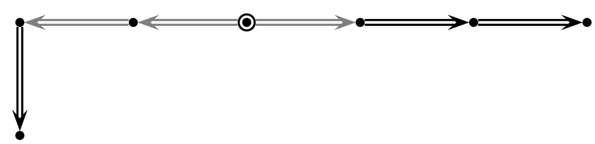


FIGURE 13. Tree $T_{\sigma'}$ is obtained from the graph T'' depicted on Figure 12 by gluing edges as prescribed in Lemma 3.

usual differential calculus by a *combinatorial calculus* in which the role of polynomials is played by certain graphs and oriented sets then the analytic proof from [**Śni03**] is valid also in this more general setup and it determines uniquely the bijection presented in this article [**JŚ06a**].

6. Postscript: operator algebras, free probability and triangular operator T

The story presented in Sections 1.1 and 1.3 is too beautiful to be true. In fact, it is not how the generalized Cauchy identities were discovered. In this section we will present the true story which also gives very strong motivations for studying these identities.

6.1. Invariant subspace conjecture. The *Voiculescu's free probability* [**VDN92, HP00**] is a non-commutative probability theory with the classical notion of independence replaced by the notion of *freeness*. Natural examples which fit nicely into the framework of the free probability include large random matrices, free products of von Neumann algebras and asymptotics of large Young diagrams. Families of operators which arise in the free probability are, informally speaking, very non-commutative and for this reason they are perfect candidates for counterexamples to the conjectures in the theory of operator algebras [**Voi96**].

Dykema and Haagerup [**DH04a**] suggested that free probability could be used to construct a counterexample for the famous *invariant subspace conjecture* (this conjecture asks if for every bounded operator x acting on an infinite-dimensional Hilbert space $\mathcal{H}$ there exists a closed subspace $\mathcal{K} \subset \mathcal{H}$ such that $\mathcal{K}$ is nontrivial in the sense that $\mathcal{K} \neq \{0\}$, $\mathcal{K} \neq \mathcal{H}$ and which is an invariant subspace of x). They also described explicitly a very good candidate for such a counterexample, namely *the triangular operator* T [**DH04a**].

For more details on the history of the search of such a counterexample within free probability theory we refer to [**Śni04**].

6.2. Combinatorics of the triangular operator T. Even though the primary description of the triangular operator T was purely analytic as a limit of certain random matrices, already in the original article [**DH04a**] Dykema and Haagerup gave a purely combinatorial description of this operator and we will present it in the following.

The triangular operator T is an element of a certain algebra (*finite von Neumann algebra*) equipped with a functional ϕ. The elements of such algebras can be uniquely determined by the values of ϕ on all polynomials in T and $T^\star$ therefore we need to specify the numbers $\phi(T^{\epsilon(1)} \cdots T^{\epsilon(n)})$ for any sequence $\epsilon(1), \dots, \epsilon(n) \in \{1, \star\}$. Dykema and Haagerup proved that that $(n/2+1)!\ \phi(T^{\epsilon(1)} \cdots T^{\epsilon(n)})$ is equal to the number of pairs $(\sigma, <)$ such that σ is a pairing compatible with ϵ and $<$ is a total order on the vertices of T_σ which is compatible with the orientation of the edges (please notice that the sequence ϵ considered above takes the values 1 and $\star$ while in the rest of this article we used the convention that ϵ takes the values $+1$ and -1, this difference is irrelavant). The Reader may easily see that the latter definition of T is very closely related to the results presented in this paper; in particular Theorem 1 can be now equivalently stated as follows (in fact it is the form in which Dykema and Haagerup stated originally their conjecture [**DH04a**]):

Theorem 6. *If* $l, m \geq 1$ *are integers then*

$$\phi\left[\left(T^l (T^\star)^l \right)^m \right] = \frac{m^{ml}}{(ml+1)!}.$$

Theorem 1 and Theorem 6 were conjectured by Dykema and Haagerup [**DH04a**] in the hope that they might be useful in the study of spectral properties of T. Literally speaking, this hope turned out to be wrong since the later construction of the hyperinvariant subspaces of T by Dykema and Haagerup [**DH04b, Haa02**] did not make use of Theorem 1 and Theorem 6, however it made use of one of the auxiliary results used in our proof [**Śni03**] of these theorems. In this way, indirectly, Theorem 1 and Theorem 6 turned out to be indeed helpful for their original purpose. Later on Aagaard and Haagerup [**AH04**] gave a different analytic proof of the generalized Cauchy identities based on very clever matrix manipulations.

As we already mentioned, Dykema and Haagerup [**DH04b, Haa02**] constructed a family of hyperinvariant subspaces of T and in this way the original motivation for studying the operator T (as a possible counterexample for the invariant subspace conjecture) ended up as a failure. Nevertheless, operator T is still regarded as a canonical example of a quasinilpotent operator and its deep understanding may give us an insight into the structure of all quasinilpotent operators.

6.3. Applications in classical probability theory. The generalized Cauchy identities and their bijective proof can be used [**JŚ06b**] to extract some information about multidimensional random walks and Brownian motions.

References

[AH04] Lars Aagaard and Uffe Haagerup, *Moment formulas for the quasi-nilpotent DT-operator*, Internat. J. Math. **15** (2004), no. 6, 581–628. MR MR2078883

[DH04a] Ken Dykema and Uffe Haagerup, *DT-operators and decomposability of Voiculescu's circular operator*, Amer. J. Math. **126** (2004), no. 1, 121–189. MR 2 033 566

[DH04b] ———, *Invariant subspaces of the quasinilpotent DT-operator*, J. Funct. Anal. **209** (2004), no. 2, 332–366. MR 2 044 226

[DY03] Kenneth Dykema and Catherine Yan, *Generating functions for moments of the quasi–nilpotent DT operator*, Adv. in Appl. Math. **30** (2003), no. 3, 545–561.

[Haa02] U. Haagerup, *Random matrices, free probability and the invariant subspace problem relative to a von Neumann algebra*, Proceedings of the International Congress of Mathematicians, Vol. I (Beijing, 2002) (Beijing), Higher Ed. Press, 2002, pp. 273–290. MR 2004e:46002

[HP00] Fumio Hiai and Dénes Petz, *The semicircle law, free random variables and entropy*, American Mathematical Society, Providence, RI, 2000. MR 2001j:46099

[JŚ06a] Artur Jeż and Piotr Śniady, *Generalized Cauchy identities, trees and multidimensional Brownian motions. Part II: Combinatorial differential calculus*, In preparation, 2006.

[JŚ06b] ———, *Generalized cauchy identities, trees and multidimensional Brownian motions. Part III: Multidimensional arc-sine laws*, In preparation, 2006.

[Kre72] G. Kreweras, *Sur les partitions non croisées d'un cycle*, Discrete Math. **1** (1972), no. 4, 333–350. MR 46 #8852

[Ran60] George N. Raney, *Functional composition patterns and power series reversion*, Trans. Amer. Math. Soc. **94** (1960), 441–451.

[Śni03] Piotr Śniady, *Multinomial identities arising from the free probability theory*, J. Combin. Theory Ser. A **101** (2003), no. 1, 1–19.

[Śni04] ———, *Generalized Cauchy identities, trees and multidimensional Brownian motions. Part I: bijective proof of generalized Cauchy identities*, Preprint `arXiv:math.CO/0412043`, 2004.

[Sta99] Richard P. Stanley, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999. MR 2000k:05026

[VDN92] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free random variables*, American Mathematical Society, Providence, RI, 1992. MR 94c:46133

[Voi96] D. Voiculescu, *The analogues of entropy and of Fisher's information measure in free probability theory. III. The absence of Cartan subalgebras*, Geom. Funct. Anal. **6** (1996), no. 1, 172–199. MR 96m:46119

Institute of Mathematics, University of Wroclaw, pl. Grunwaldzki 2/4, 50-384 Wroclaw, Poland
E-mail address: s141494@math.uni.wroc.pl

Institute of Mathematics, University of Wroclaw, pl. Grunwaldzki 2/4, 50-384 Wroclaw, Poland
E-mail address: Piotr.Sniady@math.uni.wroc.pl
URL: http://www.math.uni.wroc.pl/∼psnia

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Schubert polynomials for the affine Grassmannian

Thomas Lam

Abstract. Confirming a conjecture of Mark Shimozono, we identify polynomial representatives for the Schubert classes of the affine Grassmannian as the k-Schur functions in homology and affine Schur functions in cohomology. Our results rely on Kostant and Kumar's nilHecke ring, work of Peterson on the homology of based loops on a compact group, and earlier work of ours on non-commutative k-Schur functions.

Résumé. Nous prouvons une conjecture de Mark Shimozono en montrant que des representations polynomiales pour les classes de Schubert des Grassmanniennes affines sont des fonctions k-Schur en homologie, et des fonctions de Schur affines en cohomologie. Nous utilisons l'anneau nilHecke de Kostant et Kumar, le travail de Peterson sur l'homologie des circuits basés sur un groupe compact, et notre travail antérieur sur les fonctions de k-Schur non-commutatives.

1. Introduction

This article is an extended abstract of the paper [**11**] with the same title. Some results and many details have been omitted.

In [**3**], Bott calculated the homology and cohomology rings of the based loop spaces ΩK, where K is a compact Lie group. In type A, both $H^*(\Omega SU_n)$ and $H_*(\Omega SU_n)$ can be identified with a ring of symmetric functions: in cohomology as a quotient of the ring of symmetric functions and in homology as a subring of the ring of symmetric functions. Separately, Kostant and Kumar [**8**] have calculated the cohomology rings $H^*(\mathcal{G}/\mathcal{P})$ of homogeneous spaces of Kac-Moody groups in terms of the Schubert classes $\sigma^w \in H^*(\mathcal{G}/\mathcal{P})$. It is well known that when $\mathcal{G}$ is of affine type and $\mathcal{P}$ a maximal parabolic, then $\mathcal{G}/\mathcal{P}$ is homotopy-equivalent to the based loops on the finite-dimensional compact group associated to $\mathcal{G}$. Thus in type $\hat{A}_{n-1}$, we have $H^*(\mathcal{G}/\mathcal{P}) = H^*(\Omega SU(n))$. While some of our results generalize to all Dynkin types, we will restrict ourselves to type A for the remainder of this article.

Our main result is the identification of the Schubert classes $\sigma^w \in H^*(\mathcal{G}/\mathcal{P})$ and $\sigma_w \in H_*(\mathcal{G}/\mathcal{P})$ as explicit symmetric functions. In the homology case, these polynomials are known as the *k-Schur functions*, originally introduced by Lapointe, Lascoux and Morse [**16**] and studied thoroughly by Lapointe and Morse [**13, 14**]. In the cohomology case, these polynomials were introduced by Lapointe and Morse in [**15**] where they were called *dual k-Schur functions* and also studied by myself in [**10**] where they were called *affine Schur functions*. These results were conjectures of Mark Shimozono (in the cohomology case, the conjecture was made precise by Jennifer Morse).

Thus the k-Schur functions $s_\lambda^{(k)}(x)$ and the affine Schur functions $\tilde{F}_\lambda(x)$ can be considered affine homology and cohomology Schubert polynomials respectively. Schubert polynomials for the flag variety were introduced by Lascoux and Schützenberger [**17**] and has led to numerous developments in algebra, geometry and combinatorics. It should be expected that affine Schubert polynomials lead to many exciting developments as well. Note that since $\Omega SU(n)$ is a loop space, its homology $H_*(\Omega SU(n)) = H_*(\mathcal{G}/\mathcal{P})$ is a Hopf-algebra. Our

2000 *Mathematics Subject Classification.* Primary 05E05; Secondary 14N15.

Key words and phrases. Schubert polynomials, symmetric functions, Schubert calculus, affine Grassmannian.

I am indebted to my coauthors Luc Lapointe, Jennifer Morse and Mark Shimozono, with whom I have studied k-Schur functions and the affine Grassmannian for nearly a year. I began working on k-Schur and dual k-Schur functions more than a year ago when Jennifer first introduced them to me, and Mark explained his geometric conjectures to me.

identification of Schubert classes is actually an isomorphism of Hopf-algebras, and gives an interpretation of the Hall inner product as the natural pairing between homology and cohomology. This feature of the affine theory is lacking in the classical finite case. We will only briefly discuss the Hopf-structures in this article.

Our results rely heavily on the *nilHecke ring* $\mathbb{A}$ introduced by Kostant and Kumar [**8**], results of Peterson [**19**] on the homology of based loop spaces, and the *non-commutative k-Schur functions* by the author in [**10**]. The non-commutative k-Schur functions are elements $s_w^{(k)}$ of a commutative subalgebra $\mathbb{B} \subset \mathbb{A}$, which we call the *affine Fomin-Stanley algebra* (since it is closely related to the work in [**6**]), of the nilHecke ring. We showed in [**10**] that $\mathbb{B}$ was isomorphic to a subring of the ring of symmetric functions which can be identified via Bott's result with $H_*(\mathcal{G}/\mathcal{P})$. Peterson has constructed an isomorphism $j : H_*^T(\mathcal{G}/\mathcal{P}) \to Z_{\mathbb{A}}(S)$ of the equivariant homology $H_*^T(\mathcal{G}/\mathcal{P})$ with a certain centraliser subalgebra $Z_{\mathbb{A}}(S) \subset \mathbb{A}$ of the nilHecke ring. We show here that "evaluation at 0" takes $Z_{\mathbb{A}}(S)$ onto $\mathbb{B}$ and that the composition with Peterson's j-homomorphism takes the Schubert classes $\sigma_{(w)}$ to the non-commutative k-Schur functions $s_w^{(k)}$. Kostant and Kumar have calculated the structure constants of $H^*(\mathcal{G}/\mathcal{P})$ in terms of a coproduct Δ on $\mathbb{A}$ and we compute directly that this coproduct, when restricted to the subalgebra $\mathbb{B}$, agrees with the usual coproduct of the symmetric functions. This shows that $\mathbb{B}$, when viewed as a ring of symmetric functions, is Hopf-isomorphic to $H_*(\mathcal{G}/\mathcal{P})$.

There are many open problems related to this work, and we mention a couple: it is natural to ask for representatives in K-theory, in equivariant (co)homology and in quantum cohomology. It is also natural to ask to generalize our work from the affine Grassmannian $\mathcal{G}/\mathcal{P}$ to the affine flag variety $\mathcal{G}/\mathcal{B}$ and to generalize from type A to all Weyl types. Together with Luc Lapointe, Jennifer Morse and Mark Shimozono, we have been developing an affine version of Schensted insertion and an affine Pieri rule [**12**].

2. Equivariant homology and cohomology of $\mathcal{G}/\mathcal{P}$

Let $\mathcal{G}$ be the affine Kac-Moody Group of type $\hat{A}_{n-1}$ over $\mathbb{C}$ and let T be a Cartan subgroup of $\mathcal{G}$. Let $\mathcal{B}$ be a Borel subgroup of $\mathcal{G}$. Let $\mathcal{P}$ be a parabolic subgroup of $\mathcal{G}$. The homogeneous space $\mathcal{G}/\mathcal{P}$ is not a finite dimensional variety but an ind-variety (see [**9**]). The group $\mathcal{G}$ possesses a *Bruhat decomposition* $\mathcal{G} = \bigcup_{w \in W} \mathcal{B}w\mathcal{B}$ where W denotes the *affine symmetric group*. The Bruhat decomposition induces a decomposition of $\mathcal{G}/\mathcal{P}$ into *Schubert cells*:

$$\mathcal{G}/\mathcal{P} = \bigcup_{w \in W^P} X_w$$

where P is the parabolic subgroup of W associated to $\mathcal{P}$ and W^P denotes the elements of shortest length in W/P (see [**7**]). The Schubert classes $\sigma_w = [X_w]$ representing X_w in $H_*(\mathcal{G}/\mathcal{P})$ form a basis of the homology. We will denote the Schubert classes in homology, cohomology, equivariant homology and equivariant cohomology as follows

$$\sigma_w \in H_*(\mathcal{G}/\mathcal{P}) \ , \ \sigma^w \in H^*(\mathcal{G}/\mathcal{P}) \ , \ \sigma_{(w)} \in H_*^T(\mathcal{G}/\mathcal{P}) \ , \ \sigma^{(w)} \in H^*_T(\mathcal{G}/\mathcal{P}).$$

Throughout this paper, all homology and cohomology rings will be with $\mathbb{Z}$-coefficients.

From now on we shall assume that $\mathcal{P}$ is a maximal parabolic subgroup. The corresponding parabolic subgroup $W_0 \subset W$ is the usual symmetric group S_n and we denote the minimal-length representatives of W/W_0 by W^0. We call the elements of W^0 *Grassmannian elements*. The homogeneous space $\mathcal{G}/\mathcal{B}$ is known as the *affine flag variety* and $\mathcal{G}/\mathcal{P}$ is known as the *affine Grassmannian*. The isomorphism type of $\mathcal{G}/\mathcal{P}$ does not depend on the choice of maximal parabolic $\mathcal{P}$. It is in fact homeomorphic to $GL_n(\mathcal{K})/GL_n(\mathcal{O})$ where $\mathcal{K} = \mathbb{C}((t))$ denotes the field of Laurent series and $\mathcal{O} = \mathbb{C}[[t]]$ denotes the subring of power series.

A special feature of $\mathcal{G}/\mathcal{P}$ is that it is a group as follows. Let $K = U_n \subset GL_n$ be the compact group of type A_{n-1}. Then it is well known that $\mathcal{G}/\mathcal{P}$ is homotopy equivalent to (the identity component of) ΩK, the space of based loops into K. The group structure of ΩK induces a multiplication on (equivariant) homology, so that $H_*(\mathcal{G}/\mathcal{P})$ and $H^*(\mathcal{G}/\mathcal{P})$ are dual Hopf-algebras. Thus one can sensibly ask for *homology Schubert polynomials* representing the Schubert classes $\sigma_w \in H_*(\mathcal{G}/\mathcal{P})$. This is a feature not present in classical Schubert calculus.

The homology and cohomology rings (and their Hopf-algebra structures) of ΩK were earlier computed by Bott.

Theorem 2.1 ([**3**]). *We have the isomorphisms*

$$H_*(\mathcal{G}/\mathcal{P}) = \mathbb{Z}[\sigma_1, \sigma_2, \ldots, \sigma_{n-1}]$$

and

$$H^*(\mathcal{G}/\mathcal{P}) = \mathcal{S}H^*(\mathbb{CP}^{n-1})$$

where $\mathcal{S}$ denotes an infinite symmetric power.

These rings can be identified respectively with a subring and a quotient ring of the ring of symmetric functions. The aim of this paper is thus to identify the Schubert classes $\sigma_w \in H_*(\mathcal{G}/\mathcal{P})$ and $\sigma^w \in H^*(\mathcal{G}/\mathcal{P})$ as explicit symmetric functions.

3. NilHecke Ring

Let $\{r_i \mid i \in \mathbb{Z}/n\mathbb{Z}\}$ denote the simple generators of W and let $\{\alpha_i \mid i \in \mathbb{Z}/n\mathbb{Z}\}$ denote the simple roots of the root system of type $\hat{A}_{n-1}$ and for a real root α we let $\alpha^\vee$ denote the corresponding coroot. For each root α, we denote the corresponding reflection by r_α. Let $h^*_{\mathbb{Z}}$ denote the $\mathbb{Z}$-span of the fundamental weights, and let $S = \mathrm{Sym}(h^*_{\mathbb{Z}})$ denote the ring of polynomials in the weights so that $S = H^*_T(\text{point})$.

Let $\mathbb{A}$ denote the *affine nilHecke ring* of type $\hat{A}_{n-1}$ (see [**8**]). (Note that Kostant and Kumar define $\mathbb{A}$ over the rationals, but we have found it more convenient, following Peterson [**19**], to work over $\mathbb{Z}$.) It is the ring with a 1 given by generators $\{A_i \mid i \in \mathbb{Z}/n\mathbb{Z}\} \cup \{\lambda \mid \lambda \in h^*_{\mathbb{Z}}\}$ and the relations

$$\begin{aligned} A_i\,\lambda &= (r_i \cdot \lambda)\,A_i + \langle \lambda, \alpha_i^\vee \rangle \cdot 1 && \text{for } \lambda \in h^*_{\mathbb{Z}} \\ A_i\,A_i &= 0 && \\ A_i\,A_j &= A_j\,A_i && \text{if } |i-j| \geq 2 \\ A_i\,A_{i+1}\,A_i &= A_i\,A_{i+1}\,A_i. && \end{aligned}$$

The ring $\mathbb{A}$ acts as generalized BGG-Demazure operators on $H^*_T(X)$ for any LK-space X (here LK is the space of all loops into the unitary group U_n). The element A_i corresponds to the map $H^*_T(\mathcal{G}/\mathcal{B}) \to H^{*-2}_T(\mathcal{G}/\mathcal{B})$ obtained by integration along the fibers of the $\mathbb{P}^1$-fibration $\mathcal{G}/\mathcal{B} \to \mathcal{G}/\mathcal{P}_i$ where $\mathcal{P}_i$ are the minimal parabolic subgroups. In fact Peterson [**19**] has shown that $\mathbb{A}$ is exactly the ring of *"compact characteristic operators"*; see also [**9**]. Combinatorially, in the classical case the elements A_i act as *divided difference operators* on the Schubert polynomials.

Let $w \in W$ and let $w = s_{i_1} \cdots s_{i_l}$ be a reduced decomposition of w. Then $A_w := A_{i_1} \cdots A_{i_l}$ is a well defined element of $\mathbb{A}$. We let $A_0 := 1$. By [**8**] or [**19**, Proposition 2-7], $\{A_w \mid w \in W\}$ is an S-basis of $\mathbb{A}$. We will also identify r_i with the element $1 - \alpha_i A_i \in \mathbb{A}$ and abusing notation, we write $w \in \mathbb{A}$ for the element in the nilHecke ring corresponding to $w \in W$.

Let $\mathbb{A}_0 \subset \mathbb{A}$ denote the subring over $\mathbb{Z}$ of $\mathbb{A}$ generated by the A_i only. I called this the *affine nilCoxeter algebra* in [**10**]. There is a specialization map $\phi_0 : \mathbb{A} \to \mathbb{A}_0$ given by

$$\phi_0 : \sum_w a_w A_w \longmapsto \sum_w \phi_0(a_w) A_w$$

where ϕ_0 evaluates a polynomial $s \in S$ by setting all α_i to 0.

For later use, we note the following straightforward result, whose proof we omit; see [**8**, Proposition 4.30].

Lemma 3.1. *Let $w \in W$ and $\lambda \in S$ be of degree 1. Then*

$$A_w \lambda = (w \cdot \lambda) A_w + \sum_{r_\alpha w \lessdot w} \langle \lambda, \alpha^\vee \rangle\, A_{r_\alpha w}.$$

Here $\lessdot$ denotes a cover in strong Bruhat order.

The coefficients $\langle \lambda, \alpha^\vee \rangle$ are known as *Chevalley coefficients.*

4. The coproduct on $\mathbb{A}$

Define the coproduct map $\Delta : \mathbb{A} \to \mathbb{A} \otimes_S \mathbb{A}$ by

$$\begin{aligned}\Delta(s) &= 1 \otimes s = s \otimes 1 && \text{for } s \in S\\ \Delta(A_i) &= A_i \otimes 1 + r_i \otimes A_i = 1 \otimes A_i + A_i \otimes r_i\\ &= A_i \otimes 1 + 1 \otimes A_i - A_i \otimes \alpha_i A_i.\end{aligned}$$

This is a well defined map, which in addition is cocommutative. One can deduce from these relations that $\Delta(w) = w \otimes w$. (In the original work of [**8**], this last relation was used to define Δ, but we shall follow the set up of [**19**]).

One should be careful since the tensor product $\mathbb{A} \otimes_S \mathbb{A}$ is *not* a ring. For example,

$$(A_i \otimes 1).(1 \otimes \alpha_i) \neq (A_i \otimes 1).(\alpha_i \otimes 1)$$

However, it is shown in [**19**] that the action of $\mathbb{A}$ on $\mathbb{A} \otimes_S \mathbb{A}$ given by the above formulae still give a well defined action of $\mathbb{A}$ on $\mathbb{A} \otimes_S \mathbb{A}$. That is, $\Delta(a) = a \cdot (1 \otimes 1)$ for any $a \in \mathbb{A}$.

Note that ϕ_0 also sends $\mathbb{A} \otimes_S \mathbb{A}$ to $\mathbb{A}_0 \otimes_{\mathbb{Z}} \mathbb{A}_0$ by evaluating the coefficients at 0 when writing in the basis $\{A_w \otimes A_v\}_{w,v \in W}$.

Theorem 4.1 ([**8**]). *Let*

$$\Delta(A_w) = \sum_{u,v \in W} a_w^{u,v} A_u \otimes A_v.$$

Then $a_w^{u,v}$ are the (Schubert) structure constants of $H_T^*(\mathcal{G}/\mathcal{B})$, so that

$$\sigma^{(u)} \cdot \sigma^{(v)} = \sum_{w \in W} a_w^{u,v} \sigma^{(w)}.$$

Theorem 4.1 is in fact valid for all symmetrizable Kac-Moody groups. Since the product of two Grassmannian classes $\sigma^{(u)}$ and $\sigma^{(v)}$ (where $u, v \in W^0$) in $H^T(\mathcal{G}/\mathcal{P})$ is Grassmannian, we have the following simple result.

Lemma 4.2. *If $w \notin W^0$ and $u, v \in W^0$ then $a_w^{u,v} = 0$.*

5. Symmetric functions

We refer to [**18**] for details concerning the material of this section. Let $\Lambda = \Lambda_{\mathbb{Z}}$ denote the ring of symmetric functions over $\mathbb{Z}$ in infinitely many variables $x_1, x_2, \ldots$. We write $h_i(x)$ for the *homogeneous symmetric functions* and for a partition $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots)$, we write $h_\lambda(x) = h_{\lambda_1(x)} h_{\lambda_2(x)} \cdots$. The elements $h_1(x), h_2(x), \ldots \in \Lambda$ form a set of algebraically independent set of generators of Λ. We let $m_\lambda(x) \in \Lambda$ denote the *monomial symmetric functions*. They form a basis of the ring of symmetric functions over the integers.

Let $\Lambda_n \subset \Lambda$ denote the subring of the symmetric functions generated by $h_i(x)$ for $i \in [0, n-1]$. Let Λ^n denote the quotient of Λ given by $\Lambda^n = \Lambda / \langle m_\lambda(x) \mid \lambda_1 \geq n \rangle$. Clearly the set $\{m_\lambda(x) \mid \lambda_1 < n\}$ forms a basis of Λ^n. When giving an element $\bar{f} \in \Lambda^n$ we will usually just give a representative $f \in \Lambda$ without further comment.

The *Hall inner product*, denoted $\langle ., . \rangle : \Lambda \times \Lambda \to \mathbb{Z}$, is a symmetric non-degenerate pairing defined by $\langle h_\lambda(x), m_\mu(x) \rangle = \delta_{\lambda\mu}$. It induces a non-degenerate pairing $\langle ., . \rangle : \Lambda_n \times \Lambda^n \to \mathbb{Z}$.

It is not too difficult to see from Theorem 2.1 that $\Lambda_n \cong H_*(\mathcal{G}/\mathcal{P})$ and $\Lambda^n \cong H^*(\mathcal{G}/\mathcal{P})$.

In fact the ring of symmetric functions Λ is a Hopf algebra with coproduct given by $\Delta(h_i(x)) = \sum_{j \leq i} h_j(x) \otimes h_i(x)$. This Hopf-algebra structure gives Λ_n and Λ^n the structures of dual Hopf algebras.

6. Affine Schur functions and k-Schur functions

An integral orthonormal basis of Λ is given by the set of Schur functions $s_\lambda(x)$. We will be concerned with a set of dual bases $\{s_\lambda^{(k)}(x)\}$ of Λ_n and $\{F_\lambda(x)\}$ of Λ^n called respectively the *k-Schur functions*, and *affine Schur functions* or *dual k-Schur functions*. The k-Schur functions $\{s_\lambda^{(k)}(x)\}$ were introduced in [**16**], and were further studied in [**13, 14**]. We will give a quick "dual" definition of these functions.

DEFINITION 6.1. Let $a = a_1 a_2 \cdots a_k$ be a word with letters from $\mathbb{Z}/n\mathbb{Z}$ so that $a_i \neq a_j$ for $i \neq j$. Let $A = \{a_1, a_2, \ldots, a_k\} \subset [0, n-1]$. The word a is *cyclically decreasing* if for every i such that $i, i+1 \in A$, the letter $i+1$ precedes i in a. A permutation w is *cyclically decreasing* if $w = s_{a_1} \cdots s_{a_k}$ for some cyclically decreasing sequence $a_1 a_2 \cdots a_k$.

Now define, following [**10**], the elements $h_i \in \mathbb{A}_0 \subset \mathbb{A} : i \in [0, n-1]$ by the formula

$$h_i = \sum_w A_w$$

where the sum is over *cyclically decreasing* permutations w with length $l(w) = i$. If $I \subset [0, n-1]$ and w be the corresponding cyclically decreasing permutation. Then we will write A_I for A_w.

Let $\mathbb{B}$ denote the subalgebra of $\mathbb{A}_0 \subset \mathbb{A}$ generated by the h_i for $i \in [0, n-1]$, which we call the *affine Fomin-Stanley subalgebra*.

THEOREM 6.2 ([**10**]). *The algebra $\mathbb{B}$ is commutative. It is isomorphic to the subalgebra Λ_n of the symmetric functions generated by the homogeneous symmetric functions $h_i(x)$ for $i \in [0, n-1]$, under the map $\psi : h_i(x) \mapsto h_i$.*

Let $\langle ., . \rangle : \mathbb{A}_0 \times \mathbb{A}_0 \to \mathbb{Z}$ denote the bilinear pairing defined by $\langle A_w, A_v \rangle = \delta_{wv}$.

DEFINITION 6.3 ([**10**]). Let $w \in W$. Define the *affine Stanley symmetric functions* $\tilde{F}_w(x) \in \Lambda$ by

$$\tilde{F}_w(x) = \sum_{a=(a_1,a_2,\ldots,a_t)} \left\langle h_{a_t} h_{a_{t-1}} \cdots h_{a_1} \cdot 1, A_w \right\rangle x_1^{a_1} x_2^{a_2} \cdots x_t^{a_t},$$

where the sum is over compositions of $l(w)$ satisfying $a_i \in [0, n-1]$.

The (image in Λ^n of the) set $\{\tilde{F}_w(x) \mid w \in W^0\}$ forms a basis of Λ^n (see [**10**]). We called these functions *affine Schur functions* in [**10**]. They were earlier introduced in a different manner in [**15**], where they were called *dual k-Schur functions*. The *k-Schur functions* $\{s_w^{(k)}(x) \mid w \in W^0\}$ are the dual basis of Λ_n to the affine Schur functions under the Hall inner product. There is a bijection $w \leftrightarrow \lambda(w)$ from Grassmannian permutations $\{w \in W^0\}$ to partitions $\{\lambda \mid \lambda_1 < n\}$ obtained by taking the *code* of the permutation; see [**2**]. We make the identifications $\tilde{F}_w(x) = \tilde{F}_{\lambda(w)}(x)$ and $s_w^{(k)}(x) = s_{\lambda(w)}^{(k)}(x)$ under this bijection. Note that in the terminology of [**16**], $k = n-1$.

7. Non-commutative k-Schurs

Recall that we have an isomorphism $\psi : \Lambda_n \to \mathbb{B}$. Define $\Delta_{\mathbb{B}} : \mathbb{B} \to \mathbb{B} \otimes_{\mathbb{Z}} \mathbb{B}$ by

$$\Delta_{\mathbb{B}}(h_i) = \sum_{j \leq i} h_j \otimes h_{i-j}$$

and extending $\Delta_{\mathbb{B}}$ to a ring homomorphism. This is just the natural coproduct of the symmetric functions as explained in Section 5. The following definition is inspired by work of Fomin and Greene [**5**].

DEFINITION 7.1. Let $w \in W^0$. The *non-commutative k-Schur functions* are given by

$$s_w^{(k)} := \psi(s_w^{(k)}(x)) \in \mathbb{B}.$$

The main result we need concerning the non-commutative k-Schur functions is the following.

THEOREM 7.2 ([**10**]). *The non-commutative k-Schurs can be written in the A_w basis as*

$$s_w^{(k)} = A_w + \sum_{v \notin W^0} b_{w,v} A_v$$

where w is a Grassmannian permutation and the second term is a summation over non-Grassmannian permutations.

8. The Main Theorem

Our main theorem is the following.

Theorem 8.1. *The map $\theta : H_*(\mathcal{G}/\mathcal{P}) \to \Lambda_n$ given by*

$$\theta : \sigma_w \longmapsto s_w^{(k)}(x)$$

is an isomorphism of Hopf-algebras. The map $\theta' : H^*(\mathcal{G}/\mathcal{P}) \to \Lambda^n$ given by

$$\theta' : \sigma^w \longmapsto \tilde{F}_w(x)$$

is an isomorphism of Hopf-algebras.

In the homology case, this theorem was a conjecture of Mark Shimozono. The conjecture in the cohomology case was made precise by Jennifer Morse.

We shall prove the following technical result in Section 13.

Theorem 8.2. *The two coproducts $\Delta, \Delta_{\mathbb{B}}$ agree on $\mathbb{B}$ up to specialisation at 0:*

$$\phi_0 \circ \Delta = \Delta_{\mathbb{B}}.$$

The following theorem proves half of Theorem 8.1. Recall that $a_w^{u,v}$ are the multiplicative structure constants of $H^*(\mathcal{G}/\mathcal{P})$.

Theorem 8.3. *We have*

$$\phi_0(\Delta(s_w^{(k)})) = \sum_{u,v \in W^0\,:\, l(u)+l(v)=l(w)} a_w^{u,v}\ s_u^{(k)} \otimes s_v^{(k)}.$$

Note that since the k-Schur functions $s_w^{(k)}(x)$ are Hall-dual to the affine Schur functions $\tilde{F}_w(x)$, Theorem 8.3 immediately implies that multiplication of $\tilde{F}_w(x)$ in Λ^n agrees with the multiplication of σ^w in $H^*(\mathcal{G}/\mathcal{P})$. See also the discussion in [**10**].

Proof. By Theorems 4.1 and 7.2, we have

$$\begin{aligned}\Delta(s_w^{(k)}) &= \Delta(A_w + \sum_v b_{w,v} A_v) \\ &= \sum_{u,v} a_w^{u,x} A_u \otimes A_x + \sum_v b_{w,v} \sum_{y,z} a_v^{y,z} A_y \otimes A_z\end{aligned}$$

The polynomials $a_w^{u,x}$ are known to have (homogeneous) degree $l(u) + l(x) - l(w)$, so we get

$$\phi_0(\Delta(s_w^{(k)})) = \sum_{\substack{u,x \\ l(u)+l(x)=l(w)}} a_w^{u,x} A_u \otimes A_x + \sum_v b_{w,v} \sum_{\substack{y,z \\ l(y)+l(z)=l(v)}} a_v^{y,z} A_y \otimes A_z.$$

By Lemma 4.2, we may actually write

$$\phi_0(\Delta(s_w^{(k)})) = \sum_{u,v \in W^0 : l(u)+l(v)=l(w)} a_w^{u,x} A_u \otimes A_x + \text{other terms}. \tag{8.1}$$

The other terms involve $A_y \otimes A_z$ where one of y or z is not Grassmannian.

Now by Theorem 8.2, we have $\phi_0(\Delta(s_w^{(k)})) \in \mathbb{B} \otimes_{\mathbb{Z}} \mathbb{B}$ so we may write it as

$$\phi_0(\Delta(s_w^{(k)})) = \sum_{u,x \in W^0} c_w^{u,x} s_u^{(k)} \otimes s_x^{(k)}$$

where $c_w^{u,x}$ are some integers. Using Theorem 7.2 again, we have

$$\phi_0(\Delta(s_w^{(k)})) = \sum_{u,x \in W^0} c_w^{u,x} A_u \otimes A_x + \text{other terms},$$

where as before the other terms involve the basis elements $A_y \otimes A_z$ where one of y or z is not Grassmannian. Comparing with (8.1) we have $c_w^{u,x} = a_w^{u,x}$, as required. ∎

9. $\mathbb{B}$ nearly annihilates S

To prove Theorem 8.2, and also to obtain the multiplicative constants of the homology $H_*(\mathcal{G}/\mathcal{P})$ we first prove a technical property of the Fomin-Stanley subalgebra $\mathbb{B}$.

THEOREM 9.1. *Let $b \in \mathbb{B}$ and $s \in S$. Then*

$$\phi_0(bs) = \phi_0(s)b.$$

PROOF. We show that $\phi_0(h_i \cdot \alpha_j) = 0$ for each i and the theorem follows since h_i generate $\mathbb{B}$. Without loss of generality we will assume that $j = 1$. Let $I \subset \mathbb{Z}/n\mathbb{Z}$ be of size i. We calculate $\phi_0(A_I\alpha_1)$ explicitly. In the following $[2, r]$ is the largest interval of its form (possibly empty) contained in I which contains 2. It is possible that $[2, r]$ contains 0 but it cannot contain 1 (since then it will have size n). Also the subset I' never contains any of $0, 1, 2$. The sums over a are always over $a \in [2, r]$. The (A),(B),(C) are for marking the terms only, for later use.

I	$\phi_0(A_I\alpha_1)$
$I' \cup [2, r]$	$-\sum_a A_{I-\{a\}}$(A)
$I' \cup [2, r] \cup \{1\}$	$2A_{I-\{1\}}$(A) $+ \sum_a A_{I-\{a\}}$(C)
$I' \cup [2, r] \cup \{0\}$	$-A_{I-\{0\}}$(A) $- \sum_a A_{I-\{a\}}$(B)
$I' \cup [2, r] \cup \{0, 1\}$	$-A_{I-\{0\}}$(C) $+ A_{I-\{1\}}$(B)

For example

$$\begin{aligned}
&A_{[2,r]}A_1A_0\alpha_1 \\
&\qquad = A_{[2,r]}A_1((\alpha_1 + \alpha_0)A_0 - 1) \\
&\qquad = -A_{[2,r]}A_1 + A_{[2,r]}(-\alpha_1A_1A_0 + 2A_0 + (\alpha_1 + \alpha_0)A_1A_0 - A_0) \\
&\qquad = -A_{[2,r]-\{0\}} + A_{[2,r]-\{1\}} + \alpha_0A_{[2,r]}A_1A_0.
\end{aligned}$$

The A_t factors for $t \in I'$ always commute in these calculations.

One observes that the terms marked (A) or (B) or (C) when grouped together cancel out. We have: (A) corresponds to subsets J of size $i - 1$ such that J contains neither 1 nor 0; and (B) corresponds to subsets J of size $i - 1$ such that J contains 0 but not 1; and (C) corresponds to subsets J of size $i - 1$ such that J contains 1 but not 0. Every such subset in say case (A) will appear in all 3 case (A) terms. No other subsets (those containing both 0 and 1) appear in the sum $\sum_I A_I\alpha_1$.

For example, the subset $J = [2, 4] \cup [5, 7]$ will appear in $\phi_0(A_I\alpha_1)$ for $I = [2, 7]$ or $[1, 4] \cup [5, 7]$ or $\{0\} \cup [2, 4] \cup [5, 7]$. The multiplicities will be -1, 2, and -1 respectively, which cancel out.

□

10. An identity for finite Weyl groups

Let W^{fin} be a finite Weyl group and $H^*(K/T)$ be the cohomology of the corresponding flag variety. Also let w° denote the longest element of W^{fin}.

PROPOSITION 10.1. *Suppose that for some coefficients $\{b_u \in \mathbb{Z}\}_{u \in W^{\mathrm{fin}}}$ the following identity holds in $\mathbb{Z}W^{\mathrm{fin}}$ for all integral weights $\lambda \in h^*_{\mathbb{Z}}$*

$$\sum_{u \in W^{\mathrm{fin}};\ l(u)>0} b_u \sum_{ur_\alpha \lessdot u} \langle \lambda, \alpha^\vee \rangle\, ur_\alpha = 0.$$

Then $b_u = 0$ for all u.

PROOF. First apply the transformation $u \mapsto w^\circ u$ to the identity of the Proposition. Then reindexing the b_u, we obtain

$$\sum_{u \in W^{\mathrm{fin}};\ u \neq w^\circ} b_u \sum_{ur_\alpha \gtrdot u} \langle \lambda, \alpha^\vee \rangle\, ur_\alpha = 0$$

for all λ.

Let $\sigma_u^{(0)} \in H^*(K/T)$ denote the Schubert classes in the finite flag variety. By the Chevalley-Monk formula [**1**] we have

$$[\lambda] \cdot \sigma_u^{(0)} = \sum_{ur_\alpha \gtrdot u} \langle \lambda, \alpha^\vee \rangle\, \sigma_{ur_\alpha}^{(0)}$$

where $[\lambda] \in H^*(K/T)$ denotes the image of λ under the characteristic homomorphism $S(h_{\mathbb{Z}}^*) \to H^*(K/T)$. For example, if $\lambda = \omega_i$ is a fundamental weight then $[\omega_i] = \sigma_{s_i}^{(0)}$. It is well known that $\sigma_{s_1}^{(0)}, \sigma_{s_2}^{(0)}, \ldots, \sigma_{s_{n-1}}^{(0)}$ generate $H^*(K/T)$ or alternatively that the characteristic homomorphism is surjective.

Suppose that $[\lambda] \cdot \sigma = 0$ for some $\sigma \in H^*(K/T)$ and all $\lambda \in h_{\mathbb{Z}}^*$. If $l(v) + l(u) = l(w_\circ)$ we have $\sigma_v^{(0)} \cdot \sigma_u^{(0)} = \delta_{v,w_\circ u}\sigma_{w_\circ}^{(0)}$. Since $\sigma_u^{(0)} \cdot \sigma = 0$ for all $u \neq \mathrm{id}$, we find that σ must be a multiple of the class $\sigma_{w_\circ}^{(0)}$. Letting $\sigma = \sum_u b_u \sigma_u^{(0)}$ and applying the Chevalley-Monk formula we obtain the proposition. $\square$

11. The subalgebra $\mathbb{B}'$

Define a subalgebra $\mathbb{B}' \subset \mathbb{A}_0$ as follows:

$$\mathbb{B}' = \{a \in \mathbb{A}_0 \mid \phi_0(as) = \phi_0(s)a \text{ for all } s \in S\}.$$

Thus Theorem 9.1 says that $\mathbb{B} \subset \mathbb{B}'$. It turns out that $\mathbb{B}'$ is always a commutative subalgebra for all affine types, though we will not need such generality here.

Proposition 11.1. *Let $b \neq 0 \in \mathbb{B}'$ and write $b = \sum_w b_w A_w$ with $b_w \in \mathbb{Z}$. Then $b_w \neq 0$ for some $w \in W^0$.*

Proof. Let $D = \{w \in W \mid b_w \neq 0\}$. For each $w \in W$ we may uniquely write $w = x_w y_w$ where $x_w \in W^0$ and $y_w \in W_0$. Let $d = \{\min(l(y_w)) \mid w \in D\}$. We write $l_0(w) := l(y_w)$.

Suppose $d \neq 0$ and let $w \in D$ minimize $l_0(w)$. Let $\lambda \in S$ be of degree 1. Then by Lemma 3.1, $\phi_0(A_w\lambda) = \sum_{wr_\alpha \lessdot w} \langle \lambda, \alpha^\vee \rangle A_{wr_\alpha}$. We know that $w \gtrdot v$ if and only if a reduced decomposition of v is obtained from a reduced decomposition of w by removing a simple generator. Since $w = x_w y_w$, each such v satisfies $l_0(v) \geq l_0(w) - 1$. Let $D_w = \{v \lessdot w \mid l_0(v) = l_0(w) - 1\}$. Then $v \in D_w$ if and only if $v = x_v y_v$ where $x_v = x_w$ and $y_v \lessdot y_w$.

Now write $\phi_0(b\lambda) = \sum_v b'_v A_v$ and focus only on the coefficients of b'_v satisfying $l_0(v) = d-1$ and $v = xy_v$ for some fixed $x \in W^0$. If $b \in \mathbb{B}'$ then $b'_v = 0$. Thus in particular, for every $\lambda \in S$ of degree 1, we have

$$\sum_{u \in W_0} b_{xu} \sum_{ur_\alpha \lessdot u} \langle \lambda, \alpha^\vee \rangle A_x A_{ur_\alpha} = 0.$$

Factorizing A_x to the front, we see that this is impossible by Proposition 10.1. Since this is true for all $x \in W^0$ we conclude that we must have $d = 0$. $\square$

12. Peterson's j-homomorphism

To further understand the non-commutative k-Schur functions, we require a result of Peterson. Let $Z_{\mathbb{A}}(S)$ denote the centralizer of S in $\mathbb{A}$.

Theorem 12.1 ([**19**]). *There is an isomorphism $j : H_*^T(\Omega K) \to Z_{\mathbb{A}}(S)$ such that*

$$j(\sigma_{(x)}) = A_x \mod I$$

where x is a Grassmannian permutation and

$$I = \sum_{w \in W_0\,;\, w \neq \mathrm{id}} \mathbb{A} \cdot A_w.$$

Recall that $W_0 = S_n$ is the usual symmetric group.

Theorem 12.2. *We have $\phi_0(Z_{\mathbb{A}}(S)) = \mathbb{B}'$. More precisely, $\{\phi_0(j(\sigma_{(u)})) \mid u \in W^0\}$ forms a basis of $\mathbb{B}'$ over $\mathbb{Z}$.*

Proof. The fact that $\phi_0(Z_A(S)) \subset \mathbb{B}'$ is a trivial calculation. Now let $b \in \mathbb{B}'$. By Proposition 11.1 it contains a Grassmannian term A_u with non-zero coefficient b_u. By Theorem 12.1, $b - b_u\phi_0(j(\sigma_{(u)}))$ has strictly fewer Grassmannian terms and also lies in $\mathbb{B}'$. Repeating, we see that one can write b uniquely as a $\mathbb{Z}$-linear combination of the elements $\phi_0(j(\sigma_{(u)}))$. $\square$

Corollary 12.3. *The two algebras $\mathbb{B}$ and $\mathbb{B}'$ are identical (as subalgebras of $\mathbb{A}_0$) and we have*

$$\phi_0(j(\sigma_{(u)})) = s_u^{(k)}.$$

PROOF. This follows immediately from Theorems 12.2 and 7.2 together with Proposition 11.1: both $\phi_0(j(\sigma_{(u)}))$ and $s_u^{(k)}$ lie in $\mathbb{B}'$ and have a unique Grassmannian term A_u. □

Finally, we can complete the proof of our main theorem.

PROOF OF THEOREM 8.1. Let $x, y \in W^0$. If $\sigma_{(x)}\sigma_{(y)} = \sum_{z\in W^0} c_{x,y}^z \sigma_{(z)}$ in $H_T(\mathcal{G}/\mathcal{P})$ then $\sigma_x\sigma_y = \sum_{z\in W^0} \phi_0(c_{x,y}^z)\sigma_z$ in $H_*(\mathcal{G}/\mathcal{P})$. Thus $\mathbb{B} = \mathbb{B}'$ is isomorphic to $H_*(\mathcal{G}/\mathcal{P})$ and we have

$$s_x^{(k)}\, s_y^{(k)} = \sum_{z\in W^0} \phi_0(c_{x,y}^z) s_z^{(k)}.$$

This, together with Theorem 8.3 shows that θ and θ' are both algebra and co-algebra homomorphisms. The agreement of the remainder of the Hopf algebra structures is straightforward to verify. □

13. Proof of Theorem 8.2

We now return to the proof of Theorem 8.2. It will follow quickly from the following computation.

PROPOSITION 13.1. *We have*

$$\phi_0(\Delta(h_i)) = \sum_j h_j \otimes h_{i-j}$$

PROOF. Let $\beta_i = -\alpha_i$ be the negative simple roots. We use $\Delta(A_i) = A_i \otimes 1 + 1 \otimes A_i + A_i \otimes \beta_i A_i$.

Let $i_1, i_2, \ldots, i_l$ be a cyclically decreasing sequence. Thus

$$\Delta(A_{i_1}A_{i_2}\cdots A_{i_l}) = \textstyle\prod_j \Delta(A_{i_j})$$
$$= (A_{i_1} \otimes 1 + 1 \otimes A_{i_1} + A_{i_1} \otimes \beta_{i_1}A_{i_1}) \cdots (A_{i_l} \otimes 1 + 1 \otimes A_{i_l} + A_{i_l} \otimes \beta_{i_l}A_{i_l})$$

Let us expand the product, by picking one of the three terms in each parentheses. (Strictly speaking we cannot multiply within $\mathbb{A} \otimes_S \mathbb{A}$, instead we are calculating the action of $\mathbb{A}$ on $\mathbb{A} \otimes_S \mathbb{A}$ via the coproduct: $\Delta(A_i) \cdot (\Delta(A_j) \cdot (1 \otimes 1)) = \Delta(A_iA_j)$).

Because of the cyclically decreasing assumption, the only times we encounter a factor looking like $A_{i_a}\beta_{i_b}$ (where $a < b$) we have either

$$A_{i_a}\beta_{i_b} = \beta_{i_b}A_{i_a} \tag{13.1}$$

or we will have $a = b-1$ and $i_{a+1} = i_a - 1$ and

$$A_{i_a}\beta_{i_a-1} = (\beta_{i_a-1} + \beta_{i_a})A_{i_a} + 1. \tag{13.2}$$

If (13.1) ever occurs, then β_{i_b} commutes with all A_{i_c} where $c < b$ and we may ignore the term since eventually we will apply ϕ_0. Similarly, if (13.2) occurs, the contribution of the term involving β_{i_a-1} is 0 after applying ϕ_0.

Also we perform the calculation

$$A_{i+1}(\beta_i)^m = \beta_{i+1}^m A_{i+1} + \beta_{i+1}^{m-1} + \text{other terms}, \tag{13.3}$$

where the other terms involve β_i on the left somewhere (and would be killed by ϕ_0 later).

Let B and C be two subsets of $[0, n-1]$ with total size equal to $k \le n-1$. We will first describe how to obtain the term $A_B \otimes A_C$ (which occurs in $h_{|B|} \otimes h_{|C|}$) from $\Delta(h_k)$. Define a sequence of integers ("current degree") $(\mathrm{cd}(i) : i \in \mathbb{Z}/n\mathbb{Z})$ by $\mathrm{cd}(i) = \max_t\{|I \cap [i-t, i]| + |J \cap [i-t, i]| - t - 1\}$. Since $|B| + |C| < n$ we can find i so that $\mathrm{cd}(i) = 0$ and $i \notin B \cup C$.

We may assume that $i = 0$. Let $B = (b_1 > \cdots > b_g)$ and $C = (c_1 > \cdots > c_h)$. Define a sequence $(t_1, t_2, \ldots, t_{n-1}) \in \{L, R, B, E\}^{n-1}$ as follows (E = empty, L = left, R = right and B = both):

$$t_i = \begin{cases} E & \text{if } \mathrm{cd}(i) = 0 \text{ and } E \notin B \cup C \\ L & \text{if } \mathrm{cd}(i) = 0 \text{ and } E \in B \text{ but } E \notin C \\ R & \text{if } E \notin B \text{ and } (\mathrm{cd}(i) > 0 \text{ or } E \in C) \\ B & \text{otherwise.} \end{cases}$$

Now let $I = \{i \in [1, n-1] \mid t \neq E\} \subset [1, n-1]$. Then $A_B \otimes A_C$ is obtained from $\Delta(A_I)$ by picking the term $A_{i_s} \otimes 1$ if $t_{i_s} = L$, the term $1 \otimes A_{i_s}$ if $t_{i_s} = R$ and $A_{i_s} \otimes \beta_{i_s}A_{i_s}$ if $t_{i_s} = B$.

The sequence of integers $(\mathrm{cd}(i))$ tells us the current degree (in the second factor of the tensor product) in S of the term that we want to pick whenever we encounter the situation of (13.3).

For example if $\mathrm{cd}(t) = 3$ and $\mathrm{cd}(t+1) = 3$ then $t+1 \in B$ or $t+1 \in C$. In the first case we will have $(A_{t+1} \otimes 1) \cdot (a \otimes \beta_i^3 b)$, for some a and b not involving S, and there is no further choice. In the second case we get

$$(1 \otimes A_{t+1}) \cdot (a \otimes \beta_i^3 b) = a \otimes (\beta_{i+1}^3 A_{t+1} + \beta_{i+1}^2)b,$$

modulo terms involving β_i on the right. One must make a further choice between $\beta_{i+1}^3 A_{t+1}$ and β_{i+1}^2. We pick the first term since we want $t+1 \in C$ and this agrees with the degree being $\mathrm{cd}(t+1) = 3$.

Thus every term of the form $A_B \otimes A_C$ appears in the expansion of $\phi_0(\Delta(h_i))$. Conversely, one can reverse the description given above to see that every term in the expansion is indeed of that form. □

PROOF OF THEOREM 8.2. From Proposition 13.1, we have $\Delta_{\mathbb{B}}(h_i) = \phi_0(\Delta(h_i))$. Now let $a \in \mathbb{B}$ and $b \in \mathbb{B}$ and suppose we have shown that $\Delta_{\mathbb{B}}(a) = \phi_0(\Delta(a))$ and $\Delta_{\mathbb{B}}(b) = \phi_0(\Delta(b))$. Let $\Delta(a) = \sum_{w,v} A_w \otimes a_{w,v} A_v$ and $\Delta(b) = \sum_{x,y} A_x \otimes b_{x,y} A_y$, where $a_{w,v}, b_{x,y} \in S$. Then

$$\begin{aligned}
\phi_0(\Delta(ab)) &= \phi_0(\Delta(a)\Delta(b)) \\
&= \phi_0\Big(\sum_{w,v,x,y} A_w A_x \otimes a_{w,v} A_v b_{x,y} A_y\Big) \\
&= \sum_{w,v,x,y} A_w A_x \otimes \phi_0(a_{w,v}) A_v \phi_0(b_{x,y}) A_y \qquad \text{by Theorem 9.1.} \\
&= \phi_0(\Delta(a))\phi_0(\Delta(b)) \\
&= \Delta_{\mathbb{B}}(a)\Delta_{\mathbb{B}}(b) \\
&= \Delta_{\mathbb{B}}(ab).
\end{aligned}$$

Since the h_i generate $\mathbb{B}$ this completes the proof. □

References

[1] I.N. Bernstein, I.M. Gelfand, S.I. Gelfand: Schubert cells and cohomology of the spaces G/P, *Russ. Math. Surv.* **28** (1973), 1-26.

[2] A. Björner, F. Brenti: Affine permutations of type A, *Electron. J. Combin.* **3**/2 (1996), Research Paper 18.

[3] R. Bott: The space of loops on a Lie group, *Michigan Math. J.* **5** (1958), 35-61.

[4] M. Demazure: Désingularisation des variétés de Schubert généralisées., *Ann. Sci. Éc. Norm. Sup.* **7** (1974), 53-88.

[5] S. Fomin and C. Greene: Noncommutative Schur functions and their applications, *Discrete Mathematics* **193** (1998), 179-200.

[6] S. Fomin, R. Stanley: Schubert polynomials and the nilCoxeter algebra, *Adv. Math.* **103** (1994), 196-207.

[7] J. Humphreys: Reflection groups and Coxeter groups, *Cambridge Studies in Advanced Mathematics* **29** Cambridge University Press, Cambridge, 1990.

[8] B. Kostant, S. Kumar: The nil Hecke ring and the cohomology of G/P for a Kac-Moody group G, *Adv. in Math.* **62** (1986), 187-237.

[9] S. Kumar: Kac-Moody groups, their flag varieties and representation theory, *Progress in Mathematics* **204** Birkhuser Boston, Inc., Boston, MA, 2002.

[10] T. Lam: Affine Stanley symmetric functions, preprint, 2005; `arxiv:math.CO/0501335`.

[11] T. Lam: Schubert polynomials for the affine Grassmannian, in preparation, 2005.

[12] T. Lam, L. Lapointe, J. Morse, M. Shimozono: An affine insertion algorithm and affine Pieri rule, in preparation, 2005.

[13] L. Lapointe, J. Morse: Schur function analogs for a filtration of the symmetric function space, *J. Combin. Theory Ser. A* **101**(2) (2003), 191-224.

[14] L. Lapointe, J. Morse: Schur function identities, their t-analogs, and k-Schur irreducibility, *Adv. Math.* **180**(1) (2003), 222-247.

[15] L. Lapointe, J. Morse: Quantum cohomology and the k-Schur basis, preprint, 2005; `math.CO/0501529`.

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

[17] A. Lascoux and M. Schützenberger: Polynômes de Schubert, *C.R. Acad. Sci. Paris*, **294** (1982), 447-450.

[18] I. G. Macdonald: *Symmmetric Functions and Hall Polynomials,* Oxford, 1970.

[19] D. Peterson: Lecture notes at MIT, 1997.

Department of Mathematics, Harvard University, Cambridge MA, 02138, USA
E-mail address: `tfylam@math.harvard.edu`
URL: `http://www.math.harvard.edu/ tfylam`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

An algorithm to describe bijections involving Dyck paths

Yvan Le Borgne

Abstract. We use an algorithm to define bijections involving Dyck paths. This algorithm is parametrized by rewriting rules and is similar to the derivation of a word in a context-free grammar. The bijections are variations of a classical one which is based on the insertion of a peak in the last descent. A systematic study of the algorithms parametrized by a single rewriting rule leads to 6 bijections, taking into account a trivial symmetry. We obtain 6 classical or new parameters on Dyck paths, which are distributed as the length of the last descent. We have a description for 5 of these parameters. We present additional bijections appearing in several combinatorial contexts that can be defined by generalizations of the initial algorithm.

Résumé. On utilise un algorithme pour définir des bijections impliquant les chemins de Dyck. Cet algorithme est paramétré par des règles de réécriture et est proche de la dérivation d'un mot dans une grammaire algébrique. Les bijections sont des variations de la construction classique des chemins de Dyck par l'insertion d'un pic dans la dernière descente. Une étude systématique des algorithmes paramétrés par une seule règle de réécriture permet d'identifier essentiellement 6 bijections. De chacune de ces bijections on déduit un paramètres classique ou nouveau dont la distribution est identique à celle de la longueur de la dernière descente. On donne une description de 5 de ces 6 paramètres. On présente d'autres bijections définissables par des généralisations de cet algorithme et utilisées dans divers contextes combinatoires.

Introduction

The Catalan numbers $(\frac{1}{2n+1}\binom{2n+1}{n})_{n\geq 0} = 1, 1, 2, 5, 14, 42, 129, \ldots$ define a sequence which occurs as the counting sequence of more than one hundred classes of combinatorial objects: ordered trees, binary trees, triangulations of polygons, Dyck paths …(see Exercises 6.19 and 6.25 in [**8**] and its periodic update on the web). This sequence is also the expansion of an algebraic power series $C(t)$ that satisfies the functional equation $C(t) = 1 + tC(t)^2$. This equation is usually reflected on these combinatorial classes as a recursive decomposition of any object into two independent and smaller objects of the same class, if any smaller. Providing such a decomposition as regards a class usually proves that the counting sequence of this class is the Catalan sequence. Another way to fix this counting sequence consists of defining a bijection, which preserves the size of objects, between this class and another one counted by the Catalan sequence (see [**2**] for a more general discussion). Once the counting sequence has been computed, there often remain enumerative and open problems about the class: one wants to take into account not only the size of the objects, but also additional parameters. For example, in the case of two additional parameters, we need to obtain some information on the generating function

$$C(t; u, v) = \sum_{n,i,j\geq 0} c_n(i, j)u^i v^j t^n$$

where $c_n(i, j)$ is the number of objects of size n for which these two parameters equal i and j respectively. The usual decomposition of the objects may not fit well with the additional parameters. In such a case, we have to find either a new decomposition or a bijection that translates these objects and their parameters into other objects with more tractable parameters. In the literature, this kind of problem motivates many bijections between the various combinatorial interpretations of the Catalan sequence.

2000 *Mathematics Subject Classification.* Primary 05A15.
Key words and phrases. Enumeration, Dyck paths, Bijections.

The main aim of this extended abstract is not to solve a particular enumerative problem, but to propose a common framework describing some of these bijections. In an enumerative context we require that these bijections induce one-to-one maps when restricted to objects of any fixed size. Thus, there are 1!1!2!5!14!42! different restrictions of bijections to objects of size less than 5 between two classes counted by the Catalan sequence. But bijections of practical use in combinatorics are not so arbitrary. We want to define a subset $\mathcal{B}$ of these bijections satisfying the following (informal) condition. The set $\mathcal{B}$ should be *expressive*: it contains many bijections already present in the literature and many parameters with classical distributions are preserved or translated. The bijections of $\mathcal{B}$ should admit a *uniform description*, without too many *ad hoc* definitions. We formalize the notion of uniform description for maps between two classes $\mathcal{C}$ and $\mathcal{D}$. A uniform description will be a ordered pair $(\mathcal{P}, F)$ where $\mathcal{P}$ is a set whose elements are called "programs", and F is a map between $\mathcal{P} \times \mathcal{C}$ and $\mathcal{D}$ such that for any $p \in \mathcal{P}$ and $c \in \mathcal{C}$, the size of c is the size of $F(p, c) \in \mathcal{D}$. We denote by $\mathcal{M}$ the set of maps $\{F(p, .) | p \in \mathcal{P}\}$ that is the set of the partial evaluations of F on its first argument. By definition, $\mathcal{B}$ is the set of the one-to-one maps in $\mathcal{M}$. With this uniform description, we could formulate some additional wishes. For any program $p \in \mathcal{P}$, we could check if $F(p, .)$ is a bijection. Given a partial map m between a finite subset $C \subset \mathcal{C}$ and $\mathcal{D}$, we could efficiently compute, if any, at least one (or all) of the programs p compatible on C with m. Given finite subsets C and D of $\mathcal{C}$ and $\mathcal{D}$ respectively and two families of parameters on $\mathcal{C}$ and $\mathcal{D}$ respectively, we could also efficiently compute, if any, the programs in $\mathcal{P}$ that translate the size and the additional parameters when restricted to C and D. This is the end of the dream. Our modest attempt in this extended abstract is a relatively expressive set, as regards the length of the last descent and the area of Dyck paths. The description and the proofs are also relatively uniform. The last two properties (identify a description or guess a bijection preserving parameters) are not even discussed here.

In this extended abstract, we restrict the study to maps between two classes counted by the Catalan sequence: almost decreasing sequences and Dyck paths/words which stand for the classes $\mathcal{C}$ and $\mathcal{D}$ respectively. The latter class is very often used as the image set of a bijection, to prove that a class admits an algebraic recursive decomposition. The former class appears in a classical recursive step-by-step construction of Dyck paths, obtained by inserting a new peak in the last descent. In Section 1 we propose a first uniform description $(\mathcal{P}, F)$ for maps between almost decreasing sequences and Dyck paths. These maps are variations of the step-by-step construction of Dyck paths. The definition of F is an algorithm similar to the derivation of a word in a context-free grammar. This algorithm is parametrized by certain rewriting rules called insertion modes. $\mathcal{P}$ is the set of these insertion modes. In Section 2, we study all maps that are defined by the algorithm parametrized by a single insertion mode. Among the $210 = |\mathcal{P}|$ possibilities, we prove that 32 code bijections. Actually, some of these bijections are identical, and we only obtain $12 = |\mathcal{B}|$ distinct bijections. We deduce from this study classical and new parameters that have the same distribution on Dyck paths as the length of the last descent. Another such a systematic study was made in [**9**], but the approach was to define a set of parameters with the appropriate (Narayana) distribution, then to find bijections, whereas here, we define a set of bijections and then identify the parameters. In Section 3, we present generalizations of the algorithm that allow us to describe relevant bijections in several combinatorial contexts: a description of the Haiman statistic on Dyck paths and a combinatorial interpretation of the calculations involved in the kernel method.

This work summarizes a chapter of the author's PhD thesis [**7**].

1. An insertion algorithm

In this section, we define almost decreasing sequences and Dyck paths. We recall a classical bijection between these objects. Then we introduce some labelings of Dyck paths and some rewriting rules for these labels, which we respectively call Dyck buildings and insertion mode. Finally we present an algorithm parametrized by a single insertion mode that generalizes the classical bijection. This allow us to define the uniform description $(\mathcal{P}, F)$ systematically studied in Section 2.

Let w be a word. By definition, the letter a occurs $|w|_a$ times in w. We denote the empty word by ϵ. A word w over the alphabet $X \equiv \{x, \overline{x}\}$ is a *Dyck word* if $|w|_x = |w|_{\overline{x}}$ and for any prefix u of w, $|u|_x \geq |u|_{\overline{x}}$. The *size* of the Dyck word w is the number $|w|_x$. A *Dyck path* is a walk in the plane, that starts from the origin, is made up of *rises*, *i.e.* steps $(1, 1)$, and *falls*, *i.e.* steps $(1, -1)$, remains above the horizontal axis and finishes on it. The Dyck path related to a Dyck word w is the walk obtained by representing a letter x by a rise, and a letter $\overline{x}$ by a fall, see Figure 1. In the rest of the paper we identify the two notions, denoting

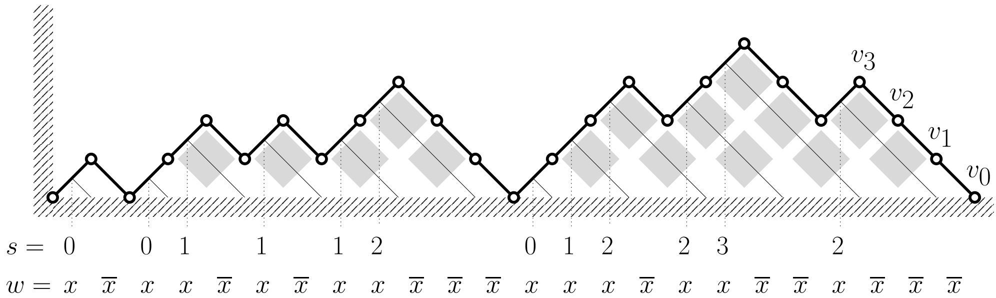


FIGURE 1. A Dyck path with its Dyck word w and its canonical almost decreasing sequence s.

them both w. A *vertex* in a Dyck path w is the origin of the plane or an endpoint of a step in w. In terms of Dyck words, a vertex corresponds to a factorization $w = uv$ where u is the subwalk between the origin and the vertex while v is the remaining subwalk. A *peak* is a vertex preceded by a rise and followed by a fall. A sequence of n non-negative integers $s = (s_k)_{k=1...n}$ is an *almost decreasing sequence* if $s_1 = 0$ and for all $k < n$, $s_{k+1} \leq 1 + s_k$. The empty sequence, denoted $\emptyset$, is an almost decreasing sequence.

The *height* of a rise is the ordinate of its starting vertex. We map a Dyck path to the sequence of the heights of its rises which is an almost decreasing sequence. We call this sequence the canonical almost decreasing sequence of this Dyck path since the map is a classical bijection. We recall a recursive step-by-step definition of the reverse of this map that we illustrate in Figure 1. We assume that we have already fixed that s is mapped to w. We want to define the image w_i of the almost decreasing sequence s, i obtained from s by the appending of the non-negative integer i. In Figure 1, the canonical sequence s ends with the value 2. Therefore, the possible values for i are 3, 2, 1, or 0. These are exactly the values not bigger than the height of the rightmost peak of w. This fact is a property called (P) of the bijection. The Dyck path w_i is obtained from w by an insertion of a factor $x\overline{x}$ in v_i, which is the rightmost vertex of w at height i. This insertion induces that the rightmost peak of w_i is the peak in the inserted factor $x\overline{x}$ thus this peak is at height $i+1$. This corresponds to the property (P) for w_i. This leads to the following step-by-step definition of the image of an almost decreasing sequence s of size n: starting from the empty path ϵ, insert a factor $x\overline{x}$ in the rightmost vertex at height s_k for k running from 1 to n.

To generalize this kind of step-by-step definition of a map, we use some labels in the Dyck words to indicate where a rise and a fall should be inserted in the path mapped to s to obtain the word image of s, i. Consider the (infinite) alphabet $L = \bigcup_{0 \leq k} \{k, \overline{k}\}$. The letters $k \in \mathbb{N}$ will be called *rising labels of index k* while the letters $\overline{k}$ will be called *falling labels of index k*. Let $L(N) = \bigcup_{0 \leq k \leq N} \{k, \overline{k}\}$. Given an alphabet A, the *projection* π_B over the alphabet $B \subseteq A$ is the morphism defined on the letters by $\pi_B(a) = a$ if $a \in B$ and $\pi_B(a) = \epsilon$ otherwise. A word w over the alphabet $X \cup L$ is a *Dyck building* if $\pi_X(w)$ is a Dyck word, and there exists a non-negative integer K, the *rank* of w, such that $|w|_k = |w|_{\overline{k}} = 1$ for $k \leq K$ and $|w|_k = |w|_{\overline{k}} = 0$ otherwise. See examples on Figure 3. When a sequence $l_1, l_2 \ldots l_i$ of labels occurs between two letters x or $\overline{x}$, we represent them, on the corresponding vertex of the Dyck path, by a stack of labels where l_1 is at the bottom and l_i at the top. In our construction, the labels of index i in a Dyck building indicate where to insert the rise and the fall when we read the value i in the almost decreasing sequence.

A step-by-step definition of these insertions requires updates of the labels during each insertion to prepare the following insertions. We use rewriting rules to describe these updates. The set $G = \{A, \overline{A}, B, \overline{B}\}$ is the alphabet of *generic labels*. A ordered pair $m = (u, v)$ of words over the alphabet $X \cup G$ is an *insertion mode* if all the letters of $X \cup G$ occur exactly once in uv, x occurs in u, $\overline{x}$ occurs in v, A occurs before $\overline{A}$ in uv and B occurs before $\overline{B}$ in uv. For instance, the pair $(BxA\overline{B}, \overline{x}\overline{A})$ is an insertion mode. The substitution of all occurrences of the letter a by the word u in the word w is denoted $w[a := u]$. Two substitutions performed in parallel are denoted by $w[a := u, b := v]$ whereas a sequence of substitutions, first of the occurrences of a and then of the occurrences of b, is denoted $w[a := u][b := v]$. For example, $abc[a := bb, b := c] = bbcc$ and $abc[a := bb][b := c] = cccc$. The *insertion* is a map ρ with three arguments: a Dyck building w, an insertion mode $m = (u, v)$ and a value k in an almost decreasing sequence. This triplet is mapped to the

Dyck building

$$\rho_m^k(w) = w[k := u, \overline{k} := v][A := k, \overline{A} := \overline{k}, B := k+1, \overline{B} := \overline{k+1}]$$

also called the result of the insertion according to m of the value k in w.

Repeating this procedure, with a fixed insertion mode m, for all values of an almost decreasing sequence leads to the algorithm on Figure 2 where comments are enclosed by $/*$ and $*/$.

Input: An almost decreasing sequence $s = (s_k)_{k=1\dots n}$
Parameter: An insertion mode m
$w_0 := 0\overline{0}$; /* Start from the "empty" building containing only labels */
For k from 1 to n do
  $w'_k := \pi_{X \cup L(s_k)}(w_{k-1})$; /* Erase in w the labels of index greater than s_k */
  $w_k := \rho_m^{s_k}(w'_k)$; /* Insert x and $\overline{x}$ in w at the location of the labels s_k and $\overline{s_k}$
    then update locally the labels of indexes s_k and $1+s_k$ */
done;
Output: $\pi_X(w_n)$; /* Erase all the labels to obtain a Dyck word */

Figure 2. Step-by-step algorithm with a single insertion mode

The word output by the algorithm with the almost decreasing sequence s as input and the insertion mode m as parameter is denoted $\Upsilon_m(s)$.

Example 1.1. In Figure 3, we trace the algorithm during the computation of $\Upsilon_{(BA\overline{B}x,\overline{A}\overline{x})}(0,1,1,2,3,1,2)$.

In terms of words :

$w_0 = 0\overline{0}$

$\longrightarrow^0 \quad w_1 = 10\overline{1}x\overline{0}\overline{x}$

$\longrightarrow^1 \quad w_2 = 21\overline{2}x0\overline{1}\overline{x}x0\overline{x}$

$\longrightarrow^1 \quad w_3 = 21\overline{2}xx0\overline{1}\overline{x}\overline{x}x\overline{0}\overline{x}$

$\longrightarrow^2 \quad w_4 = 32\overline{3}x1\overline{2}\overline{x}xx0\overline{1}\overline{x}\overline{x}x\overline{0}\overline{x}$

$\longrightarrow^3 \quad w_5 = 43\overline{4}x2\overline{3}\overline{x}x1\overline{2}\overline{x}xx0\overline{1}\overline{x}\overline{x}x\overline{0}\overline{x}$

$\longrightarrow^1 \quad w_6 = x\overline{x}x21\overline{2}x\overline{x}xx0\overline{1}\overline{x}\overline{x}\overline{x}x\overline{0}\overline{x}$

$\longrightarrow^2 \quad w_7 = x\overline{x}x32\overline{3}x1\overline{2}\overline{x}x\overline{x}xx0\overline{1}\overline{x}\overline{x}\overline{x}x\overline{0}\overline{x}$

Output: $x\overline{x}xx\overline{x}x\overline{x}xx\overline{x}\overline{x}\overline{x}x\overline{x}$

In terms of paths:

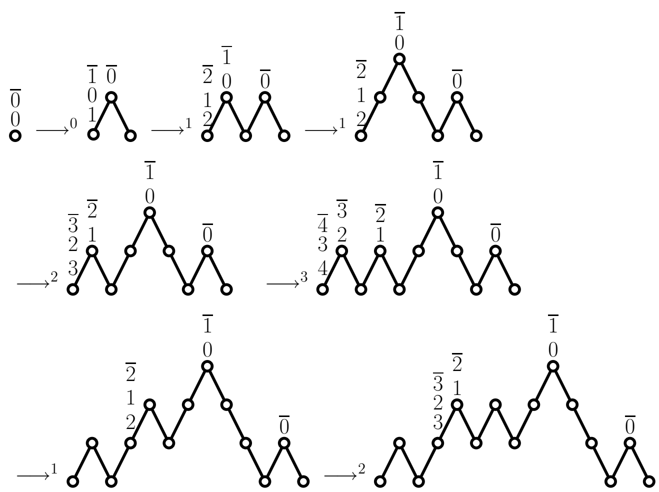


Figure 3. An example of uniform insertion according to $(BA\overline{B}x, \overline{A}\overline{x})$

First, we check that this algorithm has an expected behavior:

Lemma 1.2. *For any insertion mode m, the transformation Υ_m maps almost decreasing sequences of size n to Dyck words of size n.*

Proof. For the smallest objects, $\Upsilon_m(\emptyset) = \epsilon$ and $\Upsilon_m(0) = x\overline{x}$. For longer sequences there is an invariant in this algorithm: after k loops ($k \geq 1$), w is a Dyck building of rank $s_k + 1$ in which $\pi_X(w)$ is a Dyck word of size k and any rising label i appears before the falling label $\overline{i}$. The constraints of order on the generic labels in the definition of insertion modes implies that this invariant is preserved.

□

If $\mathcal{P}$ denotes the set of insertion modes and F is defined by $F(m,s) = \Upsilon_m(s)$, Lemma 1.2 proves that the ordered pair $(\mathcal{P}, F)$ is a uniform description.

2. A systematic study of insertion modes

There are only finitely many programs in the uniform description $(\mathcal{P}, F)$ defined in Section 1. We count them. Then we show that certain transformations on insertion mode produces maps $\Upsilon_m = F(m, .)$ that are either equal, or equivalent through a very simple involution. Then, with the help of the computer, we list counter-examples for maps Υ_m that are not one-to-one, and, with a pen and a paper, we prove that the remaining maps are bijections. Thus we fix the set $\mathcal{B}$ of 12 bijections defined by this uniform description. A trivial symmetry on Dyck paths relates each bijection to another one, so we obtain only 6 significantly distinct bijections. Each bijection in $\mathcal{B}$ is related to a parameter on Dyck path with the same distribution as the length of the last descent. We provide a description for 5 of these 6 parameters which are sometimes classical sometimes new.

2.1. Relations between insertion modes. If we consider the comma as a letter in an insertion mode, there are seven different letters in an insertion mode. Since we independently impose that A appears before $\overline{A}$, B before $\overline{B}$ and $x, \boxed{,}, \overline{x}$ in this order, there are $7!/(2! * 2! * 3!) = 210 = |\mathcal{P}|$ insertion modes.

2.1.1. *Insertion modes defining the same maps.* We define three relations of equivalence on insertion modes. Let $m_1 = (u_1, v_1)$ and $m_2 = (u_2, v_2)$ be two insertion modes. These modes are *rising-equivalent*, $m_1 \equiv_{\text{rise}} m_2$, if there exists $Y \in \{A, B\}$ such that $m_1 = (Yx, v)$ and $m_2 = (xY, v)$. They are *falling-equivalent*, $m_1 \equiv_{\text{fall}} m_2$, if there exists $\overline{Y} \in \{\overline{A}, \overline{B}\}$ such that $m_1 = (u, \overline{Y}\overline{x})$ and $m_2 = (u, \overline{x}\overline{Y})$. These two modes are *peak-equivalent*, $m_1 \equiv_{\text{peak}} m_2$, if $u_1v_1 = u_2v_2$ and $A\overline{A}$, $B\overline{B}$ are factors of u_1v_1.

LEMMA 2.1. *Two insertion modes m_1 and m_2 that are either rising, falling or peak-equivalent, define the same map $\Upsilon_{m_1} = \Upsilon_{m_2}$.*

PROOF. Let m_1 and m_2 be two insertion modes, s an almost decreasing sequence of n integers. We denote $(w^1_k)_{k=1\dots n}$, respectively $(w^2_k)_{k=1\dots n}$, the sequence of Dyck buildings $(w_k)_{k=1\dots n}$ obtained in the algorithm when the input is s and the parameter m_1, respectively m_2.

• First we assume that m_1 and m_2 are rising-equivalent ($m_1 \equiv_{\text{rise}} m_2$). The key observation is that when m_1 corresponds to an insertion of a rise at the beginning of a sequence of rises x^j, m_2 corresponds to an insertion of a rise at the end of the same sequence of rises, leading to the same Dyck word. Formally, we define an equivalence $\equiv_r$ over Dyck buildings as the symmetric and transitive closure of the relation $\longrightarrow_r$ that corresponds to the commutation of a rising label i and the rise at its right : $u.x.i.v \longrightarrow_r u.i.x.v$. To prove that $w^1_k \equiv_r w^2_k$, we will prove the following stronger fact: given the Dyck buildings w and w' such that $w \longrightarrow_r w'$, the insertions of the value i in w and w' according to the modes m_1 or m_2, leads to four equivalent Dyck buildings:

$$\rho^i_{m_1}(\pi_{X\cup L(i)}(w)) \equiv_r \rho^i_{m_2}(\pi_{X\cup L(i)}(w)) \equiv_r \rho^i_{m_1}(\pi_{X\cup L(i)}(w')) \equiv_r \rho^i_{m_2}(\pi_{X\cup L(i)}(w')).$$

Since $w \longrightarrow_r w'$, there exists a label j in w and w' such that $w = u'.j.x.u''$ and $w' = u'.x.j.u''$. We discuss according to the relative values of i and j:

- $j > i$ The rising label j is erased by $\pi_{X,L(i)}$ so $\pi_{X,L(i)}(w) = \pi_{X,L(i)}(w')$. The insertion according to m_1 (respectively m_2) leads to $v.(i.x).v'$ (respectively $v.(x.i).v'$). These two buildings are clearly equivalent.
- $j = i$ We write $w = u.i.x.u'$ and $w' = u.x.i.u'$. We observe that
$$\rho^i_{m_1}(\pi_{X\cup L(i)}(w)) = u.(i.x).x.v' \longrightarrow_r \rho^i_{m_2}(\pi_{X\cup L(i)}(w)) = u.(x.i).x.v'$$
$$= \rho^i_{m_1}(\pi_{X\cup L(i)}(w')) = u.x.(i.x).v' \longrightarrow_r \rho^i_{m_2}(\pi_{X\cup L(i)}(w')) = u.x.(x.i).v'$$
where $v' = \rho^i_{m_1}(u') = \rho^i_{m_2}(u')$ by definition of $\equiv_{\text{rise}}$.
- $j < i$ We assume that $w = u.j.x.u'.i.u''$ and $w' = u.x.j.u'.i.u''$. We observe that
$$\rho^i_{m_1}(\pi_{X\cup L(i)}(w)) = u.j.x.u'.(i.x).v'' \longrightarrow_r \rho^i_{m_2}(\pi_{X\cup L(i)}(w)) = u.j.x.u'.(x.i).v''$$
$$\longrightarrow_r \rho^i_{m_2}(\pi_{X\cup L(i)}(w')) = u.x.j.u'.(x.i).v'' \longrightarrow_r \rho^i_{m_1}(\pi_{X\cup L(i)}(w')) = u.x.j.u'.(i.x).v''$$
where $v'' = \rho^i_{m_1}(u'') = \rho^i_{m_2}(u'')$ by the definition of $\equiv_{\text{rise}}$. The careful reader will check that the relative positions of the rising labels i and j do not perturb the proof.

For $k = 0$, the assumption $w_0^1 \equiv_r w_0^2$ is satisfied since $w_0^1 = 0\overline{0} = w_0^2$. By induction, for all $k \leq n$, $w_k^1 \equiv_r w_k^2$ and in particular for $k = n$. Since the projection over the alphabet X is the same for two buildings $\equiv_r$-equivalent :

$$\Upsilon_{m_1}(s) = \pi_X(w_n^1) = \pi_X(w_n^2) = \Upsilon_{m_2}(s).$$

• If $m_1 \equiv_{\text{fall}} m_2$, the proof of $\Upsilon_{m_1} = \Upsilon_{m_2}$ is symmetric to the previous one.

• We assume that $m_1 = (u_1, v_1)$ and $m_2 = (u_2, v_2)$ are peak-equivalent ($m_1 \equiv_{\text{peak}} m_2$). We check by induction on k that $w_k^1 = w_k^2$ and that the labels of index i, if they appear, appear as the factor $i\overline{i}$. This assumption is satisfied for $k = 0$ since $w_0^1 = 0\overline{0} = w_0^2$. We suppose that it is satisfied for $k-1$. Since $s_k\overline{s_k}$ is a factor of $w_{k-1}^1 = w_{k-1}^2$, the rewriting induced by s_k corresponds to the insertion at the same place, in terms of generic labels, of $u_1.v_1$ respectively $u_2.v_2$ which are equal by the definition of $\equiv_{\text{peak}}$. Thus $w_k^1 = w_k^2$. Moreover, $A\overline{A}$ and $B\overline{B}$ are factors of $u_1v_1 = u_2v_2$ so $s_k\overline{s_k}$ and $s_{k+1}\overline{s_{k+1}}$ are factors of $w_k^1 = w_k^2$. For any index $i < s_k$, $i\overline{i}$ remains a factor of $w_k^1 = w_k^2$ since the insertion in $w_{k-1}^1 = w_{k-1}^2$ does not insert anything between the occurrence of i and $\overline{i}$. $\square$

2.1.2. *Two symmetries on insertion modes.* The reflexion according to the vertical axis defines a natural involution over Dyck paths. We generalize this mapping to words over an alphabet $L \cup \overline{L} \cup K$. Let w be a word, the *mirror word* $mir(w)$ of w is recursively defined by $mir(\epsilon) = \epsilon$, for $l \in L$, $mir(l.w') = mir(w').\overline{l}$, for $\overline{l} \in \overline{L}$, $mir(\overline{l}.w') = mir(w').l$ and for $k \in K$, $mir(k.w) = mir(w').k$. Let $m = (u, v)$ be an insertion mode, the *mirror insertion mode* is $mir(m) = (mir(v), mir(u))$. The exchange of the (generic) labels of indexes A and B leads to the notion of *exchanged insertion mode* :

$$exc(m) = (u[A := B, B := A, \overline{A} := \overline{B}, \overline{B} := \overline{A}], v[A := B, B := A, \overline{A} := \overline{B}, \overline{B} := \overline{A}]).$$

Let $s = (s_k)_{k=1\ldots n}$ be a sequence of n integers and $i \in \mathbb{N}$; by definition the sequence $t = s \oplus i$ is such that $t_k = s_k + i$ for all $k = 1 \ldots n$. An almost decreasing sequence s admits a single decomposition

$$s = 0, t_1 \oplus 1, t_2$$

where t_1 and t_2 are almost decreasing sequences. ($t_1 \oplus 1$ is the sequence of integers before the second 0 of s, which is the beginning of t_2, if any.) We use this decomposition to recursively define a map exc over almost decreasing sequences : $exc(\emptyset) = \emptyset$ and $exc(0, t_1 \oplus 1, t_2) = 0, exc(t_2) \oplus 1, exc(t_1)$. An inductive proof shows that this map exc is indeed an involution since

$$exc(exc(s)) = exc(0, exc(t_2) \oplus 1, exc(t_1)) = (0, exc(exc(t_1)) \oplus 1, exc(exc(t_2))) = (0, t_1 \oplus 1, t_2) = s.$$

The canonical bijection between almost decreasing sequences and Dyck paths relates the involution exc to an involution on Dyck paths already considered in [**4**]. These transformations of insertion modes, almost decreasing sequences and Dyck paths define relations between maps Υ_m:

LEMMA 2.2. *For any insertion mode m,*

$$mir \circ \Upsilon_m = \Upsilon_{mir(m)}$$

and

$$\Upsilon_m \circ exc = \Upsilon_{exc(m)}.$$

PROOF. Let s be an almost decreasing sequence of n integers and let $(w_k)_{k=1\ldots n}$ the sequence of buildings constructed by the algorithm computing $\Upsilon_m(s)$.

•$mir \circ \Upsilon_m = \Upsilon_{mir(m)}$: Let $(w'_k)_{k=1\ldots n}$ be the sequence of buildings when the algorithm computes $\Upsilon_{mir(m)}(s)$. We check by induction on k that $w'_k = mir(w_k)$. Thus $mir \circ \Upsilon_m(s) = \Upsilon_{mir(m)}(s)$.

•$\Upsilon_m \circ exc = \Upsilon_{exc(m)}$: Let u be a non-empty word on the alphabet $X \cup G$. We denote by $\Upsilon_m^u(s)$ the result of the algorithm when s is any sequence of non-negative integers, and the initial value w_0 is the word u. Since the labels are not moved in the buildings $(w_k)_{0 \leq k \leq n}$ while not erased, we have $\Upsilon_m(s) = \Upsilon_m^{0\overline{0}}(s) = \Upsilon_m^0(s).\Upsilon_m^{\overline{0}}(s)$. Moreover the insertion mode does not depend on the index of the labels so $\Upsilon_m^0(s) = \Upsilon_m^1(s \oplus 1)$.

We now check by induction on the length of the almost decreasing sequence s that $\Upsilon_m^0(exc(s)) = \Upsilon_{exc(m)}^0(s)$ and $\Upsilon_m^{\overline{0}}(exc(s)) = \Upsilon_{exc(m)}^{\overline{0}}(s)$. For the empty sequence $\emptyset$, $\Upsilon_m^0(exc(\emptyset)) = \epsilon = \Upsilon_{exc(m)}^0(\emptyset)$ and $\Upsilon_m^{\overline{0}}(exc(\emptyset)) = \epsilon = \Upsilon_{exc(m)}^{\overline{0}}(\emptyset)$. A non-empty sequence s satisfies $s = 0.t_1 \oplus 1.t_2$ and $exc(s) = 0.exc(t_2) \oplus 1.exc(t_1)$. The insertion mode m is written $m = (u, v)$. By definition

$$\Upsilon_m^0(exc(s)) = u[0 := \Upsilon_m^0(exc(t_1)), 1 := \Upsilon_m^1(exc(t_2) \oplus 1), \overline{\ldots}]$$

where $\overline{\cdots}$ denotes an expression for the falling labels similar to the one for rising labels that appears before the comma. Here for example, $\overline{\cdots}$ replace $\overline{0} := \Upsilon^{\overline{0}}_m(exc(t_1)), \overline{1} := \Upsilon^{\overline{1}}_m(exc(t_2) \oplus 1)$. By the induction hypothesis, $\Upsilon^0_m(exc(t_1)) = \Upsilon^0_{exc(m)}(t_1)$ and $\Upsilon^1_m(exc(t_2) \oplus 1) = \Upsilon^0_m(exc(t_2)) = \Upsilon^0_{exc(m)}(t_2)$ and so we have

$$\Upsilon^0_m(exc(s)) = u[0 := \Upsilon^0_{exc(m)}(t_1), 1 := \Upsilon^0_{exc(m)}(t_2), \overline{\cdots}].$$

By definition $exc(m) = (u, v)[A := B, B := A, \overline{\cdots}]$ and so,

$$\Upsilon^0_{exc(m)}(s) = u[0 := 1, 1 := 0, \overline{\cdots}][0 := \Upsilon^0_{exc(m)}(t_2), 1 := \Upsilon^1_{exc(m)}(t_1 \oplus 1), \overline{\cdots}],$$

$$\Upsilon^0_{exc(m)}(s) = u[1 := \Upsilon^0_{exc(m)}(t_2), 0 := \Upsilon^1_{exc(m)}(t_1 \oplus 1), \overline{\cdots}].$$

Since $\Upsilon^1_{exc(m)}(t_1 \oplus 1) = \Upsilon^0_{exc(m)}(t_1)$ we conclude that

$$\Upsilon^0_{exc(m)}(s) = u[1 := \Upsilon^0_{exc(m)}(t_2), 0 := \Upsilon^0_{exc(m)}(t_1), \overline{\cdots}] = \Upsilon^0_m(exc(s)).$$

A symmetric proof holds for $\Upsilon^{\overline{0}}_{exc(m)}(s) = \Upsilon^{\overline{0}}_m(exc(s))$. The two equalities lead to

$$\Upsilon_m(exc(s)) = \Upsilon^0_m(exc(s))\Upsilon^{\overline{0}}_m(exc(s)) = \Upsilon^0_{exc(m)}(s)\Upsilon^{\overline{0}}_{exc(m)}(s) = \Upsilon_{exc(m)}(s).$$

□

2.2. 210 insertion modes. We have written a computer program that computes $\Upsilon_m(s)$ in a sensible amount of time for all insertion modes and all almost decreasing sequences of length at most 11.

2.2.1. *The 178 insertion modes inducing non-injective maps.* For 178 insertion modes the program gives a pair of distinct almost decreasing sequences s and s' such that $\Upsilon_m(s) = \Upsilon_m(s')$. Each of these counter-examples implies that the given Υ_m is not one-to-one. Remarkably, these counter-examples have length at most 3. We do not reproduce them here. The study of these maps may be of combinatorial interest since the generating functions of the Dyck paths in the images seem to be well known (*i.e.* present in the Sloane encyclopedia of integer sequences).

2.2.2. *The 32 insertion modes inducing one-to-one maps.* For the remaining 32 insertion modes, the program shows that there are no counter-examples involving almost decreasing sequences of length shorter than 11. We have to prove "by hand" that these modes induce bijections. Using Lemma 2.1, we identify modes that define the same map. Moreover, Lemma 2.2 indicates modes related by the involutions *mir* on Dyck paths or *exc* on almost decreasing sequences. Finally we obtain a partition of the 32 modes into three classes, see Figure 4.

In each class, the equivalences and symmetries preserve the fact that the mode induces or not a bijection. So we merely have to show that one mode in the class induces a bijection to conclude that all modes in the class induce bijections.

Theorem 2.3. *For any insertion mode m in $\{(xB\overline{B}, \overline{x}A\overline{A}), (AB\overline{B}x, \overline{x}\overline{A}), (BA\overline{B}x, \overline{A}\overline{x})\}$, the map Υ_m is a size-preserving bijection between almost decreasing sequences and Dyck paths.*

The almost decreasing sequence $s = 0, 0, 1, 0, 0, 1$ is mapped to 12 distinct Dyck paths by the 12 bijections. Thus we have the following corollary :

Corollary 2.4. The 32 insertion modes in Figure 4 define 12 different bijections.

Proof. (of Theorem 2.3) All proofs follow the same scheme. For an insertion mode m, the first key element of the proof is a conjectured *labeling map* f_m which maps a Dyck path w into a Dyck building $f_m(w)$. Roughly speaking, the map f_m recovers the labels erased by π_X in the last step of the algorithm. Then we are able to recover from $f_m(w)$ the last value of the almost decreasing sequence and the last two steps inserted in the path during the algorithm. This is only true because the mode $m = (u, v)$ is *locally reversible* : u and v both contain at least one generic label. [1] Thus an induction on the size of the Dyck paths allows us to compute the reverse map of Υ_m.

Given an almost decreasing sequence of n elements we denote $\Upsilon^+_m(s)$ the last building w_n produced by the algorithm at the end of the For-loop. Thus the output is $\Upsilon_m(s) = \pi_X(\Upsilon^+_m(s))$. Our induction hypothesis is that for any Dyck path v of size n, there exists a unique almost decreasing sequence $s = s_1 \dots s_n$ such that $\Upsilon_m(s) = v$ and $\Upsilon^+_m(s) = f_m(v)$.

[1] The mode $(x, B\overline{B}\overline{x}A\overline{A})$ is an example of a not locally reversible mode.

$$
\left\{\begin{array}{ll}
 & (xB\overline{B}, \overline{x}A\overline{A}) \\
\equiv_{\text{peak}} & (x, B\overline{B}\overline{x}A\overline{A}) \\
\equiv_{\text{peak}} & (xB, \overline{B}\overline{x}A\overline{A}) \\
\equiv_{\text{rise}} & (Bx, \overline{B}\overline{x}A\overline{A})
\end{array}\right.
\iff \text{exc}
\left\{\begin{array}{ll}
 & (xA\overline{A}, \overline{x}B\overline{B}) \\
\equiv_{\text{peak}} & (x, A\overline{A}\overline{x}B\overline{B}) \\
\equiv_{\text{peak}} & (xA, \overline{A}\overline{x}B\overline{B}) \\
\equiv_{\text{rise}} & (Ax, \overline{A}\overline{x}B\overline{B})
\end{array}\right.
$$

$$
\Updownarrow_{\text{mir}} \qquad\qquad \Updownarrow_{\text{mir}}
$$

$$
\left\{\begin{array}{ll}
 & (A\overline{A}x, B\overline{B}\overline{x}) \\
\equiv_{\text{peak}} & (A\overline{A}xB\overline{B}, \overline{x}) \\
\equiv_{\text{peak}} & (A\overline{A}xB, \overline{B}\overline{x}) \\
\equiv_{\text{fall}} & (A\overline{A}xB, \overline{x}\overline{B})
\end{array}\right.
\iff \text{exc}
\left\{\begin{array}{ll}
 & (B\overline{B}x, A\overline{A}\overline{x}) \\
\equiv_{\text{peak}} & (B\overline{B}xA\overline{A}, \overline{x}) \\
\equiv_{\text{peak}} & (B\overline{B}xA, \overline{A}\overline{x}) \\
\equiv_{\text{fall}} & (B\overline{B}xA, \overline{x}\overline{A})
\end{array}\right.
$$

$$
\left\{\begin{array}{ll}
 & (AB\overline{B}x, \overline{x}\overline{A}) \\
\equiv_{\text{fall}} & (AB\overline{B}x, \overline{A}\overline{x})
\end{array}\right.
\iff \text{exc}
\left\{\begin{array}{ll}
 & (BA\overline{A}x, \overline{x}\overline{B}) \\
\equiv_{\text{fall}} & (BA\overline{A}x, \overline{B}\overline{x})
\end{array}\right.
$$

$$
\Updownarrow_{\text{mir}} \qquad\qquad \Updownarrow_{\text{mir}}
$$

$$
\left\{\begin{array}{ll}
 & (Ax, \overline{x}B\overline{BA}) \\
\equiv_{\text{rise}} & (xA, \overline{x}B\overline{BA})
\end{array}\right.
\iff \text{exc}
\left\{\begin{array}{ll}
 & (Bx, \overline{x}A\overline{AB}) \\
\equiv_{\text{rise}} & (xB, \overline{x}A\overline{AB})
\end{array}\right.
$$

$$
\left\{\begin{array}{ll}
 & (BA\overline{B}x, \overline{A}\overline{x}) \\
\equiv_{\text{fall}} & (BA\overline{B}x, \overline{x}\overline{A})
\end{array}\right.
\iff \text{exc}
\left\{\begin{array}{ll}
 & (AB\overline{A}x, \overline{B}\overline{x}) \\
\equiv_{\text{fall}} & (AB\overline{A}x, \overline{x}\overline{B})
\end{array}\right.
$$

$$
\Updownarrow_{\text{mir}} \qquad\qquad \Updownarrow_{\text{mir}}
$$

$$
\left\{\begin{array}{ll}
 & (xA, \overline{x}B\overline{AB}) \\
\equiv_{\text{rise}} & (Ax, \overline{x}B\overline{AB})
\end{array}\right.
\iff \text{exc}
\left\{\begin{array}{ll}
 & (xB, \overline{x}A\overline{BA}) \\
\equiv_{\text{rise}} & (Bx, \overline{x}A\overline{BA})
\end{array}\right.
$$

FIGURE 4. The 32 insertion modes inducing bijections

Let u be a Dyck path of size $n+1$, we look for $s = s_1, \dots s_n, s_{n+1}$ such that $\Upsilon_m(s) = u$. We want $f_m(u) = \Upsilon_m^+(s)$ and the rank of $\Upsilon_m^+(s_1, \dots s_n, s_{n+1})$ is $s_{n+1}+1$ thus necessarily $s_{n+1} = r-1$ where r is the rank of $f_m(u)$. Since m is locally reversible, we identify in $f_m(u)$ the unique rise and fall that may be inserted during the $n+1^{\text{th}}$ loop. We remove these steps in $f_m(u)$ and consider the projection v over X which is a Dyck path of size n. The induction hypothesis gives us an almost decreasing sequence s_v such that $v = \Upsilon_m(s_v)$ and $\Upsilon_m^+(s_v) = f_m(v)$. The unique possibility for s is $(s_v, r-1)$. It remains to check that it works. During the computation of $\Upsilon_m(s)$ the first n loops are similar to those of the computation of $\Upsilon_m(s_v)$ thus $u_n = v_n = \Upsilon_m^+(v) = f_m(v)$ and so $\Upsilon_m^+(s) = \rho_m^{r-1}(\pi_{X \cup L(r-2)}(f_m(v)))$. The only equality to check is

$$\rho_m^{r-1}(\pi_{X \cup L(r-2)}(f_m(v))) = f_m(u).$$

We prove this identity only for the first insertion mode, the other cases are similar and left to the reader in this summary. We focus on the simplest case: the canonical bijection. We do so to avoid technical details that the reader can find in [**7**] and to bring the essential steps into focus.

The labeling map f_m for the insertion mode $m = (AB\overline{B}x, \overline{x}\overline{A})$ will be denoted f_{descent} : the rank r of $f_{\text{descent}}(u)$ is the height of the rightmost peak in u and, for $i \leq r$, the factors $i\overline{i}$ are inserted in the rightmost vertex of height i to produce $f_{\text{descent}}(u)$ (see Figure 6).

We assume that u is a Dyck word of size $n+1$ and that the induction hypothesis is satisfied for paths of size n. Figure 5 illustrates the proof. From $f_m(u)$, the rank r is the height of the rightmost peak. The path v is obtained by deleting the rightmost rise and the next fall. The reverse operation is the insertion of a factor $x\overline{x}$ in the rightmost vertex V in v at height $r-1$. In $f_m(v)$, this vertex V contains the factor $r-1\overline{r-1}$ implying $u = \Upsilon_m(s_v, r-1)$. Now we check that $f_m(v)$ is converted into $f_m(u)$ when we insert steps according to m and the value $r-1$. Labels of indexes greater than $r-1$ in $f_m(v)$ are erased in $\pi_{X \cup L(k-1)}(f_m(v))$ and by definition there is no label greater than r in $f_m(u)$. According to m, labels of index r appear in the new rightmost peak of u and labels of index $r-1$ in the following vertex, coinciding with the labels in the rightmost vertices at height r and $r-1$ in $f_m(u)$. Since the suffix $\overline{x}^{r-1}$ of v is also a suffix of u after the insertion, labels of index $i < r-1$ coincide with those of $f_m(u)$.

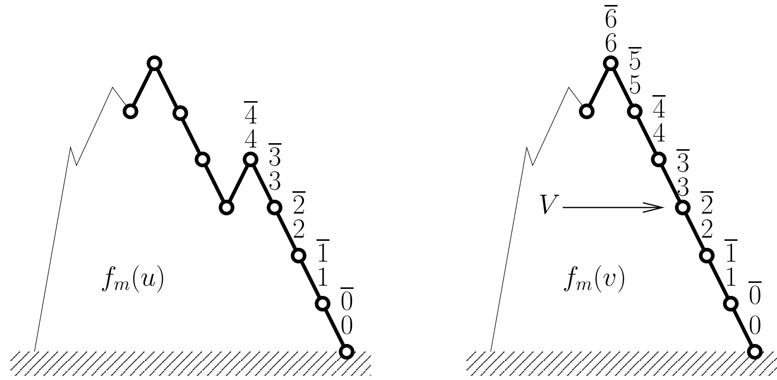


FIGURE 5. $\rho_m^{r-1}(\pi_{X\cup L(r-2)}(f_m(v))) = f_m(u)$ for $m = (xB\overline{B}, \overline{x}A\overline{A})$ and $r = 4$

The labeling map for $(AB\overline{B}x, \overline{x}\overline{A})$ is denoted f_{axis}. The rank r of $f_{\text{axis}}(u)$ is the number of falls ending on the horizontal axis. For $i < r$, the falling label $\overline{i}$ is inserted in the vertex next the $(r-i)^{\text{th}}$ fall ending on the horizontal axis. We insert the factor $01 \dots r\overline{r}$ in the first vertex of the path. See Figure 6.

The labeling map for $(AB\overline{B}x, \overline{x}\overline{A})$ is denoted f_{umbrella}. A *descent* in the Dyck path u is a maximal sequence of falls. We associate to each descent of k falls an *umbrella* of size k that is the smallest factor of u containing the k falls of the descent and the k preceding rises in u. The *center* of an umbrella is the peak preceding the k last falls. The *start* is the vertex preceding the leftmost rise of the umbrella. We consider the suffix s_u of u which is the longest concatenation of umbrellas: $s_u = u_r u_{r-1} \dots u_2 u_1$. Some additional umbrellas may appear as factors of s_u, see the umbrella of center α in Figure 6. The rank r of $f_{\text{umbrella}}(u)$ is the number of umbrellas appearing in the concatenation s_u. The label $\overline{0}$ is inserted in the center of u_1, for $i < r-1$ the factor $i\overline{i+1}$ is inserted in the center of u_{i+1} and the factor $r(r-1)\overline{r}$ in the start of u_r. This produces $f_{\text{umbrella}}(u)$.

□

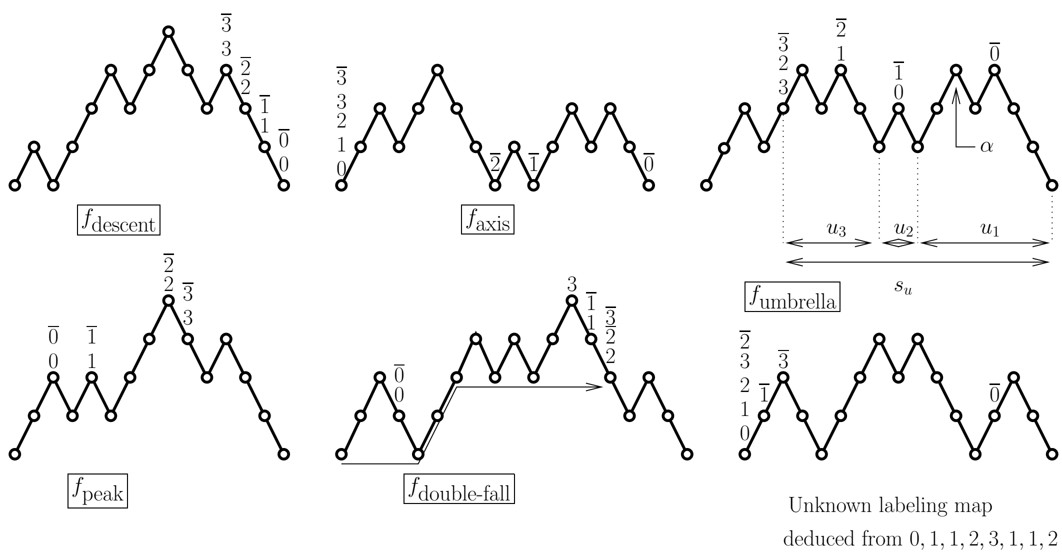


FIGURE 6. Labeling maps for five insertion modes and one unknown map

The labeling map, associated with an insertion mode inducing a bijection, defines a parameter which has the same distribution on Dyck paths as the value of the last element on the almost decreasing sequences. Some of these parameters were known to have the same distribution (in particular the length of the last descent and the number of returns to the horizontal axis). We increase the list of such parameters.

Proposition 2.1. *The following parameters have the same distribution on Dyck paths:*

- *the number of falls in the last descent,*
- *the number of falls ending on the horizontal axis,*
- *the number of umbrellas in the longest suffix which is a concatenation of umbrellas,*
- *the number of peaks before the first double fall,*
- *the number of double falls in the longest increasing prefix .*

Proof. Let u be a Dyck path. Figure 6 provides examples of the definitions of the labelings. The first three labelings already have been defined in the proof of Theorem 2.3.

The number of falls in the last descent, the number of falls ending on the horizontal axis and the number of umbrellas in the longest suffix which is a concatenation of umbrellas are respectively the rank of $f_{\text{descent}}(u)$, $f_{\text{axis}}(u)$ and $f_{\text{umbrella}}(u)$.

The labeling map for the mode $(xA\overline{A}, \overline{x}B\overline{B}) = exc(xB\overline{B}, \overline{x}A\overline{A})$ is denoted $f_{\text{double-fall}}$. The rank r of $f_{\text{double-fall}}(u)$ is the number of peaks in the longest prefix of u that does not contain a *double-fall*, *i.e.* a factor $\overline{x}\overline{x}$. For $i < r$, we insert a factor $i\overline{i}$ in the $(i+1)^{\text{th}}$ peak and the factor $r\overline{r}$ is inserted in the last vertex of the prefix (usually the first double-fall).

The labeling map for the mode $(BA\overline{A}, \overline{x}\overline{B}) = exc(AB\overline{B}x, \overline{x}\overline{A})$ is denoted $f_{\text{double-fall}}$. A *valley* in u is a vertex in the middle of a factor $\overline{x}x$. A prefix of a Dyck path is *increasing* if the height of any vertex is not strictly lower than a valley at its left. The rank r of $f_{\text{double-fall}}(u)$ is the number of double-falls in the longest increasing prefix p_u. We insert in u, for $i < r$, the factor $i\overline{i}$ in the $(i+1)^{\text{th}}$ peak, the label r in the rightmost peak in p_u and the label $\overline{r}$ in the last vertex in p_u.

□

We do not mention here the parameters equivalent up to a vertical reflexion mir. We were not able to identify the labeling map of the mode $(AB\overline{A}x, \overline{B}\overline{x})$ even we know, by Lemma 2.2, that it induces a bijection. The last building in Figure 6 is an example of a building produced by this mode.

3. Variations on the algorithm for bijections in several combinatorial contexts

We use modifications of the initial algorithm to define bijections that are relevant in several combinatorial contexts. In this extended abstract we do not emphasize these combinatorial contexts but the alteration of the algorithm. Moreover, we do not prove that the algorithms define the claimed bijections. The interested reader will find detailed proofs and other examples in [**7**].

3.1. Cyclic permutation of the labels. A *cyclic permutation of the label indexes* in a building w of rank k is denoted $Cyc(w)$. It consists of replacing for $0 \leq i \leq k$, the label i, respectively $\overline{i}$ by the label $(i+1 \mod k)$, respectively $\overline{(i+1 \mod k)}$. We generalize the algorithm presented in the Section 1 by performing a cyclic permutation at the end of each For-loop : $w_n := Cyc(w_n)$. We denote by Υ_m^{Cyc} the map defined by this algorithm parametrized by the insertion mode m.

In [**5**], the authors conjectured a formula defined by summation over integer partitions relevant for a problem in algebraic combinatorics. Haiman and Haglund, see [**6**], independently proposed two different pairs of parameter on Dyck paths that interpret this formula by summation over Dyck paths. Both pairs use the *area below the Dyck path*, that is the number of squares between the path and the horizontal axis placed as in Figure 1. We have

$$C(u, v; t) = \sum_w u^{\text{area}(w)} v^{\text{dinv}(w)} t^{\text{size}(w)} = \sum_w u^{\text{area}(w)} v^{\text{bounce}(w)} t^{\text{size}(w)}$$

where in the summations w runs over Dyck paths and $\text{dinv}(w)$, respectively $\text{bounce}(w)$, are the parameters defined by Haiman respectively Haglund. There exists another definition of $C(u, v; t)$ which is clearly symmetric in u and v. An open problem is to find a bijection that explains directly this symmetry in terms of Dyck paths.

If we consider the diagonal of squares below each rise, as in Figure 1, we remark that the area of a path is also the sum of the heights of the rises. Thus the canonical bijection shows that the area is distributed over the Dyck paths as the sum $\sum_{i=1}^{k} s_i$ on the almost decreasing sequences. In fact all the previous bijections Υ_m define parameters on Dyck path distributed as the sum of the value of the almost decreasing sequences but alas no pair of parameters has the same join distribution as $(\text{area}, \text{dinv})$ and $(\text{area}, \text{bounce})$. We show a variation of our algorithm that convert the sum of the s_i into *dinv*. In [**7**], we also provide a labeling map to define a bijection that convert the sum of the s_i into *bounce*. In the future, we plan (hope) to use these algorithms to produce other pairs of parameters whose distribution defines $C(u, v; t)$.

Let w be a Dyck path of size n and let $h = h_1, h_2 \dots h_n$ be the sequence of the height of rises in w. The parameter $dinv(w)$ counts the numbers of pairs (i, j) such that $i < j$ and $h_i \in \{h_j, 1 + h_j\}$.

We will use the labeling map f_{dinv} illustrated in Figure 7. Let k be the number of vertices at maximal height H in w and let l be the number of vertices at height $H - 1$ lying to the right of the rightmost vertex at height H. The rank of $f_{\text{dinv}}(w)$ is $k + l - 1$. For $0 \leq i \leq k - 1$ the factor $i\bar{i}$ is inserted in the $(k - i)^{\text{th}}$ vertex at height H and for $k \leq i \leq k + l - 1$, the factor $i\bar{i}$ is inserted in the $(k + l - i)^{\text{th}}$ vertex at height $H - 1$ lying to the right of the rightmost vertex at height H. In $f_{\text{dinv}}(w)$, an insertion of a rise and a fall in the labels of index k increases the parameter *dinv* by exactly k.

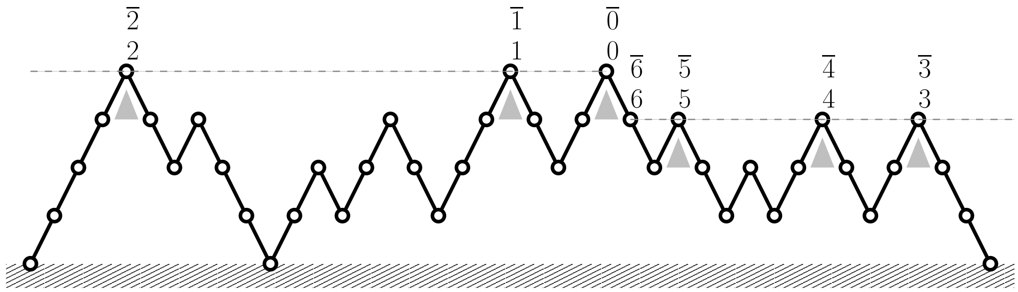


FIGURE 7. Map labeling f_{dinv} for the parameter dinv where $k = 3$ and $l = 4$

PROPOSITION 3.1. *The map $\Upsilon^{Cyc}_{(xB\overline{B},\overline{x}A\overline{A})}$ is a bijection that maps an almost decreasing sequence s to a Dyck path w such that*

$$\sum_{k=1}^{n} s_k = dinv(w).$$

The proof checks that $f_{\text{dinv}}(w)$ equals $\Upsilon^{Cyc,+}_{(xB\overline{B},\overline{x}A\overline{A})}(s)$ which is the last building w_n at the end of the For-loop in the generalized algorithm. A similar map was presented in [**1**] in terms of plane trees.

3.2. Insertion depending on the parity of the value. We define an algorithm parametrized by two insertion modes m_1 and m_2. If the value s_k in the almost decreasing sequence is even, we use m_1 to compute w_k otherwise we use m_2. For the example traced on Figure 8 we use in the even case the insertion mode $m_1 = (B\overline{B}x, A\overline{A}\overline{x})$ and in the odd case the mode $m_2 = exc(m_1) = (A\overline{A}x, B\overline{B}\overline{x})$. This defines a bijection denoted $\Upsilon_{[(B\overline{B}x,A\overline{A}\overline{x});(A\overline{A}x,B\overline{B}\overline{x})]}$, that we use in the following context.

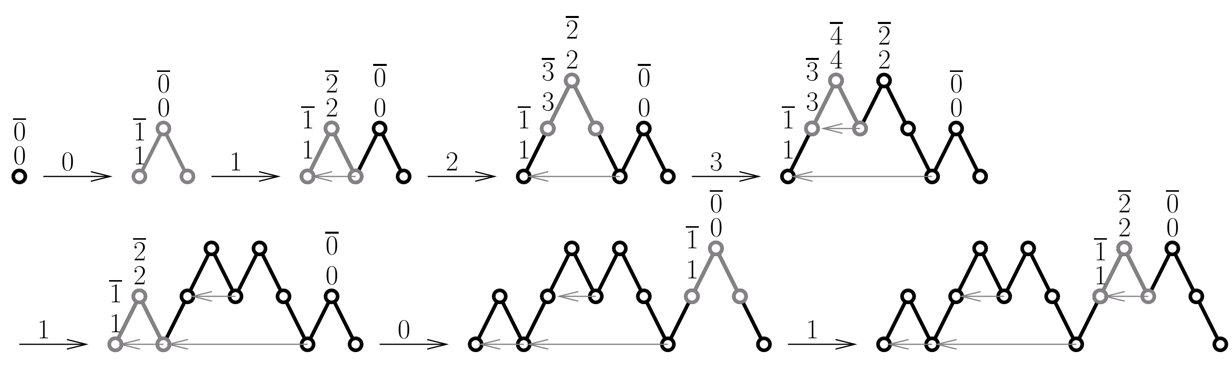


FIGURE 8. The image of $0, 1, 2, 3, 1, 0, 1$ with insertion depending on the parity

In [**7**], we interpret combinatorially the formal manipulations of generating functions involved in the solution of an equation usually used in the kernel method [**3**]. At some point we need a bijection that translates a parameter into another to conclude the interpretation. Here we only present these parameters and the description of the bijection by the extension of the algorithm that distinguish the parity of s_i's.

The first parameter is the *height* of a Dyck path w, that is the maximal value of ordinate of a vertex in w. A *ray* in a Dyck path is a segment with one endpoint a valley, the *source* of the ray, and the other one the preceding vertex of the path that is at the same height. The *ray height of a peak* is the number of rays that cross the vertical segment starting at the peak and finishing on the horizontal axis. The *ray height* of a Dyck path is the maximal height of its peaks. Figure 9 illustrates these definitions.

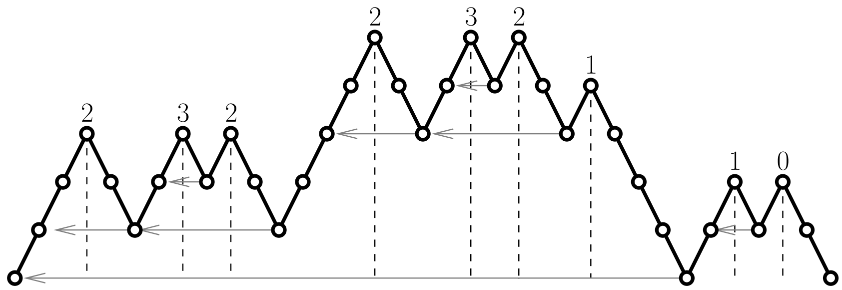


FIGURE 9. A Dyck path of ray height 3

PROPOSITION 3.2. *For any $N \geq 0$, there are as many Dyck paths of height at most $2N+1$ as Dyck paths of ray height at most N.*

The bijection $\Upsilon_{[(B\overline{B}x,A\overline{A}\overline{x});(A\overline{A}x,B\overline{B}\overline{x})]}$ maps almost decreasing sequences whose maximal value is either $2N$ or $2N+1$ to Dyck paths of ray height exactly N.

Acknowledgments. I thank Mireille Bousquet-Mélou for technical and linguistic advices. I am also grateful to Mike Robson for linguistic help.

References

[1] Barcucci, Elena and Del Lungo, Alberto and Pergola, Elisa and Pinzani, Renzo. A methodology for plane tree enumeration. Discrete Mathematics, 180 45–64, 1998.

[2] Bousquet-Mélou, Mireille, Algebraic generating functions in enumerative combinatorics and context-free languages. STACS 2005. Lecture Notes in Comput. Sci. 3404, 18–35, 2005.

[3] Bousquet-Mélou, Mireille and Petkovšek, Marko. Linear recurrences with constant coefficients: the multivariate case. Discrete Mathematics, 225(1-3) 51–75, 2000.

[4] Deutsch, Emeric, An involution on Dyck paths and its consequences, Discrete Mathematics, 204, 163–166,1999.

[5] Garsia, Adriano. M. and Haiman, Mark. A remarkable q,t-Catalan sequence and q-Lagrange inversion, Journal of Algebraic Combinatorics, 5(3) 191–244, 1996.

[6] Haglund, Jim, Conjectured statistics for the q,t-Catalan numbers, Advances in Mathematics, 175(2) 319-334, 2003.

[7] Le Borgne, Yvan, Variations combinatoires sur des classes d'objets comptées par la suite de Catalan (French), University Bordeaux I, December 2004. Available at: `www.labri.fr/~borgne`

[8] Stanley, Richard P., Enumerative combinatorics. Vol. 2, Cambridge Studies in Advanced Mathematics, 62, 1999. The Catalan addendum is available at `www-math.mit.edu/~rstan/ec/`.

[9] Sulanke, Robert A., Constraint-sensitive Catalan path statistics having the Narayana distribution, Discrete Mathematics, 204(1-3), 397–414, 1999.

CNRS/LaBRI, UNIVERSITY BORDEAUX I, TALENCE, FRANCE
E-mail address: `borgne@labri.fr`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Algebraic structures on Grothendieck groups of a tower of algebras

Huilan Li

ABSTRACT. The Grothendieck group of the tower of symmetric group algebras has a self-dual graded Hopf algebra structure. In this work, we define the general notation of a tower of algebras and study the two Grothendieck groups on this tower. Using representation theory, we prove that the two Grothendieck groups are graded Hopf algebras. Moreover, we define a paring and show that the two Grothendieck groups are dual to each other as Hopf algebras.

RÉSUMÉ. Les configurations gréées sont des objets combinatoires inspirés par l'ansatz de Bethe, et qui sont en correspondence avec les éléments cristallins de plus haut poids. Dans cette note, nous introduisons le concept de "configurations gréées généralisées", en construisant une structure cristalline dans l'espace des configurations gréées.

1. Introduction

In 1977, L. Geissinger realized that *Sym* (symmetric functions in infinite variables) is a self-dual graded Hopf algebra [**6**], which can be interpreted as the self-dual Grothendieck Hopf algebra of the tower of symmetric groups $\bigoplus_{n\geq 0}\mathbb{C}S_n$ using the work of Frobenius and Schur. After this, mathematicians have encountered many instances of combinatorial Hopf algebras that can be realized as the Grothendieck Hopf algebras of a tower of algebras. In each instance, they study a pair of dual Hopf algebras, and it turns out that this duality can be interpreted as the duality of the Grothendieck groups of an appropriate tower of algebras. For example, C. Malvenuto and C. Reutenauer established the duality between the Hopf algebra of *NSym* (noncommutative symmetric functions) and the Hopf algebra of *QSym* (quasi-symmetric functions) when looking at the combinatorics of descents [**12**]. Later, D. Krob and J.-Y. Thibon showed that this duality can be interpreted as the duality of the Grothendieck groups associated to $\bigoplus_{n\geq 0} H_n(0)$ the tower of Hecke algebras at q = 0 [**10**]. More recently, N. Bergeron, F. Hivert, and J.-Y. Thibon showed that if one uses $\bigoplus_{n\geq 0} HCl_n(0)$ the tower of Hecke-Clifford algebras at q = 0, then one gets a similar interpretation for the duality between the $Peak$ algebra and its dual [**2**].

In this work, we study the algebraic structure on the Grothendieck groups $G_0(A)$ and $K_0(A)$ in the more general case where $(A = \bigoplus_{n\geq 0} A_n, \rho_{m,n})$ is a graded algebra and each component A_n is an algebra. We will call A a tower of algebras if it satisfies some conditions. No formal study of this kind has been done so far. Up to this point it was not clear what were the right conditions to impose on a tower of algebra to get the desired algebraic structure on their Grothendieck groups. Here, we find a list of axioms on a tower of algebras which will imply that their Grothendieck groups are graded Hopf algebras. Moreover, we define a paring and show that the corresponding Grothendieck groups are dual to each other as Hopf algebras if the tower of algebras satisfying an additional condition.

This paper is divided into 5 sections as follows. Section 1 is the introduction. In Section 2 we recall some definitions and propositions about bialgebras and Grothendieck groups. In Section 3 we discuss the axioms on a tower of algebras $(A = \bigoplus_{n\geq 0} A_n, \rho_{m,n})$ with ρ preserving unities so that their Grothendieck

2000 *Mathematics Subject Classification.* Hopf algebras 16W30; Grothendieck groups 18F30.
Key words and phrases. graded algebra, Hopf algebra, Grothendieck group, representation.
The author thanks NSERC and CRC.

groups are graded Hopf algebras. Moreover, we define a paring and show that the Grothendieck groups are dual to each other as Hopf algebras. In Section 4 we weaken the condition of ρ and modify the definitions of inductions and restrictions to get the similar results as above. In Section 5 we will give some examples to indicate that the Grothendieck groups of a tower of algebras satisfying these axioms are Hopf algebras dual to each other, and these axioms are necessary.

2. Notations and Propositions

In this section there is a brief review of some ideas from the theory of bialgebras [**6**] and Grothendieck groups [**8**] which is useful for later discussion.

Definition 2.1. Let K be a commutative ring. A K-*algebra* B is a K-module with *multiplication* $\pi : B \otimes_K B \to B$ and *unit map* $\mu : K \to B$ satisfying *associativity* and *unitary property*, i.e., $\pi(\pi \otimes 1) = \pi(1 \otimes \pi)$ and $\pi(\mu \otimes 1) = \pi(1 \otimes \mu)$, where 1 is the identity map of module B. Denote this algebra by the triple (B, π, μ).

A K-*coalgebra* C is a K-module with *comultiplication* $\Delta : C \to C \otimes C$ and *counit map* $\epsilon : C \to R$ satisfying *coassociativity* and *counitary property*, i.e., $(\Delta \otimes 1)\Delta = (1 \otimes \Delta)\Delta$ and $(\epsilon \otimes 1)\Delta = (1 \otimes \epsilon)\Delta$, where 1 is the identity map of module C. Denote this coalgebra by the triple (C, Δ, ϵ).

If a K-module B is simultaneously an algebra and a coalgebra it is called a *bialgebra* provided these structures are compatible in the sense that the comultiplication and counit are algebra homomorphisms. Explicitly this means that $\epsilon(\mu(1)) = 1$, $\epsilon(gh) = \epsilon(g)\epsilon(h)$, $\Delta\mu(1) = \mu(1) \otimes \mu(1)$, and $\Delta(gh) = \Sigma g_i h_p \otimes g'_i h'_p$ if $\Delta(g) = \Sigma g_i \otimes g'_i$ and $\Delta(h) = \Sigma h_p \otimes h'_p$, where 1 is the unity of K and $gh = \pi(g \otimes h)$. This is equivalent to requiring that the multiplication and unit map are coalgebra homomorphisms. Denote this bialgebra by the 5-tuple $(B, \pi, \mu, \Delta, \epsilon)$.

A K-linear map $\gamma : H \to H$ on a bialgebra H is an *antipode* if for all h in H, $\Sigma h_i \gamma(h'_i) = \epsilon(h)1_H = \Sigma\gamma(h_i)h'_i$ when $\Delta h = \Sigma h_i \otimes h'_i$. A *Hopf algebra* is a bialgebra with antipode.

Definition 2.2. An algebra B is a *graded algebra* if there is a direct sum decomposition $B = \bigoplus B_i$ $(i \geq 0)$ such that the product of homogeneous of degrees p and q is homogeneous of degree $p + q$, that is, $\pi(B_p \otimes B_q) \subseteq B_{p+q}$, and $\mu(K) \subseteq B_0$.

A coalgebra C is a *graded coalgebra* if there is a direct sum decomposition $C = \bigoplus C_i$ $(i \geq 0)$ such that $\Delta(C_n) \subseteq \bigoplus(C_k \otimes C_{n-k})$ and $\epsilon(C_n) = 0$ if $n \geq 1$.

A bialgebra $H = \bigoplus H_i$ over K is called *graded connected* if it is $\mathbb{Z}$-graded, concentrated in nonnegative degrees, and satisfies $H_0 = K1_H$, where K is a field.

It is a known fact that a connected bialgebra is a connected Hopf algebra [**17**].

The coassociativity and counitary property are dual to associativity and unitary property, respectively. It is natural to expect the dual of a coalgebra to be an algebra and vice versa. In fact, if a module is a graded bialgebra with all homogeneous components finitely generated, then its graded dual is also a graded bialgebra [**6**].

The definition of Grothendieck groups is introduced in [**8**]. Let B be an arbitrary algebra. Denote

$$\begin{aligned} {}_B\mathcal{M} &= \text{the category of all left } B\text{-modules},\\ {}_B\text{mod} &= \text{the category of all finitely generated left } B\text{-modules},\\ \mathcal{P}(B) &= \text{the category of all finitely generated projective left } B\text{-modules}. \end{aligned}$$

Definition 2.3. Let $\mathcal{C}$ be one of the above categories. Let $\mathbf{F}$ be the free abelian group generated by symbols (M), one for each isomorphism class of modules M in $\mathcal{C}$. Let $\mathbf{F_0}$ be the subgroup of $\mathbf{F}$ generated by all expressions

$$(M) - (L) - (N)$$

arising from all short exact sequences

$$0 \to L \to M \to N \to 0$$

in $\mathcal{C}$. The *Grothendieck group* $K_0(\mathcal{C})$ of the category $\mathcal{C}$ is defined by

$$K_0(\mathcal{C}) = \mathbf{F}/\mathbf{F_0},$$

an abelian additive group. For $M \in \mathcal{C}$, let $[M]$ denote its image in $K_0(\mathcal{C})$.

Each $x \in K_0(\mathcal{C})$ is expressible as a difference $[M] - [N]$ with M, $N \in \mathcal{C}$, though not in a unique manner. Furthermore, it may occur that $x = 0$ even though M is not isomorphic to N.

Definition 2.4. The *Grothendieck group* $G_0(B)$ of the algebra B is defined by

$$G_0(B) = K_0({}_B\mathrm{mod}).$$

The *Grothendieck group* $K_0(B)$ of the algebra B is defined by

$$K_0(B) = K_0(\mathcal{P}(B)).$$

Thus, $G_0(B)$ is generated by expressions $[M]$, one for each isomorphism class (M) of finitely generated left B-modules M, with relations

$$[M] = [M'] + [M'']$$

for each short exact sequence $0 \to M' \to M \to M'' \to 0$ of finitely generated left B-modules.

$K_0(B)$ is generated by expressions $[P]$, one for each isomorphism class (P) of finitely generated left B-modules P, with relations

$$[P \oplus P'] = [P] + [P']$$

for all $P, P' \in \mathcal{P}(B)$. (Note that each short exact sequence $0 \to P' \to P \to P'' \to 0$ of modules from $\mathcal{P}(B)$ must split, because P'' is a projective B-module. Hence, the defining relations for $K_0(B)$ can be expressed in the simpler form involving direct sums, rather than exact sequences from $\mathcal{P}(B)$.)

Now let B be a finite-dimensional algebra over a field K. Let $\{V_1, \cdots, V_s\}$ be a complete list of nonisomorphic simple B-modules. Then their projective covers $\{P_1, \cdots, P_s\}$ are a complete list of nonisomorphic indecomposable projective B-modules [**13**]. With these lists, we have

Proposition 2.1.

$$G_0(B) = \bigoplus_{i=1}^{s} \mathbb{Z}[V_i]$$

is a free abelian group with basis $\{[V_1], \cdots, [V_s]\}$. *And*

$$K_0(B) = \bigoplus_{i=1}^{s} \mathbb{Z}[P_i]$$

is a free abelian group with basis $\{[P_1], \cdots, [P_s]\}$.

Let A be an algebra and $B \subseteq A$ a subalgebra. Let M be a (left) A-module and N a (left) B-module, then the *induction* of N from B to A is $\mathrm{Ind}_B^A N = A \otimes_B N$ an A-module and the *restriction* of M from A to B is $\mathrm{Res}_B^A M = \mathrm{Hom}_A(A, M)$ a B-module.

3. Grothendieck groups of a tower of algebras (Preserving unities)

In this section, first we list all the axioms we need on a graded algebra $(A = \bigoplus_{n\geq 0} A_n, \rho_{m,n})$ with ρ preserving unities. Then we define the inductions and restrictions on their Grothendieck groups $G_0(A)$ and $K_0(A)$ respectively. After this, we use these definitions to construct the multiplications and comultiplications on $G_0(A)$ and $K_0(A)$ and show that $G_0(A)$ and $K_0(A)$ are graded connected Hopf algebras with these operators. Moreover, we define a paring on the Grothendieck groups $G_0(A)$ and $K_0(A)$. It develops that they are dual to each other as Hopf algebras.

Let $A = \bigoplus_{n\geq 0} A_n$, we call it *a tower of algebras* over field $K = \mathbb{C}$ if the following conditions are satisfied:

(1) A_n is a finite-dimensional algebra with unit, for each n. $A_0 \cong K$.

(2) There is an external graded multiplication $\rho_{m,n} : A_m \otimes A_n \to A_{m+n}$ such that

(a) $\rho_{m,n}$ is an injective homomorphism of algebras, for all m and n (sending $1_m \otimes 1_n$ to 1_{m+n});

(b) ρ is associative, that is, $\rho_{l+m,n} \cdot (\rho_{l,m} \otimes 1_n) = \rho_{l,m+n} \cdot (1_l \otimes \rho_{m,n}) := \rho_{l,m,n}$, for all l, m, n.

(3) A_{n+m} is a two-sided projective $A_n \otimes A_m$-module by the action defined to be $a \cdot (b \otimes c) = a\rho_{m,n}(b \otimes c)$ and $(b \otimes c) \cdot a = \rho_{m,n}(b \otimes c)a$, for $a \in A_{m+n}$, $b \in A_m$ and $c \in A_n$.

(4) For every primitive idempotent g in A_{m+n}, $A_{m+n}g \cong \bigoplus(A_m \otimes A_n)(e \otimes f)$ as (left) $A_m \otimes A_n$-modules if and only if $gA_{m+n} \cong \bigoplus(e \otimes f)(A_m \otimes A_n)$ as (right) $A_m \otimes A_n$-modules for the same indexing of idempotents $(e \otimes f)$'s in $A_m \otimes A_n$.

(5) The following equality holds

$$\begin{array}{rl} & [\mathrm{Res}^{A_{m+n}}_{A_k\otimes A_{m+n-k}}\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}(M\otimes N)] \\ = & \sum_{t+s=k}[\widetilde{\mathrm{Ind}}^{A_k\otimes A_{m+n-k}}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}(\mathrm{Res}^{A_m}_{A_t\otimes A_{m-t}}M\otimes \mathrm{Res}^{A_n}_{A_s\otimes A_{n-s}}N)] \end{array}$$

for all $0<k<m+n$, M an A_m-module and N an A_n-module. We will explain the notations later.

Why we need these conditions? We can give a brief explanation here. Condition (1) guarantees that their Grothendieck groups are grade connected; with conditions (2) and (3) the inductions and restrictions are well defined; with (4) the duality holds; with (5) the multiplication and comultiplication are compatible. We will come up to the details later.

Now we define the inductions on $G_0(A)$ as follows:

$$\begin{array}{rcccl} i_{m,n} & : & G_0(A_m)\bigotimes_{\mathbb{Z}}G_0(A_n) & \to & G_0(A_{m+n}) \\ & & [M]\otimes[N] & \mapsto & [\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M\otimes N], \end{array}$$

where

$$\begin{array}{rcl} \mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M\otimes N & = & A_{m+n}\bigotimes_{A_m\otimes A_n}(M\otimes N) \\ & = & \dfrac{A_{m+n}\otimes M\otimes N}{< a\otimes[(b\otimes c)(w\otimes u)]-[a\rho_{m,n}(b\otimes c)]\otimes w\otimes u >}, \end{array}$$

for $a\in A_{m+n},\ b\in A_m,\ c\in A_n,\ w\in M$ and $u\in N$. Here let $k=t+s$, define the twisted induction

$$\begin{array}{rl} & \widetilde{\mathrm{Ind}}^{A_k\otimes A_{m+n-k}}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}(M_1\otimes M_2)\otimes(N_1\otimes N_2) \\ = & (A_k\otimes A_{m+n-k})\widetilde{\bigotimes}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}((M_1\otimes M_2)\otimes(N_1\otimes N_2)). \end{array}$$

This means

$$\begin{array}{rl} & (a\otimes b)\otimes[(c_1\otimes c_2)\cdot(w_1\otimes w_2)\otimes(d_1\otimes d_2)\cdot(u_1\otimes u_2)] \\ \equiv & [a\rho_{t,s}(c_1\otimes d_1)\otimes b\rho_{m-t,n-s}(c_2\otimes d_2)]\otimes(w_1\otimes w_2\otimes u_1\otimes u_2), \end{array}$$

where $a\in A_k,\ b\in A_{m+n-k},\ c_1\in A_t,\ c_2\in A_{m-t},\ d_1\in A_s,\ d_2\in A_{n-s},\ w_i\in M_i,\ u_i\in N_i$. Also define the restrictions

$$\begin{array}{rcccl} r_{k,l} & : & G_0(A_n) & \to & G_0(A_k)\bigotimes_{\mathbb{Z}}G_0(A_l) \text{ with } k+l=n \\ & & [N] & \mapsto & [\mathrm{Res}^{A_n}_{A_k\otimes A_l}N], \end{array}$$

where $\mathrm{Res}^{A_n}_{A_k\otimes A_l}N=\mathrm{Hom}_{A_n}(A_n,N)$ is an $A_k\otimes A_l$-module by the action defined to be $((b\otimes c)\cdot f)(a)=f(a\rho_{k,l}(b\otimes c))$, for $a\in A_n,\ b\in A_k,\ c\in A_l$ and $f\in\mathrm{Hom}_{A_n}(A_n,N)$.

Proposition 3.1. *i and r are well defined.*

Proof. Assume $[M]=[M']+[M'']$. Since A_{m+n} is a (right) projective $A_m\otimes A_n$-module, it is not difficult to get that

$$0\to A_{m+n}\otimes_{A_m\otimes A_n}(M'\otimes N)\to A_{m+n}\otimes_{A_m\otimes A_n}(M\otimes N)\to A_{m+n}\otimes_{A_m\otimes A_n}(M''\otimes N)\to 0$$

is exact as left A_{m+n}-modules by the properties of tensor product and short exact sequence. Hence

$$[\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M\otimes N]=[\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M'\otimes N]+[\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M''\otimes N].$$

Similarly,

$$[\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M\otimes N]=[\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M\otimes N']+[\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}M\otimes N'']$$

for $[N]=[N']+[N'']$. Hence i is well defined on $G_0(A)$.

Assume $[N]=[N']+[N'']$. Since $\mathrm{Hom}_{A_n}(A_n,M)\cong M$ for all A_n-modules M, it is clear that

$$0\to\mathrm{Hom}_{A_n}(A_n,N')\to\mathrm{Hom}_{A_n}(A_n,N)\to\mathrm{Hom}_{A_n}(A_n,N'')\to 0$$

is exact, which is also exact as $A_k\otimes A_l$-modules. Hence

$$[\mathrm{Res}^{A_n}_{A_k\otimes A_l}N]=[\mathrm{Res}^{A_n}_{A_k\otimes A_l}N']+[\mathrm{Res}^{A_n}_{A_k\otimes A_l}N''].$$

Therefore, all r are well defined. □

Let $G_0(A) = \bigoplus_{n\geq 0} G_0(A_n)$. We construct the multiplication and comultiplication by i and r and define the unit and counit as follows:

$$\begin{array}{rcl}
\pi & : & G_0(A)\bigotimes_{\mathbb{Z}} G_0(A) \to G_0(A) \\
 & & \text{by } \pi|_{G_0(A_k)\bigotimes G_0(A_l)} = i_{k,l} \\
\Delta & : & G_0(A) \to G_0(A)\bigotimes_{\mathbb{Z}} G_0(A) \\
 & & \text{by } \Delta|_{G_0(A_n)} = \sum_{k+l=n} r_{k,l} \\
\mu & : & \mathbb{Z} \to G_0(A) \\
 & & \text{by } \mu(a) = a[K] \in G_0(A_0), \text{ for } a \in \mathbb{Z} \\
\epsilon & : & G_0(A) \to \mathbb{Z} \\
 & & \text{by } \epsilon([M]) = \left\{ \begin{array}{ll} a & \text{if } [M] = a[K], \text{ where } a \in \mathbb{Z} \\ 0 & \text{otherwise.} \end{array} \right.
\end{array}$$

Later we will prove the associativity of π, the unitary property of μ, the coassociativity of Δ and the counitary property of ϵ, which imply that $(G_0(A), \pi, \mu)$ is an algebra and $(G_0(A), \Delta, \epsilon)$ is a coalgebra. We will also show the compatibility of the algebra and coalgebra structures to indicate that $(G_0(A), \pi, \mu, \Delta, \epsilon)$ is a graded connected bialgebra.

Now we define the inductions and restrictions on $K_0(A)$ analogously. As before,

$$\begin{array}{rccl}
i'_{m,n} & : & K_0(A_m)\bigotimes_{\mathbb{Z}} K_0(A_n) & \to & K_0(A_{m+n}) \\
 & & [P]\otimes[Q] & \mapsto & [\mathrm{Ind}_{A_m\otimes A_n}^{A_{m+n}} P\otimes Q],
\end{array}$$

where

$$\begin{array}{rcl}
\mathrm{Ind}_{A_m\otimes A_n}^{A_{m+n}} P\otimes Q & = & A_{m+n}\bigotimes_{A_m\otimes A_n}(P\otimes Q) \\
 & = & \dfrac{A_{m+n}\otimes P\otimes Q}{< a\otimes[(b\otimes c)(p\otimes q)] - [a\rho_{m,n}(b\otimes c)]\otimes p\otimes q >},
\end{array}$$

$a\in A_{m+n},\ b\in A_m,\ c\in A_n,\ p\in P$ and $q\in Q$. Let $k = t+s$. Denote

$$\begin{array}{rl}
 & \widetilde{\mathrm{Ind}}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}^{A_k\otimes A_{m+n-k}}(P_1\otimes P_2)\otimes(Q_1\otimes Q_2) \\
= & (A_k\otimes A_{m+n-k})\widetilde{\bigotimes}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}((P_1\otimes P_2)\otimes(Q_1\otimes Q_2))
\end{array}$$

the twisted induction with the same meaning as above. And set

$$\begin{array}{rccl}
r'_{k,l} & : & K_0(A_n) & \to & K_0(A_k)\bigotimes_{\mathbb{Z}} K_0(A_l) \text{ with } k+l=n \\
 & & [R] & \mapsto & [\mathrm{Res}_{A_k\otimes A_l}^{A_n} R],
\end{array}$$

where $\mathrm{Res}_{A_k\otimes A_l}^{A_n} R = \mathrm{Hom}_{A_n}(A_n, R)$ as a left projective $A_k\otimes A_l$-module by the action defined to be $((b\otimes c)\cdot f)(a) = f(a\rho_{k,l}(b\otimes c))$, $a\in A_n,\ b\in A_k,\ c\in A_l$ and $f\in \mathrm{Hom}_{A_n}(A_n, R)$.

Proposition 3.2. *i' and r' are well defined.*

Proof. To show that i' are well defined, we only need that $\mathrm{Ind}_{A_m\otimes A_n}^{A_{m+n}} P\otimes Q = A_{m+n}\bigotimes_{A_m\otimes A_n}(P\otimes Q)$ is a projective A_{m+n}-module for all projective A_m-module P and all projective A_n-module Q. This is straightforward by the properties of tensor product and short exact sequence and the property of projective modules that there is a module P' such that $P\oplus P'$ is a free module for the projective module P.

Assume R is a projective A_n-module. Since $\mathrm{Hom}_{A_n}(A_n, M)\cong M$ for all A_n-modules M, we can get that $\mathrm{Hom}_{A_n}(A_n, R)$ is a summand of some free A_n-module by the property of projective modules. Hence, $\mathrm{Hom}_{A_n}(A_n, R)$ is a $A_k\otimes A_l$-module for all k and l with $n = k+l$. Therefore, r' are well defined □

Let $K_0(A) = \bigoplus_{n\geq 0} K_0(A_n)$. Using i' and r' we also define the multiplication, comultiplication, unit and counit on $K_0(A)$.

$$\begin{array}{rcl}
\pi' & : & K_0(A) \bigotimes_{\mathbb{Z}} K_0(A) \to K_0(A) \\
 & & \text{by } \pi'|_{K_0(A_k)\bigotimes K_0(A_l)} = i'_{k,l} \\
\Delta' & : & K_0(A) \to K_0(A) \bigotimes_{\mathbb{Z}} K_0(A) \\
 & & \text{by } \Delta'|_{K_0(A_n)} = \sum_{k+l=n} r'_{k,l} \\
\mu' & : & \mathbb{Z} \to K_0(A) \\
 & & \text{by } \mu'(a) = a[K] \in K_0(A_0), \text{ for } a \in \mathbb{Z} \\
\epsilon' & : & K_0(A) \to \mathbb{Z} \\
 & & \text{by } \epsilon'([M]) = \begin{cases} a & \text{if } [M] = a[K], \text{ where } a \in \mathbb{Z} \\ 0 & \text{otherwise.} \end{cases}
\end{array}$$

Similarly, we will realize that $(K_0(A), \pi', \mu')$ is an algebra and $(K_0(A), \Delta', \epsilon')$ is a coalgebra later. It will also be verified that the compatibility of these algebra and coalgebra structures hold, i.e., $(K_0(A), \pi', \mu', \Delta', \epsilon')$ is a graded connected bialgebra.

Theorem 3.1. *(i) π is associative and $\big(G_0(A),\ \pi,\ \mu\big)$ is an algebra. So is $\big(K_0(A),\ \pi',\ \mu'\big)$.*

(ii) Δ is coassociative and $\big(G_0(A),\ \Delta,\ \epsilon\big)$ is a coalgebra. So is $\big(K_0(A),\ \Delta',\ \epsilon'\big)$.

(iii) Δ and ϵ are algebra homomorphisms and $G_0(A)$ is a graded connected bialgebra. Hence $G_0(A)$ is a graded Hopf algebra. So is $K_0(A)$.

Proof. (i) We only need to check the associativity of π, i.e., $i_{l+m,n} \cdot (i_{l,m} \otimes 1_n) = i_{l,m+n} \cdot (1_l \otimes i_{m,n})$. Form the associativity of ρ and the definition of i, we can check it directly. Same for π'.

(ii) We only need to show the coassociativity of Δ, i.e., $(r_{l,m} \otimes 1) \cdot r_{l+m,n} = (1 \otimes r_{m,n}) \cdot r_{l,m+n}$. Form the definition of r and the Adjointness Theorem [**8**], we can check it directly. Similarly for Δ'.

(iii) Using the definition of compatibility of algebra and coalgebra structures, we show that $G_0(A)$ is a graded bialgebra since condition (5) holds. From condition (1), we know that $G_0(A)$ is a graded connected bialgebra. Hence a graded Hopf algebra. Similarly for $K_0(A)$. ☐

Next we define a pairing on $K_0(A) \times G_0(A)$. With this pairing we can consider the duality between $K_0(A)$ and $G_0(A)$. The pairing is defined as follows:

$$<,>: K_0(A) \times G_0(A) \to \mathbb{Z}$$

such that

$$<[P],[M]>= \begin{cases} \dim_K\big(\mathrm{Hom}_{A_n}(P,M)\big) & \text{if } [P] \in K_0(A_n) \text{ and } [M] \in G_0(A_n) \\ 0 & \text{otherwise.} \end{cases}$$

and with the same notation $<,>: (K_0(A) \otimes K_0(A)) \times (G_0(A) \otimes G_0(A)) \to \mathbb{Z}$ by

$$<[P]\otimes[Q],[M]\otimes[N]>= \begin{cases} \dim_K\big(\mathrm{Hom}_{A_k\otimes A_l}(P\otimes Q, M\otimes N)\big) & \text{if } [P]\otimes[Q] \in K_0(A_k)\otimes K_0(A_l) \\ & \text{ and } [M]\otimes[N] \in G_0(A_k)\otimes G_0(A_l) \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 3.3. *$<,>$ is a well-defined bilinear pairing on $K_0(A) \times G_0(A)$ satisfying the following identities*

$$\begin{array}{rcl}
<[P]\otimes[Q],[M]\otimes[N]> & = & <[P],[M]><[Q],[N]> \\
<\pi'([P]\otimes[Q]),[M]> & = & <[P]\otimes[Q],\Delta[M]> \\
<\Delta'[P],[M]\otimes[N]> & = & <[P],\pi([M]\otimes[N])> \\
<\mu'(1),[M]> & = & \epsilon([M]) \\
<[P],\mu(1)> & = & \epsilon'([P]).
\end{array}$$

Proof. It is straightforward to check the linearity by the properties of short exact sequences and direst sums of modules .

The identity $<[P]\otimes[Q],[M]\otimes[N]>=<[P],[M]><[Q],[N]>$ is trivial.

To show $<\pi'([P]\otimes[Q]),[M]>=<[P]\otimes[Q],\Delta[M]>$, it is equivalent to prove that $<i'_{k,l}([P]\otimes[Q]),[M]>=<[P]\otimes[Q],r_{k,l}[M]>$, for all $[P] \in K_0(A_k)$, $[Q] \in K_0(A_l)$ and $[M] \in G_0(A_{k+l})$, which can be reached by the Adjointness Theorem.

To show $<\Delta'[P],[M]\otimes[N]>=<[P],\pi([M]\otimes[N])>$, we only need to prove that $<r'_{k,l}[P],[M]\otimes[N]>=<[P],i_{k,l}([M]\otimes[N])>$, for all $[P] \in K_0(A_{k+l})$, $[M] \in K_0(A_k)$ and $[N] \in G_0(A_l)$. We can simplify this by proving that $< r'_{k,l}[P], [M] \otimes [N] >=< [P], i_{k,l}([M] \otimes [N]) >$ holds when P is an indecomposable projective A_{k+l}-module. We know that each indecomposable projective module corresponding to a primitive idempotent. We establish this identity by the following lemma and condition (4).

Lemma 3.2. [**15**] *Let B be a finite-dimensional algebra over field K, M a left B-module and e a primitive idempotent. Then $Hom_B(Be, M) \cong eM$ as vector spaces.*

$< \mu'(1), [M] >= \epsilon([M])$ and $< [P], \mu(1) >= \epsilon([P])$ follow from the definitions of μ and μ'. □

To get the duality between $G_0(A)$ and $K_0(A)$ these identities are not enough. We should verify that their bases are orthonormal to each other. Let $\{V_1, \cdots, V_s\}$ be a complete list of nonisomorphic simple A_n-modules. Then the set of their projective covers $\{P_1, \cdots, P_s\}$ is a complete list of nonisomorphic indecomposable projective A_n-modules. Then

Proposition 3.4. $< [P_i], [V_j] >= \delta_{i,j}$ *for* $1 \leq i, j \leq s$.

Proof. This follows from the property of simple modules and the Schur's Lemma. □

Theorem 3.3 (Main Result 1). *$(G_0(A), \pi, \mu, \Delta, \epsilon)$ and $(K_0(A), \pi', \mu', \Delta', \epsilon')$ are both graded connected bialgebras. Hence both are graded Hopf algebras. And they are dual to each other with respect to the pairing.*

Proof. This follows directly from Theorem 3.1 and Propositions 3.3 and 3.4. □

4. Grothendieck groups of a tower of algebras (Not preserving unities)

In [**1**], N. Bergeron, C. Holhweg, M. Rosas, and M. Zabrocki consider a semi-tower of algebras with ρ not preserving unities. If we only weaken the condition of ρ and modify the definitions of inductions and restrictions can we get a similar result? In this section, we will do this job. The structure of this section is parallel to Section 3.

Let $A = \bigoplus_{n\geq 0} A_n$, we call it *a tower of algebras* over field $K = \mathbb{C}$ if the following conditions are satisfied:
(1) A_n is a finite-dimensional algebra with unit, for each n. $A_0 \cong K$.
(2) There is an external graded multiplication $\rho_{m,n} : A_m \otimes A_n \to A_{m+n}$ such that

(a) $\rho_{m,n}$ is an injective homomorphism of algebras, for all m and n (but $\rho_{m,n}(1_m \otimes 1_n) \neq 1_{m+n}$ for some or all m and n);

(b) ρ is associative, that is, $\rho_{l+m,n} \cdot (\rho_{l,m} \otimes 1_n) = \rho_{l,m+n} \cdot (1_l \otimes \rho_{m,n}) := \rho_{l,m,n}$, for all l, m, n.

(3) A_{n+m} is a two-sided projective $A_n \otimes A_m$-module by the action defined to be $a \cdot (b \otimes c) = a\rho_{m,n}(b \otimes c)$ and $(b \otimes c) \cdot a = \rho_{m,n}(b \otimes c)a$, for $a \in A_{m+n}$, $b \in A_m$ and $c \in A_n$.
(4) For every primitive idempotent g in A_{m+n}, $A_{m+n}g \cong \bigoplus(A_m \otimes A_n)(e \otimes f)$ as (left) $A_m \otimes A_n$-modules if and only if $gA_{m+n} \cong \bigoplus(e \otimes f)(A_m \otimes A_n)$ as (right) $A_m \otimes A_n$-modules for the same indexing of idempotents $(e \otimes f)$'s in $A_m \otimes A_n$.
(5) The following equalities hold

$$\begin{array}{rl} & [\mathrm{Res}^{A_{m+n}}_{A_k \otimes A_{m+n-k}} \mathrm{Ind}^{A_{m+n}}_{A_m \otimes A_n}(M \otimes N)] \\ = & \sum_{t+s=k} [\widetilde{\mathrm{Ind}}^{A_k \otimes A_{m+n-k}}_{A_t \otimes A_{m-t} \otimes A_s \otimes A_{n-s}} (\mathrm{Res}^{A_m}_{A_t \otimes A_{m-t}} M \otimes \mathrm{Res}^{A_n}_{A_s \otimes A_{n-s}} N)] \end{array}$$

for all $0 < k < m + n$, M an A_m-module and N an A_n-module, and

$$\begin{array}{rl} & [\mathrm{Res}^{A_{m+n}}_{A_k \otimes A_{m+n-k}} \mathrm{Ind}^{A_{m+n}}_{A_m \otimes A_n}(P \otimes Q)] \\ = & \sum_{t+s=k} [\widetilde{\mathrm{Ind}}^{A_k \otimes A_{m+n-k}}_{A_t \otimes A_{m-t} \otimes A_s \otimes A_{n-s}} (\mathrm{Res}^{A_m}_{A_t \otimes A_{m-t}} P \otimes \mathrm{Res}^{A_n}_{A_s \otimes A_{n-s}} Q)] \end{array}$$

for all $0 < k < m + n$, P a projective A_m-module and Q a projective A_n-module. We will explain the notations later.

The definition of inductions on $G_0(A)$ is

$$\begin{array}{rcll} i_{m,n} & : & G_0(A_m) \bigotimes_{\mathbb{Z}} G_0(A_n) & \to & G_0(A_{m+n}) \\ & & [M] \otimes [N] & \mapsto & [\mathrm{Ind}^{A_{m+n}}_{A_m \otimes A_n} M \otimes N], \end{array}$$

which is same as the one in Section 3. Let $k=t+s$, define the twisted induction

$$\begin{aligned} & \widetilde{\mathrm{Ind}}^{A_k\otimes A_{m+n-k}}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}(M_1\otimes M_2)\otimes(N_1\otimes N_2) \\ = \; & (A_k\otimes A_{m+n-k})\widetilde{\bigotimes}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}((M_1\otimes M_2)\otimes(N_1\otimes N_2)), \end{aligned}$$

which is also same as the one in Section 3. Define the restrictions r on $G_0(A)$ by

$$\begin{array}{rcccl} r_{k,l} & : & G_0(A_n) & \to & G_0(A_k)\bigotimes_{\mathbb{Z}} G_0(A_l) \text{ with } k+l=n \\ & & [N] & \mapsto & [\mathrm{Res}^{A_n}_{A_k\otimes A_l}N], \end{array}$$

where $\mathrm{Res}^{A_n}_{A_k\otimes A_l}N=\{u\in N\mid \rho_{k,l}(1_k\otimes 1_l)u=u\}\subseteq N$ is an $A_k\otimes A_l$-module by the action defined to be $(b\otimes c)\cdot u=\rho_{k,l}(b\otimes c)u$, for $u\in \mathrm{Res}^{A_n}_{A_k\otimes A_l}N$, $b\in A_k$ and $c\in A_l$. When ρ preserving unities, we have $\mathrm{Res}^{A_n}_{A_k\otimes A_l}N=N\cong \mathrm{Hom}_{A_n}(A_n,N)$. This coincides with the restrictions r in Section 3.

Proposition 4.1. *i and r are well defined.*

Proof. For i, it follows from Proposition 3.1 since they have the same definition.

For r, we know $\mathrm{Res}^{A_n}_{A_k\otimes A_l}N=\rho_{k,l}(1_k\otimes 1_l)N$ and $\rho_{k,l}(1_k\otimes 1_l)$ is an idempotent in A_n, hence $N=\rho_{k,l}(1_k\otimes 1_l)N\oplus(1-\rho_{k,l}(1_k\otimes 1_l))N$. From the properties of short exact sequence and homomorphisms of modules which can be written as a direct sum, one can get that all r are well defined. □

As in Section 3, we define $\pi,\ \Delta,\ \mu$ and ϵ by the inductions i and restrictions r on $G_0(A)$. Later we will prove that $G_0(A)$ is a graded bialgebra with these operators.

Now we define inductions and restrictions on $K_0(A)$ as follows:

$$\begin{array}{rcccl} i'_{m,n} & : & K_0(A_m)\bigotimes_{\mathbb{Z}} K_0(A_n) & \to & K_0(A_{m+n}) \\ & & [P]\otimes[Q] & \mapsto & [\mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}P\otimes Q], \end{array}$$

where $P=A_me_m,\ Q=A_ne_n$ for some primitive idempotents $e_m\in A_m$ and $e_n\in A_n$, and

$$\begin{aligned} & \mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}P\otimes Q \\ = \; & \mathrm{Ind}^{A_{m+n}}_{A_m\otimes A_n}A_me_m\otimes A_ne_n \\ := \; & A_{m+n}\rho_{m,n}(e_m\otimes e_n), \end{aligned}$$

which is a projective $A_m\otimes A_n$-module. Here i' is only defined on the basis of $K_0(A_m)\otimes K_0(A_n)$. To get induction we only need i' to satisfy linearity. i.e.,

$$\begin{aligned} & i'((a[P']+b[P''])\otimes(c[Q']+d[Q''])) \\ = \; & aci'([P']\otimes[Q'])+adi'([P']\otimes[Q''])+bci'([P'']\otimes[Q'])+bdi'([P'']\otimes[Q'']), \end{aligned}$$

where $a,b,c,d\in\mathbb{Z}, P',P''\in K_0(A_m)$ and $Q',Q''\in K_0(A_n)$ are indecomposable. Hence i' is well defined. And when ρ preserving unities, this i' coincides with the inductions i' in Section 3.

Let $k=t+s$, define the twisted induction

$$\begin{aligned} & \widetilde{\mathrm{Ind}}^{A_k\otimes A_{m+n-k}}_{A_t\otimes A_{m-t}\otimes A_s\otimes A_{n-s}}(A_te_1\otimes A_{m-t}e_2)\otimes(A_sf_1\otimes A_{n-s}f_2) \\ := \; & A_k\rho_{t,s}(e_1\otimes f_1)\otimes A_{m+n-k}\rho_{m-t,n-s}(e_2\otimes f_2), \end{aligned}$$

where $e_1,\ e_2,\ f_1$ and f_2 are primitive idempotents in $A_t,\ A_{m-t},\ A_s$ and A_{n-s} respectively.

Set

$$\begin{array}{rcccl} r'_{k,l} & : & K_0(A_n) & \to & K_0(A_k)\bigotimes_{\mathbb{Z}} K_0(A_l) \text{ with } k+l=n \\ & & [R] & \mapsto & [\mathrm{Res}^{A_n}_{A_k\otimes A_l}R], \end{array}$$

where $\mathrm{Res}^{A_n}_{A_k\otimes A_l}R=\{x\in R\mid \rho_{k,l}(1_k\otimes 1_l)x=x\}$ as a left projective $A_k\otimes A_l$-module.

Proposition 4.2. *r' is well defined.*

Proof. To show r' well defined, there are three steps. Let R be a projective A_n-module.

1. $\rho_{k,l}(1_k\otimes 1_l)$ is an idempotent and $\mathrm{Res}^{A_n}_{A_k\otimes A_l}R=\rho_{k,l}(1_k\otimes 1_l)R$.

2. $\mathrm{Res}^{A_n}_{A_k\otimes A_l}R$ is an $A_k\otimes A_l$-module.

3. $\mathrm{Res}^{A_n}_{A_k\otimes A_l}R$ is a projective $A_k\otimes A_l$-module. Here we verify that $\mathrm{Res}^{A_n}_{A_k\otimes A_l}R$ is a summand of R by lemma 3.2, step 1 and the property of idempotents.

□

As before, using the definitions of inductions i' and restrictions r' we construct π', Δ', μ' and ϵ' on $K_0(A)$. Later we will prove that $K_0(A)$ with these operators is a graded bialgebra.

THEOREM 4.1. *(i) π is associative and $\big(G_0(A),\ \pi,\ \mu\big)$ is an algebra. So is $\big(K_0(A),\ \pi',\ \mu'\big)$.*
(ii) Δ is coassociative and $\big(G_0(A),\ \Delta,\ \epsilon\big)$ is a coalgebra. So is $\big(K_0(A),\ \Delta',\ \epsilon'\big)$.
(iii) Δ and ϵ are algebra homomorphisms and $G_0(A)$ is a graded bialgebra. Hence $G_0(A)$ is a graded Hopf algebra. So is $K_0(A)$.

PROOF. (i) For $G_0(A)$, it holds from Theorem 3.1(i).
For the associativity of π' in $K_0(A)$, we need to show $i'_{l+m,n}\cdot(i'_{l,m}\otimes 1_n)=i'_{l,m+n}\cdot(1_l\otimes i'_{m,n})$. One can get it by the associativity of ρ and the definition of i'.
(ii) We only need to show the coassociativity of Δ, that is, $(r_{l,m}\otimes 1)\cdot r_{l+m,n}=(1\otimes r_{m,n})\cdot r_{l,m+n}$. This follows from the associativity of ρ and the definition of r. Similarly for $\big(K_0(A),\ \Delta',\ \epsilon'\big)$.
(iii) From condition (5), one can prove that $G_0(A)$ is a graded bialgebra by the definition of compatibility of algebra and coalgebra structures. Do the similar work to $K_0(A)$. From condition (1), we know that $G_0(A)$ is a graded connected bialgebra. Hence a graded Hopf algebra. Similarly for $K_0(A)$. $\square$

Define a pairing $<,>: K_0(A)\times G_0(A)\to\mathbb{Z}$ by

$$<[P],[M]>=\begin{cases}\dim_K\big(\mathrm{Hom}_{A_n}(P,M)\big) & \text{if } [P]\in K_0(A_n) \text{ and } [M]\in G_0(A_n)\\ 0 & \text{otherwise.}\end{cases}$$

and with the same notation $<,>: (K_0(A)\otimes K_0(A))\times(G_0(A)\otimes G_0(A))\to\mathbb{Z}$ by

$$<[P]\otimes[Q],[M]\otimes[N]>=\begin{cases}\dim_K\big(\mathrm{Hom}_{A_k\otimes A_l}(P\otimes Q,M\otimes N)\big) & \text{if } [P]\otimes[Q]\in K_0(A_k)\otimes K_0(A_l)\\ & \ \text{and } [M]\otimes[N]\in G_0(A_k)\otimes G_0(A_l)\\ 0 & \text{otherwise.}\end{cases}$$

PROPOSITION 4.3. *$<,>$ is a well-defined bilinear pairing on $K_0(A)\times G_0(A)$ satisfying the following identities*

$$\begin{aligned}<[P]\otimes[Q],[M]\otimes[N]> &= <[P],[M]><[Q],[N]>\\ <\pi'([P]\otimes[Q]),[M]> &= <[P]\otimes[Q],\Delta[M]>\\ <\Delta'[P],[M]\otimes[N]> &= <[P],\pi([M]\otimes[N])>\\ <\mu'(1),[M]> &= \epsilon([M])\\ <[P],\mu(1)> &= \epsilon'([P]).\end{aligned}$$

PROOF. The bilinearity and the first identity are same as Proposition 3.3.
To show $<\pi'([P]\otimes[Q]),[M]>=<[P]\otimes[Q],\Delta[M]>$, we only need to prove that $<i'_{k,l}([P]\otimes[Q]),[M]>=<[P]\otimes[Q],r_{k,l}[M]>$, for all $[P]\in K_0(A_k)$, $[Q]\in K_0(A_l)$ and $[M]\in G_0(A_{k+l})$. Without loss of generality, let $P=A_ke_k$ and $Q=A_le_l$ for some primitive idempotents $e_k\in A_k$ and $e_l\in A_l$. Using Lemma 3.2 one can get it straightforwardly.
To show $<\Delta'[P],[M]\otimes[N]>=<[P],\pi([M]\otimes[N])>$, we only need to prove that $<r'_{k,l}[P],[M]\otimes[N]>=<[P],i_{k,l}([M]\otimes[N])>$, for all $[P]\in K_0(A_{k+l})$, $[M]\in K_0(A_k)$ and $[N]\in G_0(A_l)$. One can get it from Lemma 3.2 and condition (4).
$<\mu'(1),[M]>=\epsilon([M])$ and $<[P],\mu(1)>=\epsilon([P])$ follow from the definitions of μ and μ'. $\square$

THEOREM 4.2 (Main Result 2). *$(G_0(A),\pi,\mu,\Delta,\epsilon)$ and $(K_0(A),\pi',\mu',\Delta',\epsilon')$ are both graded connected bialgebras. Hence both are graded Hopf algebras. And they are dual to each other with respect to the pairing.*

PROOF. This follows directly from Theorem 4.1 and Propositions 4.3 and 3.4. $\square$

5. Some examples

In this section, we will verify that $\bigoplus_{n\geq 0}S_n$, $\bigoplus_{n\geq 0}H_n(0)$ and $\bigoplus_{n\geq 0}HCl_n(0)$ satisfy all the axioms listed in Section 3. They are towers of algebras and we already know that their Grothendieck groups are dual Hopf algebras, respectively. And we will discuss some graded algebras which don't satisfy some axiom are not towers of algebras. Consequently, their Grothendieck groups are not dual Hopf algebras.

Let $A=\bigoplus_{n\geq 0}A_n$ with $A_n=\mathbb{C}S_n$. Here

$$\rho_{m,n}:\mathbb{C}S_m\otimes\mathbb{C}S_n\to\mathbb{C}S_{m+n}$$

is defined to be $\rho_{m,n}(\sigma \otimes \tau) = \sigma(1)\sigma(2)\cdots\sigma(m)(\tau(1)+m)(\tau(2)+m)\cdots(\tau(n)+m)$ if we use the one line notation of permutations, where $\sigma \in S_m$ and $\tau \in S_n$. It is easy to check that ρ's preserve unities and have the associativity. Since $\mathbb{C}S_n$ is a semi-simple algebra, we know that $\mathbb{C}S_{m+n}$ is a two-sided projective $\mathbb{C}S_m \otimes \mathbb{C}S_n$-module and satisfies condition (4). Condition (5) is just the Mackey's Formula [**16**]. Hence $A = \mathbb{C}S_n$ is a tower of algebra and the Grothendieck group $G_0(A) = K_0(A)$ is a self-dual graded Hopf algebra.

For $\bigoplus_{n\geq 0} H_n(0)$ of 0-Hecke algebras, the ρ's are defined by $\rho_{m,n}(T_i \otimes 1) = T_i$ and $\rho_{m,n}(1 \otimes T_j) = T_{j+m}$, where T_i's and T_j's are the generators of $H_m(0)$ and $H_n(0)$, $1 \leq i \leq m-1$ and $1 \leq i \leq n-1$. For $\bigoplus_{n\geq 0} HCl_n(0)$ of 0-Hecke-Clifford algebras, the ρ's are defined by $\rho_{m,n}(T_i \otimes 1) = T_i$, $\rho_{m,n}(C_k \otimes 1) = C_k$, $\rho_{m,n}(1 \otimes T_j) = T_{j+m}$ and $\rho_{m,n}(1 \otimes C_l) = C_{l+m}$, where T_i's with C_k's and T_j's with C_l's are the generators of $HCl_m(0)$ and $HCl_n(0)$ respectively, $1 \leq i \leq m-1$, $1 \leq k \leq m$, $1 \leq i \leq n-1$ and $1 \leq l \leq n$. We will also check that these two satisfy all the axioms listed in section 3. In the introduction we have mentioned that their Grothendieck groups are dual graded Hopf algebras.

Now we describe an example not satisfying condition (5). In [**1**], $(\mathbf{\Pi}, \wedge) = \bigoplus_{n\geq 0}(\mathbb{C}\Pi_n, \wedge)$ of the partition lattice algebras with

$$\rho_{m,n} : (\mathbb{C}\Pi_m, \wedge) \otimes (\mathbb{C}\Pi_n, \wedge) \to (\mathbb{C}\Pi_{m+n}, \wedge)$$

defined by $\rho_{m,n}(A \otimes B) = A|B$, where $A|B = \{A_1, A_2, \ldots, A_{l(A)}, B_1 + m, B_2 + m, \ldots, B_{l(B)} + m\}$. Here ρ's do not preserve unities. Although $\left(\bigoplus_{n\geq 0}(\mathbb{C}\Pi_n, \wedge), \{\rho_{m,n}\}\right)$ satisfies conditions (1)-(4) in section 4, there is no similar Mackey's fomula (5), i.e., the operations of induction and restriction are not compatible as a bialgebra. Hence the Grothendieck groups $G_0(\mathbf{\Pi}, \wedge)$ and $K_0(\mathbf{\Pi}, \wedge)$ do not have the Hopf algebra structure although the operation of restriction on $G_0(\mathbf{\Pi}, \wedge)$ is dual to the operation of induction on $K_0(\mathbf{\Pi}, \wedge)$ and the induction on $G_0(\mathbf{\Pi}, \wedge)$ is dual as graded operations to restriction on $K_0(\mathbf{\Pi}, \wedge)$.

If one consider a direct sum of algebras that does not satisfy conditions (3) then the inductions and restrictions may not be well defined. Hence we can not construct the multiplication and comultiplication. Consequently, its Grothendieck groups are not Hopf algebras. If it does not satisfy condition (4), then its Grothendieck groups are graded Hopf algebras respectively but not necessarily dual to each other.

References

[1] N. Bergeron; C. Hohlweg; M. Rosas; M. Zabrocki *Grothendick bialgebra, Partition lattices and symmetric functions in noncommutative variables*, arXiv:math.CO/0506360 v1 17 Jun 2005

[2] Bergeron, Nantel; Hivert, Florent; Thibon, Jean-Yves *The peak algebra and the Hecke-Clifford algebras at* $q = 0$. J. Combin. Theory Ser. A 107 (2004), no. 1, 1–19.

[3] Bergeron, Nantel; Mykytiuk, Stefan; Sottile, Frank; van Willigenburg, Stephanie *Noncommutative Pieri operators on posets*. In memory of Gian-Carlo Rota. J. Combin. Theory Ser. A 91 (2000), no. 1-2, 84–110.

[4] M. Aguiar; N. Bergeron; F. Sottile *Combinatorial Hopf algebras and Generalized Dehn-Sommervile Relations*, arXiv:math.CO/0310016 v1 1 Oct 2003

[5] Childs, Lindsay N.; Greither, Cornelius; Moss, David J.; Sauerberg, Jim; Zimmermann *Hopf algebras, polynomial formal groups, and Raynaud orders*. Mem. Amer. Math. Soc. 136 (1998), no. 651, viii+118 pp.

[6] Geissinger, Ladnor *Hopf algebras of symmetric functions and class functions*, Combinatoire et Combinatoire et représentation du groupe symétrique (Actes Table Ronde C.N.R.S., Univ. Louis-Pasteur Strasbourg, Strasbourg, 1976), pp. 168–181. Lecture Notes in Math., Vol. 579, Springer, Berlin, 1977.

[7] Zelevinsky, Andrey V. *Representations of finite classical groups. A Hopf algebra approach*, Lecture Notes in Mathematics, 869. Springer-Verlag, Berlin-New York, 1981.

[8] Curtis, Charles W.; Reiner, Irving *Methods of representation theory. Vol. I. With applications to finite groups and orders*, Reprint of the 1981 original. Wiley Classics Library. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1990.

[9] Gelfand, Israel M.; Krob, Daniel; Lascoux, Alain; Leclerc, Bernard; Retakh, Vladimir S.; Thibon, Jean-Yves *Noncommutative symmetric functions* , Adv. Math. 112 (1995), no. 2, 218–348.

[10] Krob, Daniel; Thibon, Jean-Yves *Noncommutative symmetric functions. IV. Quantum linear groups and Hecke algebras at* $q = 0$, J. Algebraic Combin. 6 (1997), no. 4, 339–376.

[11] Macdonald, I. G. *Symmetric functions and Hall polynomials*, Second edition. With contributions by A. Zelevinsky. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1995. x+475 pp. ISBN: 0-19-853489-2

[12] Malvenuto, Clauda; Reutenauer, Christophe *Duality between quasi-symmetric functions and the solomon descent algebra*, J. Algebra 177 (1995), no. 3, 967–982.

[13] M. Auslander; I. Reiten; S. O.Smalø, *Representation theory of Artinian algebras*, Cambridge University Press, 1995

[14] Moss E. Sweedler *Hopf algebras*, W. A. Benjamin, Inc., 1969
[15] H. Nagao; Y. Tsushima *Representations of finite groups*, Academic Press, Inc., 1989
[16] Serge Lang *Algebra*, 3rd ed., Addison-Wesley Pub. Co., c1993.
[17] J. W. Milnor; J. C. Moore *On the structure of Hopf algebras*, Ann. of Math. 81 (1965), no. 2, 211–264.

Department of Mathematics and Statistics, York University, Canada, M3J 1P3
E-mail address: lihuilan@mathstat.yorku.ca

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Classifying ascents and descents with specified equivalences mod k

Jeffrey Liese

Abstract. Given a permutation τ of length j, we say that a permutation σ has a τ-match starting at position i, if the elements in position $i, i+1, \ldots, i+j-1$ in σ have the same relative order as the elements of τ. If Υ is set of permutations of length j, then we say that a permutation σ has an Υ-match starting at position j if it has a τ-match at position j for some $\tau \in \Upsilon$. A number of recent papers have studied the distribution of τ-matches and Υ-matches in permutations. In this paper, we consider a more refined pattern matching condition where we take into account conditions involving the equivalence classes of the elements mod k for some integer $k \geq 2$. In general, when one includes parity conditions or conditions involving equivalence mod k, then the problem of counting the number of pattern matchings becomes more complicated. In this paper, we prove explicit formulas for the number of permutations of n which have s τ-equivalence mod k matches when τ is of length 2. We also show that similar formulas hold for Υ-equivalence mod k matches for certain subsets of permutations of length two.

Résumé. Étant donnée une permutation τ de longueur j, on dit qu'une permutation σ a un τ-motif débutant en position i si les éléments en position $i, i+1, \ldots, i+j-1$ de σ ont le même ordre relatif que les éléments de τ. Si Υ est un ensemble de permutations de longueur j, alors on dit que σ a un Υ-motif en position i si σ a un τ-motif en position i pour une permutation τ de Υ. Plusieurs travaux récents ont portés sur la distribution des τ-occurrences et Υ-occurrences dans les permutations. Dans ce travail, nous étudions un raffinement de la notion de motif prenant en compte de conditions basée sur les classes d'équivalences des éléments mod k. De manière générale, lorsque l'on prend en compte la parité ou l'équivalence mod k, le problème de l'énumération du nombre d'occurrences d'un motif devient plus compliqué. Nous démontrons une formule explicite pour le nombre de permutations de n qui ont s τ-motifs équivalents mod k quand τ est de longueur 2. Nous montrons aussi que des formules similaires existent pour les Υ-motifs quand Υ est limité à certains sous-ensembles de permutations de longueur 2.

1. Introduction

Given any sequence $\sigma = \sigma_1 \cdots \sigma_n$ of distinct integers, we let $red(\sigma)$ be the permutation that results by replacing the i-th largest integer that appears in the sequence σ by i. For example, if $\sigma = 2\ 7\ 5\ 4$, then $red(\sigma) = 1\ 4\ 3\ 2$. Given a permutation τ in the symmetric group S_j, we define a permutation $\sigma = \sigma_1 \cdots \sigma_n \in S_n$ to have a τ-match at place i provided $red(\sigma_i \cdots \sigma_{i+j-1}) = \tau$. Let τ-$mch(\sigma)$ be the number of τ-matches in the permutation σ. To prevent confusion, we note that a permutation not having a τ-match is different than a permutation being τ-avoiding. A permutation is called τ-avoiding if there are no indices $i_1 < \cdots < i_j$ such that $red[\sigma_{i_1} \cdots \sigma_{i_j}] = \tau$. For example, if $\tau = 2\ 1\ 4\ 3$, then the permutation $3\ 2\ 1\ 4\ 6\ 5$ does not have a τ-match but it does not avoid τ since $red[2\ 1\ 6\ 5] = \tau$.

In the case where $|\tau| = 2$, then τ-$mch(\sigma)$ reduces to familiar permutation statistics. That is, if $\sigma = \sigma_1 \cdots \sigma_n \in S_n$, let $Des(\sigma) = \{i : \sigma_i > \sigma_{i+1}\}$ and $Rise(\sigma) = \{i : \sigma_i < \sigma_{i+1}\}$. Then it is easy to see that $(2\ 1)$-$mch(\sigma) = des(\sigma) = |Des(\sigma)|$ and $(1\ 2)$-$mch(\sigma) = rise(\sigma) = |Rise(\sigma)|$.

A number of recent publications have analyzed the distribution of τ-matches in permutations. See, for example, [**EN03, Kit03, Kit**]. A number of interesting results have been proved. For example, let τ-nlap(σ)

2000 *Mathematics Subject Classification.* Primary 17B10; Secondary 05E10.

Key words and phrases. ascents, descents, parity, pattern matching, permutations.

This paper adapted from a section of the Doctoral thesis of Jeffrey Liese with thanks to the direction and assistance provided by the thesis advisor, Jeffrey Remmel.

be the maximum number of non-overlapping τ-matches in σ where two τ-matches are said to overlap if they contain any of the same integers. Then Kitaev [**Kit03, Kit**] proved the following.

THEOREM 1.1.

$$\sum_{n\geq 0}\frac{t^n}{n!}\sum_{\sigma\in S_n} x^{\tau\text{-nlap}(\sigma)} = \frac{A(t)}{(1-x)+x(1-t)A(t)} \tag{1.1}$$

where $A(t)=\sum_{n\geq 0}\frac{t^n}{n!}|\{\sigma\in S_n : \tau\text{-mch}(\sigma)=0\}|$.

In other words, if the exponential generating function for the number of permutations in S_n without any τ-matches is known, then so is the exponential generating function for the entire distribution of the statistic τ-nlap.

In this paper, we consider a more refined pattern matching condition where we take into account conditions involving equivalence mod k for some integer $k\geq 2$. That is, suppose we fix $k\geq 2$ and we are given some sequence of distinct integers $\tau=\tau_1\cdots\tau_j$. Then we say that a permutation $\sigma=\sigma_1\cdots\sigma_n\in S_n$ has a τ-k-equivalence match at place i provided $red(\sigma_i\cdots\sigma_{i+j-1})=red(\tau)$ and for all $s\in\{0,\dots,j-1\}$, $\sigma_{i+s}=\tau_{1+s}$ mod k. For example, if $\tau=1\ 2$ and $\sigma=5\ 1\ 7\ 4\ 3\ 6\ 8\ 2$, then σ has τ-matches starting at positions 2, 5, and 6. However, if $k=2$, then only the τ-match starting at position 5 is a τ-2-equivalence match. Later, it will be explained that the τ-match starting a position 2 is a (1 3)-2-equivalence match and the τ-match starting a position 6 is a (2 4)-2-equivalence match. Let τ-k-$emch(\sigma)$ be the number of τ-k-equivalence matches in the permutation σ. Let τ-k-$enlap(\sigma)$ be the maximum number of non-overlapping τ-k-equivalence matches in σ where two τ-matches are said to overlap if they contain any of the same integers.

More generally, if Υ is a set of sequences of distinct integers of length j, then we say that a permutation $\sigma=\sigma_1\cdots\sigma_n\in S_n$ has a Υ-k-equivalence match at place i provided there is a $\tau\in\Upsilon$ such that $red(\sigma_i\cdots\sigma_{i+j-1})=red(\tau)$ and for all $s\in\{0,\dots,j-1\}$, $\sigma_{i+s}=\tau_{1+s}\mod k$. Let Υ-k-$emch(\sigma)$ be the number of Υ-k-equivalence matches in the permutation σ and Υ-k-$enlap(\sigma)$ be the maximum number of non-overlapping Υ-k-equivalence matches in σ.

In this paper, we shall begin the study of the polynomials

$$T_{\tau,k,n}(x) = \sum_{\sigma\in S_n} x^{\tau\text{-}k\text{-}emch(\sigma)} = \sum_{s=0}^{n} T^s_{\tau,k,n}x^s \text{ and} \tag{1.2}$$

$$U_{\Upsilon,k,n}(x) = \sum_{\sigma\in S_n} x^{\Upsilon\text{-}k\text{-}emch(\sigma)} = \sum_{s=0}^{n} U^s_{\Upsilon,k,n}x^s. \tag{1.3}$$

In particular, we shall focus on certain special cases of these polynomials where we consider only patterns of length 2. That is, fix $k\geq 2$ and let A_k equal the set of all sequences $(a\ b)$ such that $1\leq a<b\leq 2k$ where there is no lexicographically smaller sequence $x\ y$ having the property that $x\equiv a\mod k$ and $y\equiv b\mod k$. For example,

$$A_4=\{1\ 2, 1\ 3, 1\ 4, 1\ 5, 2\ 3, 2\ 4, 2\ 5, 2\ 6, 3\ 4, 3\ 5, 3\ 6, 3\ 7, 4\ 5, 4\ 6, 4\ 7, 4\ 8\}.$$

Let $D_k=\{b\ a : a\ b\in A_k\}$ and $E_k=A_k\cup D_k$. Thus E_k consists of all k-equivalence patterns of length 2 that we could possibly consider. Note that if $\Upsilon=A_k$, then Υ-k-$emch(\sigma)=rise(\sigma)$ and if $\Upsilon=D_k$, then Υ-k-$emch(\sigma)=des(\sigma)$.

Our goal is to give explicit formulas for the coefficients of $T^s_{\tau,k,n}$ and $U^s_{\Upsilon,k,n}$. First we shall show that we can use inclusion-exclusion to find a formula for $U^s_{\Upsilon,k,n}$ for any $\Upsilon\subset E_k$ in terms of certain rook numbers of a sequences of boards associated with Υ. While this approach is straightforward, it is unsatisfactory since it reduces the computation of $U^s_{\Upsilon,k,n}$ to another difficult problem, namely, computing rook numbers for general boards. However, we can give two other more direct formulas for the coefficients $T^s_{\tau,k,n}$ where $\tau\in E_k$. For example, in the case where $\tau = (1\ k)$, our results will imply that for all $0 \le s \le n$ and for all $0 \le j \le k-1$,

$$
\begin{aligned}
(1.4)\ T^s_{(1\ k),k,kn+j} &= \sum_{r=s}^{n} (-1)^{r-s}(kn+j-r)!\binom{r}{s} S_{n+1,n+1-r} \\
&= ((k-1)n+j)!\sum_{r=0}^{s}(-1)^{s-r}((k-1)n+j+r)^n\binom{(k-1)n+j+r}{r}\binom{kn+j+1}{s-r} \\
&= ((k-1)n+j)!\sum_{r=0}^{n-s}(-1)^{n-s-r}(1+r)^n\binom{(k-1)n+j+r}{r}\binom{kn+j+1}{n-s-r}
\end{aligned}
$$

where $S_{n,k}$ is the Stirling number of the second kind, i.e., $S_{n,k}$ is the number of partitions of an n-set into k parts. These formulas lead to interesting identities in their own right. For example, we see that for all $k \ge 2$, $0 \le s \le n$ and $0 \le j \le k-1$,

$$
\sum_{r=0}^{s}(-1)^{s-r}((k-1)n+j+r)^n\binom{(k-1)n+j+r}{r}\binom{kn+j+1}{s-r} =
\sum_{r=0}^{n-s}(-1)^{n-s-r}(1+r)^n\binom{(k-1)n+j+r}{r}\binom{kn+j+1}{n-s-r}.
$$

The general problem of finding explicit expressions for the coefficients $U^s_{\Upsilon,k,n}$ for arbitrary Υ is open. However, Kitaev and Remmel [**KR05, KR06**] have developed formulas for $U^s_{\Upsilon,k,n}$ in certain other special cases. In particular, Kitaev and Remmel studied permutation statistics which classified the descents of a permutation according to whether either the first element or the second element of a descent pair is equivalent to $0 \mod k$. In our language, they computed explicit formulas for $U^s_{\Upsilon,k,n}$ where either $\Upsilon = \{b\ a : (b\ a) \in D_k\ \&\ b \equiv 0 \mod k\}$ or $\Upsilon = \{b\ a : (b\ a) \in D_k\ \&\ a \equiv 0 \mod k\}$. In this paper, we shall generalize some of their results by deriving explicit formulas for $U^s_{\Upsilon,k,n}$ in the special cases where Υ is a subset of the form $\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$ where for all i,j $y_i \equiv y_j \mod k$ and either $\Upsilon \subseteq A_k$ or $\Upsilon \subseteq D_k$.

The outline of this paper is as follows. In section 2, we shall discuss some of the previous results of Kitaev and Remmel [**KR05, KR06**] and give some examples of the polynomials $T_{\tau,k,n}(x)$. In section 3, we will show how to one can use inclusion-exclusion to derive an $U_{\Upsilon,k,n}(x)$ in terms of certain rook numbers. In section 4, we shall prove formulas in the case where Υ consists of a sequences of pairs $\{(x_1,y_1),\ldots,(x_t,y_t)\} \subseteq A_k$ such that for all i and j, $y_i = y_j \mod k$. Using the bijection which sends each permutation $\sigma = \sigma_1 \cdots \sigma_n$ to its reverse, $\sigma^r = \sigma_n \cdots, \sigma_1$, one can show that the same formulas hold for $\Upsilon^r = \{(y_1,x_1),\ldots,(y_t,x_t)\} \subseteq D_k$. We shall also see that the identities that result by equating the different formulas for any given coefficient are interesting in their own right. Then, we shall make a few comments about the problem of finding $U_{\Upsilon,k,n}(x)$ for arbitrary Υ.

2. Previous results and examples

In this section, we shall state some previous results and give some examples of the polynomials $T_{\tau,k,n}(x)$ and $U_{\Upsilon,k,n}(x)$. As mentioned in the introduction, Kitaev and Remmel [**KR05, KR06**], found explicit formulas for the coefficients $U^s_{\Upsilon,k,n}$ in certain special cases. In particular, they studied descents according to the equivalence class $\mod k$ of either the first or second element in a descent pair. That is, for any set $X \subseteq \{0,1,2,\ldots\}$, define

- $\overleftarrow{Des}_X(\sigma) = \{i : \sigma_i > \sigma_{i+1}\ \&\ \sigma_i \in X\}$ and $\overleftarrow{des}_X(\sigma) = |\overleftarrow{Des}_X(\sigma)|$
- $\overrightarrow{Des}_X(\sigma) = \{i : \sigma_i > \sigma_{i+1}\ \&\ \sigma_{i+1} \in X\}$ and $\overrightarrow{des}_X(\sigma) = |\overrightarrow{Des}_X(\sigma)|$

In [**KR05**], Kitaev and Remmel studied the following polynomials.

(1) $R_n(x) = \sum_{\sigma\in S_n} x^{\overleftarrow{des}_E(\sigma)} = \sum_{k=0}^n R_{k,n}x^k$,
(2) $P_n(x,z) = \sum_{\sigma\in S_n} x^{\overrightarrow{des}_E(\sigma)} z^{\chi(\sigma_1\in E)} = \sum_{k=0}^n \sum_{j=0}^1 P_{j,k,n} z^j x^k$
(3) $M_n(x) = \sum_{\sigma\in S_n} x^{\overleftarrow{des}_O(\sigma)} = \sum_{k=0}^n M_{k,n}x^k$, and
(4) $Q_n(x,z) = \sum_{\sigma\in S_n} x^{\overrightarrow{des}_O(\sigma)} z^{\chi(\sigma_1\in O)} = \sum_{k=0}^n \sum_{j=0}^1 Q_{j,k,n} z^j x^k$.

where $E = \{0, 2, 4, \ldots, \}$ is the set of even numbers, $O = \{1, 3, 5, \ldots\}$ is the set of odd numbers, and for any statement A, we let $\chi(A) = 1$ is A is true and $\chi(A) = 0$ if A is false. Thus, for example, in our language, $R_n(x) = U_{\Upsilon,2,n}(x)$ where $\Upsilon = \{2\ 1, 4\ 2\}$ and $P_n(x, 1) = U_{\Upsilon,2,n}(x)$ where $\Upsilon = \{3\ 2, 4\ 2\}$. In this case, there are some surprisingly simple formulas for the coefficients of this polynomials. For example, Kitaev and Remmel [**KR05**] proved the following.

THEOREM 2.1.

$$R_{k,2n} = \binom{n}{k}^2 (n!)^2, \tag{2.1}$$

$$R_{k,2n+1} = (k+1)\binom{n}{k+1}^2 (n!)^2 + (2n+1-k)\binom{n}{k}^2 (n!)^2 = \frac{1}{k+1}\binom{n}{k}^2 ((n+1)!)^2, \tag{2.2}$$

$$P_{1,k,2n} = \binom{n-1}{k}\binom{n}{k+1}(n!)^2, \tag{2.3}$$

$$P_{0,k,2n} = \binom{n-1}{k}\binom{n}{k}(n!)^2, \tag{2.4}$$

$$P_{0,k,2n+1} = (k+1)\binom{n}{k}\binom{n+1}{k+1}(n!)^2 = (n+1)\binom{n}{k}^2 (n!)^2, \textit{ and} \tag{2.5}$$

$$P_{0,k,2n+1} = \binom{n}{k}(n!)^2 \left(n\binom{n-1}{k} + (k+1)\binom{n}{k}\right). \tag{2.6}$$

In [**KR06**], Kitaev and Remmel studied the polynomials

(1) $A_n^{(k)}(x) = \sum_{\sigma\in S_n} x^{\overleftarrow{des}_{kN}(\sigma)} = \sum_{j=0}^{\lfloor \frac{n}{k}\rfloor} A_{j,n}^{(k)} x^j$ and
(2) $B_n^{(k)}(x,z) = \sum_{\sigma\in S_n} x^{\overrightarrow{des}_{kN}(\sigma)} z^{\chi(\sigma_1\in kN)} = \sum_{j=0}^{\lfloor \frac{n}{k}\rfloor}\sum_{i=0}^{1} B_{i,j,n}^{(k)} z^i x^j$.

where $kN = \{0, k, 2k, \ldots\}$. Again both $A_n^{(k)}(x)$ and $B_n^{(k)}(x,z)$ are special cases of $U_{\Upsilon,k,n}(x)$. When $k \geq 2$, the formulas for $A_n^{(k)}(x)$ and $B_n^{(k)}(x,z)$ become more complicated. Nevertheless, certain nice formulas arise. For example, Kitaev and Remmel [**KR06**] proved the following.

THEOREM 2.2. *For all $0 \leq j \leq k-1$ and all $n \geq 0$, we have*

$$A_{s,kn+j}^{(k)} = ((k-1)n+j)!\sum_{r=0}^{s}(-1)^{s-r}\binom{(k-1)n+j+r}{r}\binom{kn+j+1}{s-r}\prod_{i=0}^{n-1}(r+1+j+(k-1)i) \tag{2.7}$$

$$= ((k-1)n+j)!\sum_{r=0}^{n-s}(-1)^{n-s-r}\binom{(k-1)n+j+r}{r}\binom{kn+j+1}{n-s-r}\prod_{i=1}^{n}(r+(k-1)i) \tag{2.8}$$

In general, when one includes parity conditions or conditions involving equivalence mod k, then the problem of counting the number of pattern matchings become more complicated. For example, if $\tau = 2\ 1$, then the number of permutations of S_n with no τ-matches is 1 since the only permutation of S_n with no (2 1)-matches is the identity permutation $1\ 2\ \cdots\ n-1\ n$. However, according to Theorem 2.1, the number of permutations of S_m with no $\{(2\ 1), (4\ 2)\}$-2-equivalences matches is $(n!)^2$ if $m = 2n$ and is $((n+1)!)^2$ if $m = 2n+1$. Similarly, the analogue of the Kitaev's result (1.1) fails to hold in general. For example, in the case where $k = 2$ and $\tau = 1\ 2$, then (1.4) implies that for $n \geq 1$, $T^0_{(1\ 2),2,2n} = n^n(n!)$ and $T^0_{(1\ 2),2,2n+1} = (n+1)^n((n+1)!)$,

$$A(t) = \sum_{n\geq 0}\frac{t^n}{n!}|\{\sigma\in S_n : (1\ 2)\text{-}2\text{-}emch(\sigma) = 0\}| = 1 + \sum_{n\geq 1}\frac{t^{2n}}{(2n)!}n^n(n!) + \sum_{n\geq 0}\frac{t^{2n+1}}{(2n+1)!}(n+1)^n(n+1)!.$$

Moreover for any $\sigma \in S_n$, $(1\ 2)$-2-$emch(\sigma) = (1\ 2)$-2-$enlap(\sigma)$. But is easy to check that

$$\sum_{n\geq 0}\frac{t^n}{n!}\sum_{\sigma\in S_n} x^{(1\ 2)\text{-}2\text{-}emch(\sigma)} \neq \frac{A(t)}{(1-x)+x(1-t)A(t)}.$$

Next we give some examples of our polynomials. Here is a table that lists $T_{(a\ b),k,n}(x)$ for all possible values of a and b where $k=3$ and $2 \leq n \leq 8$.

$T_{(12),3,2}(x) = 1 + x$	$T_{(13),3,2}(x) = 2$	$T_{(14),3,2}(x) = 2$
$T_{(12),3,3}(x) = 4 + 2x$	$T_{(13),3,3}(x) = 4 + 2x$	$T_{(14),3,3}(x) = 6$
$T_{(12),3,4}(x) = 18 + 6x$	$T_{(13),3,4}(x) = 18 + 6x$	$T_{(14),3,4}(x) = 18 + 6x$
$T_{(12),3,5}(x) = 54 + 60x + 6x^2$	$T_{(13),3,5}(x) = 96 + 24x$	$T_{(14),3,5}(x) = 96 + 24x$
$T_{(12),3,6}(x) = 384 + 312x + 24x^2$	$T_{(13),3,6}(x) = 384 + 312x + 24x^2$	$T_{(14),3,6}(x) = 600 + 120x$
$T_{(12),3,7}(x) = 3000 + 1920x + 120x^2$	$T_{(13),3,7}(x) = 3000 + 1920x + 120x^2$	$T_{(14),3,7}(x) = 3000 + 1920x + 120x^2$
$T_{(12),3,8}(x) = 15000 + 20520x + 4680x^2 + 120x^3$	$T_{(13),3,8}(x) = 25920 + 13680x + 720x^2$	$T_{(14),3,8}(x) = 25920 + 13680x + 720x^2$

Glancing at these values, certain things become apparent. First, observe that for each of these polynomials all the coefficients are divisible by the coefficient of the highest power of x appearing in the polynomial. Second, one can observe that polynomials $T_{(ab),3,n}(x)$ depend only on b. Finally, one can also observe that for any given n, the function $T_{(ab),k,n}(x)$ takes at most three distinct values. For example when $n=5$, one can see that all the polynomials $T_{(ab),3,5}(x)$ are equal to one of $T_{(12),3,5}(x)$, $T_{(13),3,5}(x)$, or $T_{(36),3,5}(x)$ and that these three polynomials are distinct. All of these facts are true in general for any k and n since they follow from our closed forms for $T_{(ab),k,n}(x)$.

$T_{(23),3,2}(x) = 2$	$T_{(24),3,2}(x) = 2$	$T_{(25),3,2}(x) = 2$
$T_{(23),3,3}(x) = 4 + 2x$	$T_{(24),3,3}(x) = 6$	$T_{(25),3,3}(x) = 6$
$T_{(23),3,4}(x) = 18 + 6x$	$T_{(24),3,4}(x) = 18 + 6x$	$T_{(25),3,4}(x) = 24$
$T_{(23),3,5}(x) = 96 + 24x$	$T_{(24),3,5}(x) = 96 + 24x$	$T_{(25),3,5}(x) = 96 + 24x$
$T_{(23),3,6}(x) = 384 + 312x + 24x^2$	$T_{(24),3,6}(x) = 600 + 120x$	$T_{(25),3,6}(x) = 600 + 120x$
$T_{(23),3,7}(x) = 3000 + 1920x + 120x^2$	$T_{(24),3,7}(x) = 3000 + 1920x + 120x^2$	$T_{(25),3,7}(x) = 4320 + 720x$
$T_{(23),3,8}(x) = 25920 + 13680x + 720x^2$	$T_{(24),3,8}(x) = 25920 + 13680x + 720x^2$	$T_{(25),3,8}(x) = 25920 + 13680x + 720x^2$

$T_{(34),3,2}(x) = 2$	$T_{(35),3,2}(x) = 2$	$T_{(36),3,2}(x) = 2$
$T_{(34),3,3}(x) = 6$	$T_{(35),3,3}(x) = 6$	$T_{(36),3,3}(x) = 6$
$T_{(34),3,4}(x) = 18 + 6x$	$T_{(35),3,4}(x) = 24$	$T_{(36),3,4}(x) = 24$
$T_{(34),3,5}(x) = 96 + 24x$	$T_{(35),3,5}(x) = 96 + 24x$	$T_{(36),3,5}(x) = 120$
$T_{(34),3,6}(x) = 600 + 120x$	$T_{(35),3,6}(x) = 600 + 120x$	$T_{(36),3,6}(x) = 600 + 120x$
$T_{(34),3,7}(x) = 3000 + 1920x + 120x^2$	$T_{(35),3,7}(x) = 4320 + 720x$	$T_{(36),3,7}(x) = 4320 + 720x$
$T_{(34),3,8}(x) = 25920 + 13680x + 720x^2$	$T_{(35),3,8}(x) = 25920 + 13680x + 720x^2$	$T_{(36),3,8}(x) = 25920 + 13680x + 720x^2$

3. Finding the coefficients for $U_{\Upsilon,k,n}(x)$ by inclusion-exclusion

In this section, we shall show how we can use inclusion-exclusion to obtain an expression for $U_{\Upsilon,k,n}(x)$ for any $\Upsilon \subset E_k$. The idea is as follows. Suppose that we fix k and $\Upsilon \subseteq E_k$. Given any two element sequence $ab \in E_k$, we shall write $ab \approxeq xy \mod k$ if (i) $x \equiv a \mod k$, (ii) $y \equiv b \mod k$, (iii) $a < b$ implies $x < y$, and (iv) $a > b$ implies $x > y$. Then for each $n \geq 1$, we let $\Upsilon_n = \{xy : 1 \leq x, y \leq n \ \&\ xy \approxeq ab \mod k \text{ where } (ab) \in \Upsilon\}$. For each $xy \in \Upsilon_n$, we let $C_{xy,n}$ equal the set of all $\sigma = \sigma_1 \cdots \sigma_n \in S_n$ such that there exist an $1 \leq i < n$ such that $\sigma_i = x$ and $\sigma_{i+1} = y$. Given $\sigma \in S_n$, we define $Pr_{\Upsilon,n}(\sigma)$, the property set of σ relative to Υ, to be the set of all $xy \in \Upsilon_n$ such that $\sigma \in C_{xy,n}$. Then we define the following.

1. For each $T \subseteq \Upsilon_n$, let $E_{=T,\Upsilon,n} = \{\sigma \in S_n : Pr_{\Upsilon,n}(\sigma) = T\}$ and $\beta_{T,\Upsilon,n} = |E_{=T,\Upsilon,n}|$.
2. For each $T \subseteq \Upsilon_n$, let $E_{\supseteq T,\Upsilon,n} = \{\sigma \in S_n : Pr_{\Upsilon,n}(\sigma) \supseteq T\}$ and $\alpha_{T,\Upsilon,n} = |E_{\supseteq T,\Upsilon,n}|$.
3. For each $r \geq 0$, let $\beta_{r,\Upsilon,n} = \sum_{S \subseteq \Upsilon_n, |S|=r} \beta_{S,\Upsilon,n}$ and $\alpha_{r,\Upsilon,n} = \sum_{S \subseteq \Upsilon_n, |S|=r} \alpha_{S,\Upsilon,n}$.

It is an easy consequence of the inclusion-exclusion principle that

$$\sum_{t \geq 0} \beta_{t,\Upsilon,n} x^t = \sum_{t \geq 0} \alpha_{t,\Upsilon,n} (x-1)^t. \tag{3.1}$$

It is also easy to see from our definitions that

$$\sum_{t \geq 0} \beta_{t,\Upsilon,n} x^t = U_{\Upsilon,k,n}(x). \tag{3.2}$$

Thus we get an expression for $U_{\Upsilon,k,n}(x)$ by calculating the RHS of (3.1).

Next we observe that it is easy to compute $\alpha_{T,\Upsilon,n}$. We say that $T \subseteq \Upsilon_n$ is consistent if there does not exist distinct ab and cd in T such that either $a = c$ or $b = d$. For example, if $k = 4$ and $\Upsilon = \{12, 34, 32, 46\}$, then $\Upsilon_7 = \{12, 16, 56, 34, 32, 72, 76, 46\}$. Then $T_1 = \{12, 16, 34\}$ and $T_2 = \{12, 32, 76\}$ are not consistent while $T_3 = \{12, 34, 46\}$ is consistent. First we claim that if T is consistent, then $\alpha_{T,\Upsilon,n} = (n-|T|)!$. That is, we need to construct $E_{\supseteq T,\Upsilon,n}$ which consists of all permutations $\sigma \in S_n$ such that each pattern in T occurs consecutively in σ. We do this by first constructing the maximal blocks of elements of $\{1, \ldots, n\}$ where xy occurs consecutively in a block if and only if $xy \in T$. For example, if $n = 7$ and $T = T_3$ as given above, then the maximal blocks constructed from T are 12, 346, 5 and 7. Then it is easy to see that any permutation of the maximal blocks constructed from T corresponds to a permutation $\sigma \in E_{\supseteq T,\Upsilon,n}$. For example, the permutation of the maximal blocks 346 5 12 7 corresponds to the permutation 3 4 6 5 1 2 7 $\in E_{\supseteq T_3,\Upsilon,7}$. Now it is easy to see that the number of maximal blocks of $\{1, \ldots, n\}$ constructed from T is $n - |T|$. Thus $\alpha_{T,\Upsilon,n} = |E_{\supseteq T,\Upsilon,n}| = (n - |T|)!$. Of course, if T is inconsistent, there there is no permutation $\sigma \in S_n$ such that all the sequences in T occur consecutively in σ. In this situation, $\alpha_{T,\Upsilon,n} = 0$.

Thus to compute $\alpha_{t,\Upsilon,n}$, we need only count the number of consistent subsets of size t in Υ_n. We can think of this problems as counting the number of rook placements of size t in a certain board associated with Υ_n. That is, given Υ_n, let $B_{\Upsilon,n}$ be the set of all (x, y) such that $xy \in \Upsilon_n$. For example, if $k = 4$ and $\Upsilon = \{12, 34, 32, 46\}$ so that $\Upsilon_7 = \{12, 16, 56, 34, 32, 72, 76, 46\}$, then $B_{\Upsilon,7}$ consists of the shaded squares on the board pictured in Figure 1.

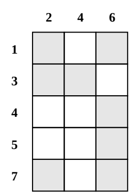


FIGURE 1. The board $B_{\Upsilon,7}$.

Given any board $B \subseteq \{1, \ldots, n\} \times \{1, \ldots, n\}$, we let $r_k(B)$ denote the number of placements of k rooks in B such that no two rooks lie in the same row or the same column. It is then easy to see that number of consistent subsets of size t in Υ_n equals $r_t(B_{\Upsilon,n})$ and thus, $\alpha_{t,\Upsilon,n} = (n-t)!r_t((B_{\Upsilon,n})$. It follows that

$$\begin{aligned} U_{\Upsilon,k,n}(x) &= \sum_{t\geq 0} \beta_{t,\Upsilon,n} x^t = \sum_{t\geq 0} \alpha_{t,\Upsilon,n}(x-1)^t \\ &= \sum_{t\geq 0} (n-t)! r_t(B_{\Upsilon,n}) \sum_{s=0}^{t} (-1)^{t-s} \binom{t}{s} x^s = \sum_{s\geq 0} x^s \sum_{t=s}^{n} (n-t)!(-1)^{t-s} \binom{t}{s} r_t(B_{\Upsilon,n}). \end{aligned}$$

The problem with formula (3.3) is that we obtain an expression for the coefficients of $U_{\Upsilon,k,n}(x)$ in terms of the numbers $r_t(B_{\Upsilon,n})$ which are not easy to compute in general. There are however some special cases of (3.3) where the numbers $r_t(B_{\Upsilon,n})$ are familiar. That is, suppose $\Upsilon = \{(1k)\}$. Then it is easy to see that $B_{\Upsilon,kn+j}$ consists of the set of squares $\{(1 + ik, jk) : 0 \leq i < j \leq n\}$. For example, if $k = 3$ and $\Upsilon = \{(13)\}$, then $B_{\Upsilon,12}$ consists of the shaded squares on the board pictured in Figure 2.

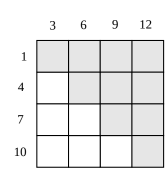


FIGURE 2. The board $B_{\{13\},12}$.

It is well known that the Stirling number of the second kind, $S_{n+1,k}$, is the number of placements of $n+1-k$ rooks on the staircase board, consisting of columns of heights $0, 1, \ldots, n$ reading from right to left, so that no two rooks lie in the same row or column. It then easily follows that

$$T^s_{(1k),k,kn+j} = U^s_{\{(1k)\},k,kn+j} = \sum_{r=s}^{n} (-1)^{r-s} \binom{r}{s} (kn+j-r)! S_{n+1,n+1-r}. \tag{3.3}$$

Another case that involves the Stirling numbers is when $\Upsilon = D_k$. As pointed out in the introduction, in that case, Υ-k-$emch(\sigma) = des(\sigma)$. In this case the board the $B_{\Upsilon,n}$ equals $\{(j,i) : 0 \leq i < j \leq n\}$ which is equivalent to a staircase board with column heights $0, 1, \ldots, n-1$.

It is also well known that the Eulerian numbers, $E_{m,n}$ counts the number of permutations in S_m that have exactly n descents. Thus we can derive the following formula for the Eulerian numbers in terms of the Stirling numbers.

$$E_{n,s} = U^s_{A_k,k,n}(x) = \sum_{r=s}^{n} (-1)^{r-s} \binom{r}{s} (n-r)! S_{n,n-r}. \tag{3.4}$$

In some other cases, we have been able to derive formulas that involve sums over products of Stirling numbers. In such cases, the board $B_{\Upsilon,n}$ naturally breaks up as a disjoint union of staircase boards. However, because of lack of space, we shall not give such examples in this paper.

4. Finding the coefficients of $U_{\Upsilon,k,n}$ by iterating recursions

In this section, we shall give an alternative approach to finding the $U_{\Upsilon,k,n}$ that exploits the fact that we can find simple recursion for the polynomials $U_{\Upsilon,k,n}$.

Given any permutation $\sigma = \sigma_1 \cdots \sigma_n \in S_n$, we label with the integers from 0 to n (from left to right) the possible positions of where we can insert $n+1$ to get a permutation in S_{n+1}. In other words, inserting $n+1$ in position 0 means that we insert $n+1$ at the beginning of σ and for $i \geq 1$, inserting $n+1$ in position i means we insert $n+1$ immediately after σ_i. In such a situation, we let $\sigma^{(i)}$ denote the permutation of S_{n+1} that results by inserting $n+1$ in position i.

Throughout the rest of this section, we shall assume that $k \geq 2$ and $\Upsilon \subseteq A_k$ is a subset of the form $\{(x_1,y_1),(x_2,y_2),\ldots,(x_t,y_t)\}$ where for all i,j $y_i \equiv y_j \mod k$. Now, define $y = min(\{y_1,\ldots,y_t\})$ and $\alpha = |\{x_i : x_i < y\}|$. We then let $Asc_{\Upsilon,k}(\sigma) = \{i : \sigma_i < \sigma_{i+1}\ \&\ \sigma_i \equiv x_j \mod k\ \&\ \sigma_{i+1} \equiv y_j \mod k$ for some $(x_j,y_j) \in \Upsilon\}$. We shall call the elements of $Asc_{\Upsilon,k}(\sigma)$ the Υ-ascents of σ.

For $j = y-k+1,\ldots,y-1$, let Δ_{kn+j} be the operator which sends x^s to $sx^{s-1} + (kn+j-s)x^s$ and Γ_{kn+y} be the operator that sends x^s to $((k-|\Upsilon|)n+y+s-\alpha)x^s + (|\Upsilon|n+\alpha-s)x^{s+1}$. Then we have the following.

Theorem 4.1. *Given Υ, y, and α as described above, the polynomials $\{U_{\Upsilon,k,n}(x)\}_{n\geq 1}$ satisfy the following recursions.*

(1) $U_{\Upsilon,k,1}(x) = 1$,
(2) *For* $j = y-k+1,\ldots,y-1$, $U_{\Upsilon,k,kn+j}(x) = \Delta_{kn+j}(U_{\Upsilon,k,kn+j-1}(x))$, *and*
(3) $U_{\Upsilon,k,kn+y}(x) = \Gamma_{kn+y}(U_{\Upsilon,k,kn+y-1}(x))$.

Proof. Part (1) is trivial.

For part (2), fix j such that $y-k+1 \leq j \leq y-1$. Now suppose $\sigma = \sigma_1 \cdots \sigma_{kn+j-1} \in S_{kn+j-1}$ and $asc_{\Upsilon,k}(\sigma) = s$. It is then easy to see that if we insert $kn+j$ in position i where $i \in Asc_{\Upsilon,k}(\sigma)$, then $asc_{\Upsilon,k}(\sigma^{(i)}) = s-1$. However, if we insert $kn+j$ in position i where $i \notin Asc_{\Upsilon,k}(\sigma)$, then $asc_{\Upsilon,k}(\sigma^{(i)}) = s$. Thus $\{\sigma^{(i)} : i = 0,\ldots,kn+j-1\}$ gives a contribution of $sx^{s-1} + (kn+j-s)x^s$ to $U_{\Upsilon,k,kn+j}$.

For part (3), suppose $\sigma = \sigma_1 \cdots \sigma_{kn+y-1} \in S_{kn+y-1}$ and $asc_{\Upsilon,k}(\sigma) = s$. In this situation we can create a Υ-ascent, but we can't lose one. That is, if we place $kn+y$ after any element equivalent to $x_i \mod k$ for some $(x_i,y_i) \in \Upsilon$ which isn't already part of a Υ-ascent, we would create an additional Υ-ascent. There are $|\Upsilon|n+\alpha-s$ such locations. This means that the number of locations that keep the number of ascents the same must be $(k-|\Upsilon|)n+y+s-\alpha$ as the two must sum to $kn+y$. Thus $\{\sigma^{(i)} : i = 0,\ldots,kn+y-1\}$ gives a contribution of $((k-|\Upsilon|)n+y+s-\alpha)x^s + (|\Upsilon|n+\alpha-s)x^{s+1}$ to $U_{\Upsilon,k,kn+y}$. ◻

We can give combinatorial proofs of two simple formulas for the extreme coefficients of $U_{\Upsilon,k,n}(x)$.

Theorem 4.2. *Let Υ, y, and α be as described above. Then for all $k \geq 2$, for all $j = y-k, \ldots, y-1$ and n such that $kn+j > 0$,*

$$U^0_{\Upsilon,k,kn+j} = ((k-1)n+j)! \prod_{i=0}^{n-1} (k-1)n+j+1-\alpha-i(|\Upsilon|-1) \tag{4.1}$$

$$U^n_{\Upsilon,k,kn+j} = ((k-1)n+j)! \prod_{i=0}^{n-1} \alpha+i(|\Upsilon|-1) \tag{4.2}$$

Proof. Clearly when $n=0$, the only $j \in \{y-k,\ldots,y-1\}$ such that $kn+j>0$ are $j=1,\ldots y-1$. In these cases, no permutation σ of S_j can have an Υ-k- equivalence match so that $U_{\Upsilon,k,j}(x) = j!$. By convention, we assume the empty product is equal to 1 so that our formulas holds when $n=0$.

Next assume that $n \geq 1$ and $\Upsilon = \{(x_i,y_i) : i = 1,\ldots t\}$ where $x_1,\ldots x_\alpha$ consist of those x_i's such that $(x_i,y) \in \Upsilon$. Suppose that $j \in \{y-k,\ldots,y-1\}$.

First we consider those permutations $\sigma \in S_{kn+j}$ such that Υ-k-*emch*$(\sigma) = 0$. We claim that we can construct all such σ as follows. By our definition, there are $(k-1)n+j$ elements in $\{1,\ldots,kn+j\}$ which are not equivalent to $y \mod k$. We can arrange these elements in $((k-1)n+j)!$ ways. Given an arrangement τ of the elements in $\{1,\ldots,kn+j\}$ which are not equivalent to $y \mod k$, we can extend τ to a permutation $\sigma \in S_{kn+j}$ such that Υ-k-*emch*$(\sigma) = 0$ as follows. First we can insert y into τ so that we do not create any Υ-k-equivalence matches. Clearly this can be done in $(k-1)n+j+1-\alpha$ ways since all we have to do is to ensure that we do not insert y immediately after any of $x_1,\ldots x_\alpha$. Now suppose τ_1 is a sequence that results from inserting y into τ so that we do not create any Υ-k-equivalence matches. Then, the number of ways to insert $y+k$ into τ_1 so that we do not create any Υ-k-equivalence matches is $(k-1)n+j+1-\alpha-(|\Upsilon|-1)$. That is there are $(k-1)n+j+2$ possible ways to insert $y+k$ into τ_1 but that are $\alpha+|\Upsilon|$ elements z such that if we insert $y+k$ after z, then we would form an Υ-k-equivalence match. Now suppose τ_2 is a sequence that results from inserting $y+k$ into τ_1 so that we do not create any Υ-k-equivalence matches. Then, the number of ways to insert $y+2k$ into τ_2 so that we do not create any Υ-k-equivalence matches is $(k-1)n+j+1-\alpha-2(|\Upsilon|-1)$. That is there are $(k-1)n+j+3$ possible ways to insert $y+2k$ into τ_2 but that are $\alpha+2|\Upsilon|$ elements z such that if we insert $y+2k$ after z, then we would form an Υ-k-equivalence match. Continuing on in this way, we see that $U^0_{\Upsilon,k,kn+j} = ((k-1)n+j)! \prod_{i=0}^{n-1} (k-1)n+r+j+1-\alpha-i(|\Upsilon|-1)$.

Next we consider those permutations $\sigma \in S_{kn+j}$ such that Υ-k-*emch*$(\sigma) = n$. We claim that we can construct all such σ as follows. By our definition, there are $(k-1)n+j$ elements in $\{1,\ldots,kn+j\}$ which are not equivalent to $y \mod k$. We can arrange these elements in $((k-1)n+j)!$ ways. Given an arrangement τ of the elements in $\{1,\ldots,kn+j\}$ which are not equivalent to $y \mod k$, we can extend τ to a permutation $\sigma \in S_{kn+j}$ such that Υ-k-*emch*$(\sigma) = n$ as follows. Clearly, we must insert $y, y+k, \ldots, y+(n-1)k$ in such a way that each of these elements create an Υ-k-equivalence match. Thus we must insert y into τ so that it immediately follows one of $x_1,\ldots,x_\alpha$. Hence we have α ways to insert y. Now suppose τ_1 is a sequence that results from inserting y into τ so that we did create a Υ-k-equivalence match. Then the number of ways to insert $y+k$ into τ_1 so that we create another Υ-k-equivalence match is $\alpha+(|\Upsilon|-1)$ since there $\alpha+|\Upsilon|$ elements $x < y+k$ such that $(x\ (y+k))$ would be an Υ-k-equivalence match and we can not insert $y+k$ immediately before y. Now suppose τ_2 is a sequence that results from inserting $y+k$ into τ_1 so that we have created a second Υ-k-equivalence match. Then the number of ways to insert $y+2k$ into τ_2 so that we create an additional Υ-k-equivalence matches is $\alpha = 2(|\Upsilon|-1)$ since there $\alpha+2|\Upsilon|$ elements $x < y+k$ such that $(x\ (y+2k))$ would be an Υ-k-equivalence match and we can not insert $y+2k$ immediately before y or $y+2k$. Continuing on in this way, we see that $U^n_{\Upsilon,k,kn+j} = ((k-1)n+j)! \prod_{i=0}^{n-1} \alpha+i(|\Upsilon|-1)$. $\square$

This given, we can derive a general formula $U^s_{\Upsilon,k,n}$ using the recursions implicit in Theorem 4.1. It is easy to see from Theorem 4.1 that we have two following recursions for the coefficients $U^s_{\Upsilon,k,n}$.

For $y-k+1 \leq j \leq y-1$,

$$U^s_{\Upsilon,k,kn+j} = (kn+j-s)U^s_{\Upsilon,k,kn+j-1} + (s+1)U^{s+1}_{\Upsilon,k,kn+j-1} \tag{4.3}$$

Similarly, we have

$$U^s_{\Upsilon,k,kn+y} = ((k-|\Upsilon|)n+y-\alpha+s)U^s_{\Upsilon,k,kn+y-1} + (|\Upsilon|n+\alpha-s+1)U^{s-1}_{\Upsilon,k,kn+y-1} \tag{4.4}$$

We will now turn to a closed form for $U^s_{\Upsilon,k,kn+j}$. This formula was obtained by using (4.3) and iterating these recursions from the bottom up.

THEOREM 4.3. *For all $y-k \le j \le y-1$ and all $s \le n$ such that $kn+j>0$, we have*

$$U^s_{\Upsilon,k,kn+j} = ((k-1)n+j)!\left[\sum_{r=0}^{s}(-1)^{s-r}\binom{(k-1)n+r+j}{r}\binom{kn+j+1}{s-r}\Gamma(r,j,n)\right]$$

$$\text{where } \Gamma(r,j,n) = \prod_{i=0}^{n-1}((k-1)n+r+j+1-\alpha-i(|\Upsilon|-1))$$

PROOF. We shall prove by induction, first on s, and then n that our formulas hold. That is, by Theorem 4.2, our formulas hold when $s=0$ for all $n \ge 0$ and and $y-k \le j \le y-1$ if $kn+j>0$. Next assume that our formulas satisfy the recursions (4.3) and (4.4), which we will verify later in the proof. Then, we can complete the induction as follows. First assume that that our formulas hold at some s for all $n \ge s$ and $y-k \le j \le y-1$ if $kn+j>0$. Note that the recursions (4.3) and (4.4) can be rewritten as

$$U^{s+1}_{\Upsilon,k,kn+j-1} = \frac{1}{s+1}(U^s_{\Upsilon,k,kn+j} - (kn+j-s)U^s_{\Upsilon,k,kn+j-1}), \tag{4.5}$$

for $y-k+1 \le j \le y-1$, and

$$U^{s+1}_{\Upsilon,k,kn+y-1} = \frac{1}{((k-|\Upsilon|)n+y-\alpha+s+1)}(U^{s+1}_{\Upsilon,k,kn+y} - (|\Upsilon|+\alpha-s)U^s_{\Upsilon,k,kn+y-1}) \tag{4.6}$$

Thus in particular, (4.5) implies our formulas hold at $s+1$ when $n \ge s+1$ and $j = y-k, \dots, y-2$. We are then able to use (4.6) to establish that our formula holds at $s+1$ when $n \ge s+1$ and $j = y-1$.

Thus to complete our proof, we need only verify that our formulas satisfy the recursions (4.3) and (4.4). In order to simplify the algebra, we will convert the form from (4.5) to the following

$$U^s_{\Upsilon,k,kn+j} = \sum_{r=0}^{s}\frac{(-1)^{s-r}((k-1)n+r+j)!(kn+j+1)!\Gamma(r,j,n)}{(kn+j-s+r+1)!r!(s-r)!}. \tag{4.7}$$

So, for $y-k+1 \le j \le y-1$ plugging in the above form into the RHS of (4.3) gives

$$(kn+j-s)\left[\sum_{r=0}^{s}\frac{(-1)^{s-r}((k-1)n+r+j-1)!(kn+j)!\Gamma(r,j-1,n)}{(kn+j-s+r)!r!(s-r)!}\right]$$
$$+(s+1)\left[\sum_{r=0}^{s+1}\frac{(-1)^{s+1-r}((k-1)n+r+j-1)!(kn+j)!\Gamma(r,j-1,n)}{(kn+j-s+r-1)!r!(s+1-r)!}\right]$$

Removing the $s+1$ term from the second summand, recognizing that $\Gamma(r,j-1,n) = \Gamma(r-1,j,n)$ and combining the rest of the terms yields

$$\sum_{r=0}^{s}\frac{(-1)^{s-r}((k-1)n+r+j-1)!(kn+j)!\Gamma(r-1,j,n)\,[-r(kn+j+1)]}{(kn+j-s+r)!r!(s+1-r)!} + \frac{((k-1)n+s+j)!\Gamma(s+1,j-1,n)}{s!}$$

Since there is a factor of r in the numerator, we may omit the $r=0$ term from the summand, shift indices and recognize that $\Gamma(s+1,j-1,n) = \Gamma(s,j,n)$ to get

$$\sum_{r=0}^{s-1}\frac{(-1)^{s-r}((k-1)n+r+j)!(kn+j+1)!\Gamma(r,j,n)}{(kn+j-s+r+1)!r!(s-r)!} + \frac{((k-1)n+s+j)!\Gamma(s,j,n)}{s!}$$
$$= \sum_{r=0}^{s}\frac{(-1)^{s-r}((k-1)n+r+j)!(kn+j+1)!\Gamma(r,j,n)}{(kn+j-s+r+1)!r!(s-r)!} = U^s_{\Upsilon,k,kn+j}$$

Thus we have shown that our formula for $U^s_{\Upsilon,k,kn+j}$ satisfies (4.3) for $y-k+1 \le j \le y-1$. We will now show that our formula satisfies (4.4). The RHS of (4.4) becomes

$$((k-|\Upsilon|)n+s+y-\alpha)\left[\sum_{r=0}^{s}\frac{(-1)^{s-r}((k-1)n+r+y-1)!(kn+y)!\Gamma(r,y-1,n)}{(kn+y-s+r)!r!(s-r)!}\right]$$
$$+(|\Upsilon|n+\alpha-s+1)\left[\sum_{r=0}^{s-1}\frac{(-1)^{s-r-1}((k-1)n+r+y-1)!(kn+y)\Gamma(r,y-1,n)!}{(kn+y-s+r+1)!r!(s-r-1)!}\right]$$

Removing the s term from the first summand, and combining the rest of the terms yields

$$\begin{aligned}
&\sum_{r=0}^{s-1} \frac{(-1)^{s-r}((k-1)n+r+y-1)!(kn+y)!\Gamma(r,y-1,n)\,[(kn+y+1)(kn-n|\Upsilon|+r+y-\alpha)]}{(kn+y-s+r+1)!r!(s-r)!}\\
&+\frac{((k-|\Upsilon|)n+s+y-\alpha)((k-1)n+y+s-1)!\Gamma(s,y-1,n)}{s!}\\
= &\sum_{r=0}^{s-1} \frac{(-1)^{s-r}((k-1)n+r+y-1)!(kn+y+1)!\Gamma(r,y-1,n)}{(kn+y-s+r+1)!r!(s-r)!}\\
&+\frac{((k-|\Upsilon|)n+s+y-\alpha)((k-1)n+y+s-1)!\Gamma(s,y-1,n)}{s!}\\
= &\sum_{r=0}^{s} \frac{(-1)^{s-r}((k-1)n+r+y-1)!(kn+y+1)!\Gamma(r,y-1,n)((k-|\Upsilon|)n+r+y-\alpha)}{(kn+y-s+r+1)!r!(s-r)!}\\
= &\sum_{r=0}^{s} \frac{(-1)^{s-r}((k-1)n+r+y-1)!(kn+y+1)!\Gamma(r,y-k,n+1)}{(kn+y-s+r+1)!r!(s-r)!} = U^s_{\Upsilon,k,kn+y}
\end{aligned}$$

Thus we have shown that our formula for $U^s_{\Upsilon,k,kn+j}$ satisfies (4.4) as desired. □

Here is another formula for $U^s_{\Upsilon,k,kn+j}$. This one was obtained by iterating the recursions (4.3) and (4.4) from the top down.

THEOREM 4.4. *For all $y-k \le j \le y-1$ and all $s \le n$ such that $kn+j>0$, we have*

$$U^{n-s}_{\Upsilon,k,kn+j} = ((k-1)n+j)!\left[\sum_{r=0}^{s}(-1)^{s-r}\binom{(k-1)n+r+j}{r}\binom{kn+j+1}{s-r}\Omega(r,n)\right] \tag{4.8}$$

$$\text{where } \Omega(r,n) = \prod_{i=0}^{n-1}(r+\alpha+i(|\Upsilon|-1)).$$

PROOF. We shall prove by induction, first on s and then on $kn+j$ that our formulas hold. Theorem 4.2 proves our formulas hold when $s=0$ for all $n \ge 0$ and $y-k \le j \le y-1$ such that $kn+j>0$. Now assume that our formulas for $U^{n-s}_{\Upsilon,k,kn+j}$ satisfy the the recursions, (4.3) and (4.4), which we will verify later in the proof. Then, we can complete our induction as follows. Assume that our formulas for $U^{n-s}_{\Upsilon,k,kn+j}$ hold at s for all $n \ge s$ and and $y-k \le j \le y-1$ such that $kn+j>0$. Then, the recursions can be rewritten as

$$U^{n-(s+1)}_{\Upsilon,k,kn+j} = (kn+j-n+s+1)U^{n-(s+1)}_{\Upsilon,k,kn+j-1} + (n-s)U^{n-s}_{\Upsilon,k,kn+j-1} \tag{4.9}$$

for $y-k+1 \le j \le y-1$, and

$$U^{(n+1)-(s+1)}_{\Upsilon,k,k(n+1)+y-k} = ((k-|\Upsilon|)n+y-\alpha+n-s)U^{n-s}_{\Upsilon,k,kn+y-1}) + (|\Upsilon|+\alpha-n+s+1)U^{n-(s+1)}_{\Upsilon,k,kn+y-1} \tag{4.10}$$

It is easy to see that the recursions (4.10) and (4.10) will allow us to prove our formulas hold for $U^{n-(s+1)}_{\Upsilon,k,kn+j}$, for all $n \ge s+1$ and $y-k \le j \le y-1$ such that $kn+j>0$, by induction on $kn+j$ so long as we can prove a base case. In the base case, we can prove the recursion

$$U^{(s+1)-(s+1)}_{\Upsilon,k,k(n+1)+y-k} = (k-|\Upsilon|)n+y-\alpha+s-s)U^{s-s}_{\Upsilon,k,kn+y-1} + (|\Upsilon|+\alpha-s+s+1)U^{s-(s+1)}_{\Upsilon,k,kn+y-1} \tag{4.11}$$

if we interpret each term in the sense of the RHS of (4.8). The problem is that our formulas make sense even in the case

$$U^{s-(s+1)}_{\Upsilon,k,kn+y-1} = ((k-1)n+y-1)!\left[\sum_{r=0}^{s+1}(-1)^{s+1-r}\binom{(k-1)n+r+y-1}{r}\binom{kn+y}{s+1-r}\Omega(r,s)\right]. \tag{4.12}$$

However, by our definitions, it must be the case that $U^{s-(s+1)}_{\Upsilon,k,kn+y-1} = U^{-1}_{\Upsilon,k,kn+y-1} = 0$. Thus in order to establish the base case, we need an independent proof that the RHS of (4.12) is 0. In fact, we can prove much more. That is, we can give a direct combinatorial proof that

$$U^{n+1}_{\Upsilon,k,kn+j} = ((k-1)n+j)!\left[\sum_{r=0}^{n+1}(-1)^{n+1-r}\binom{(k-1)n+r+j}{r}\binom{kn+j+1}{n+1-r}\Omega(r,n)\right] = 0$$

for any $y-k \leq j \leq y-1$. We will not give this combinatorial proof here due to lack of space.

Thus to complete our induction, we need only show that our formulas satisfy the recursions (4.3) and (4.4). In order to simplify the algebra, we will again convert the form from (4.8) to the following

$$U^s_{\Upsilon,k,kn+j} = \sum_{r=0}^{n-s} \frac{(-1)^{n-s-r}((k-1)n+r+j)!(kn+j+1)!\Omega(r,n)}{((k-1)n+j+s+r+1)!r!(n-s-r)!}$$

So, for $y-k+1 \leq j \leq y-1$ plugging in the above form into the RHS of (4.3) gives

$$\begin{aligned}&(kn+j-s)\left[\sum_{r=0}^{n-s} \frac{(-1)^{n-s-r}((k-1)n+r+j-1)!(kn+j)!\Omega(r,n)}{((k-1)n+j+s+r)!r!(n-s-r)!}\right]\\ &+(s+1)\left[\sum_{r=0}^{n-s-1} \frac{(-1)^{n-s-r-1}((k-1)n+r+j-1)!(kn+j)!\Omega(r,n)}{((k-1)n+j+s+r+1)!r!(n-s-r-1)!}\right]\end{aligned}$$

Removing the $n-s$ term from the first summand, and combining the rest of the terms yields

$$\begin{aligned}&\sum_{r=0}^{n-s-1} \frac{(-1)^{n-s-r}((k-1)n+r+j-1)!(kn+j)!\Omega(r,n)\left[(kn+j+1)((k-1)n+j+r)\right]}{((k-1)n+j+s+r+1)!r!(n-s-r)!}\\ &+\frac{(kn+j-s)!\Omega(n-s,n)}{(n-s)!}\\ =\quad &\sum_{r=0}^{n-s-1} \frac{(-1)^{n-s-r}((k-1)n+r+j)!(kn+j+1)!\Omega(r,n)}{((k-1)n+j+s+r+1)!r!(n-s-r)!}+\frac{(kn+j-s)!\Omega(n-s,n)}{(n-s)!}\\ =\quad &\sum_{r=0}^{n-s} \frac{(-1)^{n-s-r}((k-1)n+r+j)!(kn+j+1)!\Omega(r,n)}{((k-1)n+j+s+r+1)!r!(n-s-r)!} = U^s_{\Upsilon,k,kn+j}\end{aligned}$$

Thus we have shown that our formula for $U^s_{\Upsilon,k,kn+j}$ satisfies (4.3) for $y-k+1 \leq j \leq y-1$. We will now show that our formula satisfies (4.4). The RHS of (4.4) becomes

$$\begin{aligned}&((k-|\Upsilon|)n+s+y-\alpha)\left[\sum_{r=0}^{n-s} \frac{(-1)^{n-s-r}((k-1)n+r+y-1)!(kn+y)!\Omega(r,n)}{((k-1)n+y+s+r)!r!(n-s-r)!}\right]\\ &+(|\Upsilon|n+\alpha-s+1)\left[\sum_{r=0}^{n-s+1} \frac{(-1)^{n-s-r+1}((k-1)n+r+y-1)!(kn+y)!\Omega(r,n)}{((k-1)n+y+s+r-1)!r!(n-s-r+1)!}\right]\end{aligned}$$

Removing the $n-s+1$ term from the second summand, and combining the rest of the terms yields

$$\begin{aligned}&\sum_{r=0}^{n-s} \frac{(-1)^{n-s-r}((k-1)n+r+y-1)!(kn+y)!\Omega(r,n)\left[(-1)(\alpha+r+n(|\Upsilon|-1))(kn+y+1)\right]}{((k-1)n+y+s+r)!r!(n-s-r+1)!}\\ &+\frac{(|\Upsilon|n+\alpha-s+1)(kn+y-s)!\Omega(n-s+1,n)}{(n-s+1)!}\\ &\sum_{r=0}^{n-s+1} \frac{(-1)^{n-s-r+1}((k-1)n+r+y-1)!(kn+y+1)!\Omega(r,n)(\alpha+r+n(|\Upsilon|-1))}{((k-1)n+y+s+r)!r!(n-s-r+1)!}\\ &\sum_{r=0}^{n-s+1} \frac{(-1)^{n-s-r+1}((k-1)n+r+y-1)!(kn+y+1)!\Omega(r,n+1)}{((k-1)n+y+s+r)!r!(n-s-r+1)!} = U^s_{\Upsilon,k,kn+y}\end{aligned}$$

Thus we have shown that our formula for $U^s_{\Upsilon,k,kn+j}$ satisfies (4.4) as desired. $\square$

5. Conclusion and perspectives

This paper can be regarded as some initial results on the study of pattern matching in permutations that include conditions on the equivalence class modulo k of the elements of the pattern. In particular, we studied the polynomials

$$T_{\tau,k,n}(x) = \sum_{\sigma\in S_n} x^{\tau\text{-}k\text{-}emch(\sigma)} = \sum_{s=0}^{n} T^s_{\tau,k,n}x^s \text{ and } U_{\Upsilon,k,n}(x) = \sum_{\sigma\in S_n} x^{\Upsilon\text{-}k\text{-}emch(\sigma)} = \sum_{s=0}^{n} U^s_{\Upsilon,k,n}x^s.$$

We developed a number of explicit formulas for these polynomials in the case where τ is a two-element sequence or when Υ is a set of ascents of the form $\{(x_1,y_1),\dots,(x_t,y_t)\}$ where for all i and j, $y_i \equiv y_j \mod k$ or a set of descents of the form $\{(y_1,x_1),\dots,(y_t,x_t)\}$ where for all i and j, $y_i \equiv y_j \mod k$. Our formulas for the coefficients of these polynomials lead to a number of interesting identities. For example, it follows from Theorems 4.3 and 4.4 that we have Υ is set of ascents of the form $\{(x_1,y_1),\dots,(x_t,y_t)\}$ where for all i and j, $y_i \equiv y_j \mod k$, $y = min(\{y_1,\dots,y_k\})$, and $\alpha = |\{x_i : x_i < y\}|$, then

$$\left[\sum_{r=0}^{s}(-1)^{s-r}\binom{(k-1)n+r+j}{r}\binom{kn+j+1}{s-r}\Gamma(r,j,n)\right]$$
$$= \left[\sum_{r=0}^{n-s}(-1)^{n-s-r}\binom{(k-1)n+r+j}{r}\binom{kn+j+1}{n-s-r}\Omega(r,n)\right]$$

where $\Gamma(r,j,n) = \prod_{i=0}^{n-1}((k-1)n+r+j+1-\alpha-i(|\Upsilon|-1))$ and $\Omega(r,n) = \prod_{i=0}^{n-1}(r+\alpha+i(|\Upsilon|-1))$. It would be nice to have a more general explanation as to how these types of identities arise.

Also the results of this paper give rise to a number of interesting bijective questions. For example, our formulas show that many of the polynomials $T_{(ab),n,kn+j}(x)$ are identical for certain values of a, b, n and j. One can ask to give a bijective proof of such facts. We have not been able to do this in all cases, but we can give can a bijective proof that $T_{(ab),k,kn+j}(x) = T_{(cd),k,kn+j}(x)$ where for all n and $1 \le j \le k$ whenever $n - \chi(b>k) + \chi(j \ge b \mod k) = n - \chi(d>k) + \chi(j \ge d \mod k)$.

There is still much work to be done on the structure of the polynomials $T_{\tau,k,n}(x)$ and $U_{\Upsilon,k,n}(x)$. First one can consider generalized Wilf equivalence questions, i.e., given k, for which patterns α and β do we have $T_{\alpha,k,n}(x) = T_{\beta,k,n}(x)$ for all n. We can also consider more complicated sets of patterns. We should note that when we consider more complicated patterns, the problems get considerably harder. For example, consider $U_{\Upsilon,k,kn+j}(x)$ where $k=3$ and $\Upsilon = \{12,23\}$. We can no longer get simple recursions for the coefficients $U^s_{\Upsilon,k,kn+j}$ since we need to keep track of more information than just the number of Υ-k-equivalence matches. That is, let

$$A_n(x,y) = \sum_{\sigma\in S_n} x^{(12)\text{-}3\text{-}emch(\sigma)}y^{(23)\text{-}3\text{-}emch(\sigma)} = \sum_{r,s\ge 0} A^{s,t}_n x^s y^t.$$

Using the methods of this paper, we can derive simple recursions for the coefficients of $A^{r,s}_n$

$$\begin{aligned}
A^{s,t}_{3n+1} &= (s+1)A^{s+1,t}_{3n} + (t+1)A^{s,t+1}_{3n} + (3n+1-s-t)A^{s,t}_{3n}\\
A^{s,t}_{3n+2} &= (2+n-s)A^{s-1,t}_{3n+1} + (t+1)A^{s,t+1}_{3n+1} + (2n+1+s-t)A^{s,t}_{3n+1}\\
A^{s,t}_{3n+3} &= (s+1)A^{s+1,t}_{3n+2} + (2+n-t)A^{s,t-1}_{3n+2} + (2n+2+t-s)A^{s,t}_{3n+2}.
\end{aligned}$$

These recursions are more difficult to iterate, but we have found explicit formulas similar to the ones described in this paper for the coefficients $A^{r,s}_n$ when either r is the maximum power of x that appears in $A_n(x,y)$ or s is the maximum power of y that appears in $A_n(x,y)$. Similarly, we can use extend the inclusion-exclusion approach of section 3 to show that $A_n(x,y) = \sum_{k,l}(n-k-l)!r_k(B_{(12),n})r_l(B_{(23),n})$.

Similar problems arise when we consider patterns of length ≥ 3. For example, if one is going to study the number of (123)-k-equivalence matches, then to develop simple recursive formulas, one needs to also keep track of the number of (12)-k-equivalence matches so that one ends up studying polynomials like

$$B_n(x,y) = \sum_{\sigma\in S_n} x^{(12)\text{-}3\text{-}emch(\sigma)}y^{(123)\text{-}3\text{-}emch(\sigma)} = \sum_{r,s\ge 0} B^{r,s}_n x^s y^t.$$

Finally, we should note we have derived q-analogues of the results of this paper.

References

[EN03] Sergi Elizalde and Marc Noy, *Consecutive patterns in permutations*, Adv. in Appl. Math. **30** (2003), no. 1-2, 110–125, Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).

[Kit] Sergey Kitaev, *Partially ordered generalized patterns*, to appear in Discrete Math.

[Kit03] Sergey Kitaev, *Generalized patterns in words and permutations*, Ph.D. thesis, Chalmers University of Technology and Göteborg University, 2003.

[KM03] Sergey Kitaev and Toufik Mansour, *Partially ordered generalized patterns and k-ary words*, Ann. Comb. **7** (2003), no. 2, 191–200.

[KR05] S. Kitaev and J. Remmel: Classifying descents according to parity,(2005), preprint.

[KR06] S. Kitaev and J. Remmel: Classifying descents according to equivalence mod k,(2006), preprint.

[MR05] Anthony Mendes and Jeffrey Remmel, *Permutations and words counted by consecutive patterns*, to appear in a special volume in honor of A. Regev.

Department of Mathematics, University of California San Diego, La Jolla CA 92093-0112. USA
E-mail address: jliese@math.ucsd.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Enumerating Bases of Self-Dual Matroids

Molly Maxwell

ABSTRACT. We define involutively self-dual matroids and prove a relationship between the bases and self-dual bases of these matroids. We use this relationship to prove an enumeration formula for the higher dimensional spanning trees in a class of cell complexes. This gives a new proof of Tutte's theorem that the number of spanning trees of a central reflex is a perfect square and solves a problem posed by Kalai about higher dimensional spanning trees in simplicial complexes. We also give a weighted version of the latter result.

The critical group of a graph is a finite abelian group whose order is the number of spanning trees of the graph. We prove that the critical group of a central reflex is a direct sum of two copies of an abelian group. We conclude with an analogous result in Kalai's setting.

RÉSUMÉ. Nous définissons la notion de matroide auto-dual par involution et nous démontrons une relation entre les bases et les bases auto-duales de ces matroides. Nous utilisons le relation pour démontrer une formule d'énumération pour les arbres couvrants de dimension supérieure dans une classe de complexes de cellules. Ceci mène à une nouvelle démonstration d'un théorème de Tutte – le nombre d'arbres couvrants d'un *central reflex* est un carré parfait – et résoud un problème posé par Kalai concernant les arbres couvrants de dimension supérieure à 1 de complexes simpliciaux. Nous donnons également une version pondérée de ce dernier résultat.

Le groupe critique d'un graphe est un groupe abélien fini dont l'ordre est le nombre d'arbres couvrants du graphe. Nous prouvons que le groupe critique d'un *central reflex* est la somme directe de deux copies d'un groupe abéliens. Nous concluons avec un résultat analogue dans le cadre posé par Kalai.

1. Introduction

A *matroid* $\mathcal{M}$ is a finite set E along with a collection $\mathcal{I}$ of subsets of E called *independent sets* which satisfy the following conditions:

(1) The empty set $\emptyset$ is in $\mathcal{I}$.
(2) If $I_1 \in \mathcal{I}$ and $I_2 \subseteq I_1$, then $I_2 \in \mathcal{I}$.
(3) If $I_1, I_2 \in \mathcal{I}$ and $|I_2| > |I_1|$, then there exists $e \in I_2 \backslash I_1$ such that $I_1 \cup \{e\} \in \mathcal{I}$.

The *bases* $\mathcal{B}$ of a matroid $\mathcal{M}$ are the maximal independent sets. The bases satisfy the conditions:

(1) $\mathcal{B}$ is non-empty.
(2) If $B_1, B_2 \in \mathcal{B}$ and $e \in B_1 \backslash B_2$, then there exists $e' \in B_2 \backslash B_1$ with

$$(B_1 \backslash \{e\}) \cup \{e'\} \in \mathcal{B}.$$

For a matroid $\mathcal{M}$, its *dual matroid* $\mathcal{M}^{\perp}$ has bases

$$\mathcal{B}(\mathcal{M}^{\perp}) := \{E \backslash B : B \in \mathcal{B}(\mathcal{M})\}.$$

DEFINITION 1.1. A matroid $\mathcal{M}$ is said to be *involutively self-dual* if it can be represented by an $n \times 2n$ $\mathbb{Z}$-valued matrix with columns indexed by $E = \{e_1, \ldots, e_n, \tilde{e}_1, \ldots, \tilde{e}_n\}$ of the form

2000 *Mathematics Subject Classification.* Primary 05B35; Secondary 05C05.

Key words and phrases. matroid, Pfaffian, duality, regular cell complex, simplicial complex, simplicial matroid, central reflex, critical group.

This work partially satisfies the requirements for the author's doctoral dissertation at the University of Minnesota, under the supervision of Professor Vic Reiner, and is partially supported by NSF grant DMS-0245379.

$$
\begin{array}{rc}
 & \begin{array}{ccc|ccc} e_1 & \dots & e_n & \tilde{e}_1 & \dots & \tilde{e}_n \end{array} \\
M = & \left[\begin{array}{c|c} \qquad N \qquad & \qquad I \qquad \end{array}\right],
\end{array}
$$

such that the matrix

$$
\begin{array}{rc}
 & \begin{array}{ccc|ccc} e_1 & \dots & e_n & \tilde{e}_1 & \dots & \tilde{e}_n \end{array} \\
M^{\perp} := & \left[\begin{array}{c|c} \qquad -I \qquad & \qquad -N \qquad \end{array}\right]
\end{array}
$$

satisfies Rowspace$(M^{\perp}) =$ Rowspace$(M)^{\perp}$ (or equivalently $N^T = -N$). In this case, the map $\phi : E \to E$ given by $e_i \mapsto \tilde{e}_i$ is a fixed-point free involution which induces a matroid isomorphism $\mathcal{M} \to \mathcal{M}^{\perp}$.

A basis B is said to be *self-dual* if it contains *exactly* one of e_i and $\tilde{e}_i$ from each pair. Equivalently, B is self-dual if $\phi(E \backslash B) = B$. From the matrix M, we see that $B_0 := \{\tilde{e}_1, \dots, \tilde{e}_n\}$ is a self-dual basis of $\mathcal{M}$.

In this paper, we use the method of Pfaffians to prove the following result.

THEOREM 1.2. *If $\mathcal{M}$ is an involutively self-dual matroid, then*

$$
\sum_{\textit{bases } B \textit{ of } M} \det(M|_B)^2 = \left(\sum_{\substack{\textit{self-dual} \\ \textit{bases } B \textit{ of } M}} |\det(M|_B)| \right)^2 .
$$

A matrix is *unimodular* if all non-singular square submatrices have determinant ± 1.

COROLLARY 1.3. If $\mathcal{M}$ is an involutively self-dual matroid and the associated matrix M is unimodular, then the number of bases of $\mathcal{M}$ equals the square of the number of self-dual bases of $\mathcal{M}$.

THEOREM 1.4. *Let $\mathcal{M}$ be an involutively self-dual matroid and let A be the concatenated matrix* $A := \begin{bmatrix} M \\ M^{\perp} \end{bmatrix}$. *Then*

$$
\operatorname{coker}(MM^T) \cong \operatorname{coker}(A) \cong H \oplus H,
$$

where H is an abelian group of order

$$
\sum_{\substack{\textit{self-dual} \\ \textit{bases } B \textit{ of } \mathcal{M}}} |\det(M|_B)|.
$$

In Section 3, we show that involutively self-dual matriods arise from cellular $2k$-spheres for k odd that are isomorphic to their duals via the antipodal map. These include the central reflexes studied by Tutte and the boundaries of simplices studied by Kalai. We apply the matroid results above to prove Theorem 1.6, Proposition 1.2 and Theorem 1.9 below.

For a p-dimensional regular cell complex X, the *dual block complex* $D(X)$ of X is a partition of X into disjoint blocks such that every i-cell σ of X is associated to a unique $(p-i)$-block $D(\sigma)$ of $D(X)$. If X is self-dual, then $D(X)$ is a regular cell complex and the blocks $D(\sigma)$ are its cells.

DEFINITION 1.5. Let k be an odd positive integer. An *antipodally self-dual cell complex* X is a regular cell complex such that $|X| = \mathbb{S}^{2k}$ and $a(X) = D(X)$, where $a : \mathbb{S}^{2k} \to \mathbb{S}^{2k}$ is the antipodal map and $D(X)$ is the dual block complex of X.

For each k-cell σ of X, its dual block $D(\sigma)$ is a k-cell in $D(X)$ and its conjugate $\tilde{\sigma}$ is defined by $\tilde{\sigma} := a(D(\sigma))$. The cells σ and $\tilde{\sigma}$ are distinct k-cells of X, and when k is odd, X and $D(X)$ can be oriented in such a way that $\tilde{\tilde{\sigma}} = \sigma$. It follows that the k-cells can be partitioned into n pairs $\{\sigma,\tilde{\sigma}\}$.

Let $\mathcal{T}_k(X)$ be the set of all k-dimensional subcomplexes T of X such that

(1) T contains the $(k-1)$-skeleton of X,
(2) $Z_k(T) = \widetilde{H}_k(T) = 0$,
(3) $\widetilde{H}_{k-1}(T)$ is a finite group.

Complexes in $\mathcal{T}_k(X)$ will be called *k-dimensional spanning trees* of X. A k-dimensional spanning tree T is said to be *self-dual* if it contains *exactly* one of σ_i and $\tilde{\sigma}_i$ from each pair.

PROPOSITION 1.1. *Let k be an odd positive integer. If X is an antipodally self-dual cell complex which contains an acyclic, self-dual spanning tree T_0, then X gives rise to an involutively self-dual matroid.*

We use Proposition 1.1, Theorem 1.2 and Lemma 3.2 to obtain the following result.

THEOREM 1.6. *Let k be an odd positive interger and let X be an antipodally self-dual cell complex which contains an acyclic, self-dual spanning tree T_0. Then*

$$\sum_{T \in \mathcal{T}_k(X)} |\widetilde{H}_{k-1}(T)|^2 = \left(\sum_{\substack{\text{self-dual} \\ T \in \mathcal{T}_k(X)}} |\widetilde{H}_{k-1}(T)| \right)^2.$$

We next discuss how Theorem 1.6 implies a result of Tutte. A *central reflex* G is an embedding of a connected, directed planar graph on the sphere $\mathbb{S}^2$ with the property that the antipodal map a sends G to an embedding of its planar dual graph G^* on $\mathbb{S}^2$. When $k = 1$, the antipodally self-dual cell complexes are precisely the central reflexes with no loops and no isthmuses. We show that every central reflex G is equivalent to a central reflex G' with no loops and no isthmuses in the sense that G and G' have the same spanning tree numbers and the same critical groups. The dual block complex of a central reflex G is an embedding of the planar dual graph G^* on the sphere $\mathbb{S}^2$. For each edge e, its dual block $D(e)$ is the edge e^* which crosses e in the dual graph and its *conjugate* $\tilde{e}$ is defined by $\tilde{e} := a(e^*)$. A *self-dual spanning tree* is a spanning tree that contains exactly one of e and $\tilde{e}$ from each pair. We let $\mathcal{D}(G)$ denote the number of self-dual spanning trees of G. In [**12**], Tutte uses the theory of electrical networks to prove the following theorem.

THEOREM 1.7. *(Tutte) If G is a central reflex, then the spanning tree number $\kappa(G) = \mathcal{D}(G)^2$.*

In Section 4.1, we show that every central reflex contains a self-dual tree. Theorem 1.6 then gives a new proof of Tutte's theorem.

The critical group of a graph is an abelian group whose order is the number of spanning trees of the graph. We use Theorem 1.4 to prove the following result.

PROPOSITION 1.2. *The critical group of a central reflex G is of the form*

$$K(G) \cong H \oplus H,$$

where H is an abelian group of order $\mathcal{D}(G)$.

Theorem 1.6 also resolves a question posed by Kalai, as we now discuss. Let $\mathcal{T}(n, k)$ be the set of all simplicial complexes T on the vertex set $\{1, 2, \ldots, n\} = [n]$ such that

(1) T has a complete $(k-1)$-skeleton,
(2) T has exactly $\binom{n-1}{k}$ k-faces,
(3) $H_k(T) = 0$.

Complexes in $\mathcal{T}(n, k)$ will be called *k-dimensional spanning trees* on the vertex set $[n]$. To each vertex i we associate a variable x_i. Let $\mathbf{x}^{\deg(T)} := \prod_{i=1}^n x_i^{\deg_T(i)}$, $m_1 := \binom{n-2}{k-1}$, and $m_2 := \binom{n-2}{k}$. Kalai ([**5**, Theorem 1, Theorem 3']) proved the following analogues of Cayley's Theorem and the Cayley-Prüfer Theorem for these k-dimensional trees:

THEOREM 1.8. *(Kalai)*

$$\sum_{T \in \mathcal{T}(n,k)} |H_{k-1}(T, \mathbb{Z})|^2 = n^{\binom{n-2}{k}},$$

and more generally

$$\sum_{T \in \mathcal{T}(n,k)} |H_{k-1}(T, \mathbb{Z})|^2 \, \mathbf{x}^{\deg(T)} = (x_1 + x_2 + \cdots + x_n)^{m_2} \prod_{i=1}^n x_i^{m_1}.$$

The *blocker* or *Alexander dual* of a simplicial complex C is defined by $C^\vee := \{S \subseteq [2k+2] : S^c \notin C\}$. A complex $T \in \mathcal{T}(2k+2,k)$ is said to be *self-dual* if $T^\vee = T$.

In Section 4.2 we show that when k is odd the complete $2k$-dimensional simplicial complex on the vertex set $[2k+2]$ can be embedded on the sphere $\mathbb{S}^{2k}$ in such a way that it forms an antipodally self-dual cell complex X. In this case, $\mathcal{T}_k(X) = \mathcal{T}(2k+2,k)$ and the two descriptions of self-dual trees given above agree.

In [**5**, Problem 3], Kalai posed a problem about the relationship between the trees and the self-dual trees in these complexes. The next result gives a solution to this problem when k is odd. We apply Theorem 1.6 to prove the first assertion. In Section 4.2 we use the method of Pfaffians to prove the second assertion.

THEOREM 1.9. *If k is an odd positive integer, then*

$$\left(\sum_{\substack{\textit{self-dual}\\ T\in\mathcal{T}(2k+2,k)}} |\widetilde{H}_{k-1}(T,\mathbb{Z})|\right)^2 = \sum_{T\in\mathcal{T}(2k+2,k)} |\widetilde{H}_{k-1}(T,\mathbb{Z})|^2,$$

and more generally

$$\sum_{\substack{\textit{self-dual}\\ T\in\mathcal{T}(2k+2,k)}} |\widetilde{H}_{k-1}(T,\mathbb{Z})|\, \mathbf{x}^{\deg(T)} = (x_1^2+x_2^2+\cdots+x_{2k+2}^2)^{\frac{m_2}{2}} \prod_{i=1}^{2k+2} x_i^{m_1};$$

or in other words,

$$\left(\sum_{\substack{\textit{self-dual}\\ T\in\mathcal{T}(2k+2,k)}} |\widetilde{H}_{k-1}(T,\mathbb{Z})|\, \mathbf{x}^{\deg(T)}\right)^2 = \sum_{T\in\mathcal{T}(2k+2,k)} |\widetilde{H}_{k-1}(T,\mathbb{Z})|^2\, \mathbf{x}^{\deg(T)}\Bigg|_{x_i\to x_i^2}.$$

COROLLARY 1.10. If k is an odd positive integer, then

$$\sum_{\substack{\text{self-dual}\\ T\in\mathcal{T}(2k+2,k)}} |\widetilde{H}_{k-1}(T,\mathbb{Z})| = (2k+2)^{\binom{2k-1}{k}}.$$

2. Proofs of Theorems 1.2 and 1.4

Before we begin the proof of Theorem 1.2, we recall that for a skew-symmetric matrix A, the *Pfaffian* of A, $\mathrm{Pf}(A)$, is a polynomial in the entries of A defined, up to a sign, by the formula

$$\mathrm{Pf}(A)^2 = \det(A).$$

More information about the general theory of Pfaffians can be found in [**7**].

SKETCH PROOF OF THEOREM 1.2. Since $N^T = -N$, the matrix

$$A := \begin{bmatrix} M \\ M^\perp \end{bmatrix} = \begin{bmatrix} N & I \\ -I & -N \end{bmatrix}$$

is skew-symmetric, and hence $\mathrm{Pf}(A)^2 = \det(A)$. We prove that

$$|\mathrm{Pf}(A)| = |\det(N+I)| \tag{2.1}$$

and that

$$\det(N+I) = \sum_{\substack{\text{self-dual}\\ \text{bases} B \text{ of } \mathcal{M}}} |\det(M|_B)|. \tag{2.2}$$

Then the result follows from the fact that

$$\sum_{\text{bases } B \text{ of } M} \det(M|_B)^2 = \det(A),$$

which comes from generalized Laplace expansion along the first n rows of A, and the relation between complementary minors of M and $M^\perp$.

Proof of (2.1): We begin by noting that

$$PAQ = \begin{bmatrix} I & 0 \\ 0 & N^2 - I \end{bmatrix}, \tag{2.3}$$

where the matrices

$$P = \begin{bmatrix} I & 0 \\ -N & I \end{bmatrix}, Q = \begin{bmatrix} 0 & -I \\ -I & -N \end{bmatrix}$$

both have determinant ± 1. Since N is skew-symmetric, this implies that

$$\pm \det(A) = \det(N+I)(N-I) = \pm \det(N+I)^2$$

where the last equality uses the fact that

$$N - I = -(N^T + I) = -(N+I)^T.$$

Since $\det(A) = \mathrm{Pf}(A)^2$, it follows that

$$|\mathrm{Pf}(A)| = |\det(N+I)|.$$

Proof of (2.2): Now we set $X = N$ and $Y = I$ and use the general fact that if X and Y are $n \times n$ matrices, then

$$\det(X+Y) = \sum_{U \subseteq [n]} \det X_U,$$

where X_U denotes the matrix formed by replacing the columns in X indexed by $U \subseteq [n]$ with the corresponding columns in Y. This formula can be proved using the multilinearity of the determinant and induction.

□

In this paper, we'll let $\mathbb{Z}_d$ denote the cyclic group $\mathbb{Z}/d\mathbb{Z}$.

Proof of Theorem 1.4. In [**6**, Theorem 18], Kuperberg proves that for any skew-symmetric $2n \times 2n$ matrix A, there exists a matrix $B \in GL_{2n}(\mathbb{Z})$ such that $B^T AB$ is a direct sum of matrices of this form:

$$B^T AB = \bigoplus_{i=1}^{r} \begin{bmatrix} 0 & a_i \\ -a_i & 0 \end{bmatrix}.$$

Hence

$$\mathrm{coker}(A) \cong \bigoplus_{i=1}^{r} \mathrm{coker} \begin{bmatrix} 0 & a_i \\ -a_i & 0 \end{bmatrix} \cong \bigoplus_{i=1}^{r} \mathbb{Z}_{a_i}^2 \cong H \oplus H,$$

where $H := \bigoplus_{i=1}^{r} \mathbb{Z}_{a_i}$.

We've shown that $|\mathrm{coker}(A)| = \det(A) = |H|^2$. From the proof of Theorem 1.2, we have

$$\det(A) = \left(\sum_{\substack{\text{self-dual} \\ \text{bases } B \text{ of } M}} |\det(M|_B)| \right)^2,$$

and it follows that

$$|H| = \sum_{\substack{\text{self-dual} \\ \text{bases } B \text{ of } M}} |\det(M|_B)|.$$

□

As one might expect, the matrix N controls the behavior of $\mathrm{coker}(A)$. We make this more precise in the next proposition. Let $\mathrm{Syl}_p(G)$ denote the p-primary component of an abelian group G.

Proposition 2.1. *If a matrix A has the form*

$$A = \begin{bmatrix} N & I \\ -I & -N \end{bmatrix}$$

and is skew-symmetric, then for primes $p \neq 2$,

$$\mathrm{Syl}_p(\mathrm{coker}(A)) \cong \mathrm{Syl}_p(\mathrm{coker}(N+I)) \oplus \mathrm{Syl}_p(\mathrm{coker}(N+I)).$$

Proof. From line (2.3) we have $\mathrm{coker}(A) = \mathrm{coker}(N+I)(N-I)$. We note that $(N+I) - (N-I) = 2I$ and $N - I = -(N^T + I) = -(N+I)^T$. The result then follows from Lemma 2.1 below. □

Lemma 2.1. *(* [**4**, Lemma 16], [**1**, Proposition 3.1]*) Let G be a finite abelian group, and let α, β be two endomorphisms $G \to G$ satisfying $\beta - \alpha = m \cdot I_G$ for some $m \in \mathbb{Z}$. Then for any prime p that does not divide m, we have*

$$\mathrm{Syl}_p(\mathrm{coker}(\alpha\beta)) \cong \mathrm{Syl}_p(\mathrm{coker}(\alpha)) \oplus \mathrm{Syl}_p(\mathrm{coker}(\beta)).$$

Corollary 2.2. The group H in Theorem 1.4 is "almost" $\mathrm{coker}(N+I)$: for primes $p \neq 2$, one has

$$\mathrm{Syl}_p(H) \cong \mathrm{Syl}_p(\mathrm{coker}(N+I)).$$

Example 4.3 below shows that it is necessary to exclude $p = 2$ in the previous corollary.

3. Antipodally Self-Dual Regular Cell Complexes

We begin this section by briefly describing the dual block complex of a regular cell complex X. More information on this topic can be found in [**9**]. The dual block complex $D(X)$ of a p-dimensional regular cell complex X is a partition of X into disjoint blocks. For an i-cell τ in X, its *dual block* $D(\tau)$ is a $(p-i)$-block in $D(X)$. When X is self-dual, $D(X)$ is a regular cell complex and the dual blocks $D(\tau)$ are its cells.

Definition 3.1. Let k be an odd positive integer. An *antipodally self-dual cell complex* X is a regular cell complex such that $|X| = \mathbb{S}^{2k}$ and $a(X) = D(X)$, where $a : \mathbb{S}^{2k} \to \mathbb{S}^{2k}$ is the antipodal map and $D(X)$ is the dual block complex of X.

For each k-cell σ of X, its dual block $D(\sigma)$ is a k-cell in $D(X)$ and its conjugate $\tilde{\sigma}$ is defined by $\tilde{\sigma} := a(D(\sigma))$. The cells σ and $\tilde{\sigma}$ are distinct k-cells of X. When k is odd, we use an inductive argument similar to that in [**9**, Theorem 65.1] to orient X and $D(X)$ in such a way that $\tilde{\tilde{\sigma}} = \sigma$. It follows that the k-cells can be partitioned into n pairs $\{\sigma,\tilde{\sigma}\}$.

Let $\mathcal{T}_k(X)$ be the set of all k-dimensional subcomplexes T of X such that

(1) T contains the $(k-1)$-skeleton of X,
(2) $Z_k(T) = \widetilde{H}_k(T) = 0$,
(3) $\widetilde{H}_{k-1}(T)$ is a finite group.

Complexes in $\mathcal{T}_k(X)$ will be called *k-dimensional spanning trees* of X. A k-dimensional spanning tree T is said to be *self-dual* if it contains *exactly* one of σ_i and $\tilde{\sigma}_i$ from each pair. Equivalently, T is self-dual if $\widetilde{X \backslash T} = \{\tilde{\tau} : \tau \nsubseteq T\} = T$.

For a collection C of k-cells of X, the closure of C is the cell complex defined by $\overline{C} := C \cup X^{(k-1)}$, where $X^{(k-1)}$ denotes the $(k-1)$-skeleton of X. X gives rise to a matroid $\mathcal{M}$ by setting

- $E =$ the set of all k-cells of X,
- $\mathcal{I} =$ collections C of k-cells of X with $Z_k(\overline{C}) = \widetilde{H}_k(\overline{C}) = 0$,
- $\mathcal{B} =$ collections C of k-cells of X with $\overline{C} \in \mathcal{T}_k(X)$.

In the proof of the next proposition, we see that the boundary of each k-cell can be represented as a vector. Then the elements of $\mathcal{I}$ correspond to collections of vectors that are independent over $\mathbb{Z}$ (and hence over $\mathbb{Q}$) and the elements of $\mathcal{B}$ correspond to $\mathbb{Q}$-bases for the span of the vectors.

Proposition 3.1. *Let k be an odd positive integer. If X is an antipodally self-dual cell complex which contains an acyclic, self-dual spanning tree T_0, then X gives rise to an involutively self-dual matroid.*

Then we use Proposition 3.1, Theorem 1.2 and Lemma 3.2 to obtain Theorem 1.6.

Sketch Proof of Proposition 3.1: The k^{th} incidence matrix $I^k(X)$ is the matrix whose rows are labeled by the $(k-1)$-faces of X, whose columns are labeled by the k-faces of X, and whose entries are the incidence numbers

$$\epsilon(\sigma,\tau) = \begin{cases} 0 & \text{if } \sigma \nsubseteq \tau \\ 1 & \text{if } \sigma \subseteq \tau \text{ and } \sigma \text{ is oriented coherently with } \tau \\ -1 & \text{if } \sigma \subseteq \tau \text{ and } \sigma \text{ has the opposite orientation of } \tau. \end{cases}$$

The columnns of $I^k(X)$ represent the boundaries of the k-faces in X. We can order the columns of $I^k(X)$ so it has the form

$$\begin{array}{cc} & \begin{array}{cc} \overbrace{\tilde{\sigma}_1 \ \dots \ \tilde{\sigma}_n}^{\substack{k\text{-faces} \\ \text{not in } T_0}} & \overbrace{\sigma_1 \ \dots \ \sigma_n}^{\substack{k\text{-faces} \\ \text{in } T_0}} \end{array} \\ I^k(X) = \text{(k-1)-faces} \left\{ \begin{array}{c} \tau_1 \\ \vdots \\ \tau_m \end{array} \right. & \left[\begin{array}{c|c} \epsilon(\tau_i, \tilde{\sigma_j}) & \epsilon(\tau_i, \sigma_j) \end{array} \right]. \end{array}$$

There is a one-to-one correspondence between the $(k-1)$-cells and $(k+1)$-cells of X given by $\tau \mapsto \tilde{\tau}$. Thus the transpose of the $(k+1)^{st}$ incidence matrix can be written as

$$\begin{array}{cc} & \begin{array}{cc} \overbrace{\tilde{\sigma}_1 \ \dots \ \tilde{\sigma}_n}^{\substack{k\text{-faces} \\ \text{not in } T_0}} & \overbrace{\sigma_1 \ \dots \ \sigma_n}^{\substack{k\text{-faces} \\ \text{in } T_0}} \end{array} \\ I^{k+1}(X)^T = \text{(k+1)-faces} \left\{ \begin{array}{c} \tilde{\tau}_1 \\ \vdots \\ \tilde{\tau}_m \end{array} \right. & \left[\begin{array}{c|c} \epsilon(\tilde{\sigma}_j, \tilde{\tau}_i) & \epsilon(\sigma_j, \tilde{\tau}_i) \end{array} \right]. \end{array}$$

Again, using an inductive argument as in [**9**, Theorem 65.1], we orient X and $D(X)$ in such a way that $\epsilon(\tau_i, \sigma_j) = \epsilon(\tilde{\sigma}_j, \tilde{\tau}_i)$ and $\epsilon(\tau_i, \tilde{\sigma_j}) = \epsilon(\sigma_j, \tilde{\tau}_i)$. Thus the matrices $I^k(X)$ and $I^{k+1}(X)^T$ are of the forms

$$\begin{array}{rcl} I^k(X) & = & [\ P \mid Q\ ], \\ I^{k+1}(X)^T & = & [\ Q \mid P\ ]. \end{array} \tag{3.1}$$

We show that there exists a matrix $R \in \mathbb{Z}^{n \times m}$ such that $RI^{k+1}(X)^T = [\ I \mid N\ ]$. We define the *reduced incidence matrices* $I_r^k(X) := RI^k(X)$ and $I_r^{k+1}(X)^T := RI^{k+1}(X)^T$. These matrices are of the forms

$$\begin{array}{rcl} I_r^k(X) & = & [\ N \mid I\ ] \ =: \ M \\ I_r^{k+1}(X)^T & = & [\ I \mid N\ ] \ =: \ M^{\perp}. \end{array}$$

Since $\partial_k \ \partial_{k+1} = 0$, we have Rowspace$(M)^{\perp} =$ Rowspace$(M^{\perp})$.

□

When k is even, we can form the matrices $I^k(X)$ and $I^{k+1}(X)^T$ as above. However, our method of orienting X and $D(X)$ now yields $\epsilon(\tau_i, \sigma_j) = \epsilon(\tilde{\sigma}_j, \tilde{\tau}_i)$ and $\epsilon(\tau_i, \tilde{\sigma_j}) = -\epsilon(\sigma_j, \tilde{\tau}_i)$. Thus the matrices $I^k(X)$ and $I^{k+1}(X)^T$ have the forms

$$\begin{array}{rcl} I^k(X) & = & [\ P \mid Q\ ] \\ I^{k+1}(X)^T & = & [\ Q \mid -P\ ], \end{array}$$

and the reduced incidence matrices $I_r^k(X)$ and $I_r^{k+1}(X)^T$ have the forms

$$\begin{array}{rcl} I_r^k(X) & = & [\ N \mid I\ ] \ =: \ M \\ I_r^{k+1}(X)^T & = & [\ I \mid -N\ ] \ =: \ M^{\perp}. \end{array}$$

With this orientation, X does not give rise to an involutively self-dual matroid and the concatenated matrix $A = \left[\frac{M}{M^{\perp}}\right]$ is symmetric rather than skew-symmetric, so the matroid results do not apply. Of course this does not preclude the possibility that a different method of orienting X and $D(X)$ could yield a version of Theorem 1.6 for even k. However, the fact that certain trees had to be excluded to give a similar formula for simplicial complexes when $k = 2$ makes it seem less promising (see [**5**, page 350]).

We conclude this section with the following analogue of Kalai's Lemma 2 [**5**]. The ideas of this proof are almost exactly the same as those in Kalai's proof.

LEMMA 3.2. *Let $k \geq 1$ and let X be an antipodally self-dual cell complex, which contains an acyclic, self-dual tree T_0. Then for each collection C of k-cells of X we have*

(1) $\det I_r^k(\overline{C}) = 0$ *if and only if* $\widetilde{H}_k(\overline{C}) \neq 0$,

(2) *If* $\widetilde{H}_k(\overline{C}) = 0$, *then* $|\det I_r^k(\overline{C})| = |\widetilde{H}_{k-1}(\overline{C})|$.

PROOF. The proof of (1) is exactly the same as Kalai's proof of [**5**, Lemma 2]. For (2), we first consider the case when $k > 1$. Since $\widetilde{H}_{k-1}(X) = \widetilde{H}_{k-1}(\mathbb{S}^{2k}) = 0$ and $X^{(k-1)} \subseteq \overline{C}$, we have $B_{k-1}(X) = Z_{k-1}(X) = Z_{k-1}(\overline{C})$. The columns of $I^k(X)$ represent $B_{k-1}(X)$, while the columns of $I^k(\overline{C})$ represent $B_{k-1}(\overline{C})$. Hence,

$$\widetilde{H}_{k-1}(\overline{C}) = I^k(X)\mathbb{Z}^{2n}/I^k(\overline{C})\mathbb{Z}^n \cong RI^k(X)\mathbb{Z}^{2n}/RI^k(\overline{C})\mathbb{Z}^n = I_r^k(X)\mathbb{Z}^{2n}/I_r^k(\overline{C})\mathbb{Z}^n \cong \mathbb{Z}^n/I_r^k(\overline{C})\mathbb{Z}^n,$$

where the last congruence uses the fact that $I_r^k(X) = [\ N \mid I\ ]$ contains an $n \times n$ identity matrix.

When $k = 1$, we use part (1) along with the standard facts from graph theory and topology that for a collection C of edges of a graph G

$$\det I_r^1(C) = \begin{cases} \pm 1 & \text{if } C \text{ is a tree} \\ 0 & \text{otherwise,} \end{cases}$$

$$|\widetilde{H}_0(\overline{C})| = \begin{cases} 1 & \text{if } \overline{C} \text{ is connected} \\ \infty & \text{otherwise.} \end{cases}$$

□

4. Applications and Further Results

In this section we first discuss a class of graphs called central reflexes. We apply the results from the previous sections to show that their spanning tree numbers are perfect squares and that their critical groups have a special form. Then we discuss a class of simplicial complexes and apply the previous results to solve a problem that was posed by Kalai (see [**5**, problem 3]).

4.1. Spanning Trees and Critical Groups of Central Reflexes. Central reflexes are a special class of directed, connected self-dual graphs on $\mathbb{S}^2$ for which the graph isomorphism sending G to G^* is the antipodal map $a : \mathbb{S}^2 \to \mathbb{S}^2$. Some examples of central reflexes include odd wheels embedded on $\mathbb{S}^2$. Figure 1 shows a 5-wheel on $\mathbb{S}^2$ and a planar representation of a 5-wheel. Another interesting class of central reflexes arises from squared rectangles and is described in [**11**].

When $k = 1$, the antipodally self-dual cell complexes are precisely the central reflexes with no loops and no isthmuses. The dual block complex $D(G)$ of a central reflex is just an embedding of the planar dual graph G^* on the sphere $\mathbb{S}^{2k}$. For each edge e, its dual block $D(e)$ is the edge e^* in G^* which crosses e and its conjugate $\tilde{e}$ is defined by $\tilde{e} := a(e^*)$. See Figure 1 for some examples of conjugate edges. Central reflexes can be oriented in such a way that the property $\tilde{\tilde{e}} = e$ holds. For each conjugate pair $\{e, \tilde{e}\}$, we arbitrarily orient one edge e. Its dual edge e^* is oriented so that it crosses e from right to left. Then, since $\tilde{e} = a(e^*)$, the orientation of $\tilde{e}$ is determined. Tutte [**12**, (3.4)] proves that the property $\tilde{\tilde{e}} = e$ holds.

A *self-dual spanning tree* of a central reflex G is a spanning tree that contains exactly one edge from each conjugate pair $\{e, \tilde{e}\}$. Equivalently, a spanning tree T is self-dual if $a((E(G)\backslash T)^*) = \{\tilde{e} : e \notin T\} = T$. An example of a self-dual spanning tree is given in Figure 1. We let $\mathcal{D}(G)$ denote the number of self-dual spanning trees of G.

An edge e is a loop in G if and only if e^* is an isthmus in G^*. Since the antipodal map a is a homeomorphism, it follows that e is a loop in G if and only if $\tilde{e}$ is an isthmus in G.

In this paper, we'll let $G\backslash e$ denote deletion of e from G and G/e denote contraction of G on e. Deleting a non-isthmus edge e in G corresponds to contracting its dual edge e^* in G^*. Likewise, contracting a non-loop edge e in G corresponds to deleting its dual edge e^* in G^*. Also, the self-dual spanning trees in $G\backslash\tilde{e}/e$ correspond to the self-dual spanning trees in G that contain e, while the self-dual spanning trees in $G\backslash e/\tilde{e}$ correspond to the self-dual spanning trees in G that contain $\tilde{e}$. Tutte uses these facts to prove the following proposition [**12**, (4.4) and (4.5)].

PROPOSITION 4.1. *If G is a central reflex and e is an edge of G that is neither a loop nor an isthmus, then $G\backslash\tilde{e}/e$ and $G\backslash e/\tilde{e}$ are central reflexes and*

$$\mathcal{D}(G) = \mathcal{D}(G\backslash\tilde{e}/e) + \mathcal{D}(G\backslash e/\tilde{e}).$$

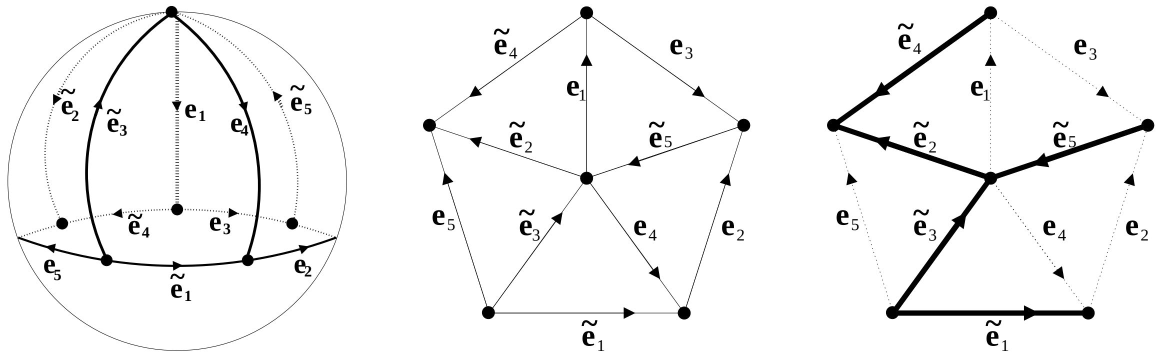


FIGURE 1. An example of a central reflex G on $\mathbb{S}^2$, a planar representation of G, and a self-dual spanning tree T.

We use this proposition and induction on the number of conjugate pairs that are not loop-isthmus pairs to prove the next lemma.

LEMMA 4.1. *If G is a central reflex, then G has at least one self-dual spanning tree.*

In [**12**], Tutte allows loops and isthmuses in central reflexes. The antipodally self-dual cell complexes are regular and hence cannot contain loops and isthmuses. However, every central reflex is equivalent to a regular central reflex in the following sense. Given a central reflex G, let G' be the graph that results from deleting all of the loops and contracting all of the isthmuses. By [**12**, (4.3)], G' is a central reflex. A spanning tree of G contains no loops and contains every isthmus, hence $\kappa(G) = \kappa(G')$.

Since $\widetilde{H}_0(T) = 0$ for any spanning tree T, Theorem 1.6 gives a new proof of Theorem 1.7.

The critical group $K(G)$ of a connected graph G is an abelian group of order $\kappa(G)$. The critical group has several equivalent interpretations. In this paper, we use the form

$$K(G) = \mathbb{Z}^{|E(G)|}/Z_1(G) \oplus B_0(G). \tag{4.1}$$

The formula $\kappa(G) = \mathcal{D}(G)^2$ suggests that the critical group of a central reflex[1] can be written as a direct sum of two copies of a group of order $\mathcal{D}(G)$. Using line (4.1) and Theorem 1.4, we obtain Proposition 1.2.

EXAMPLE 4.2. As noted above, n-wheels are central reflexes when n is odd. For an n-wheel G (with n odd), Biggs [**3**, Theorem 9.2] uses a variation of the chip-firing game to prove that

$$K(G) = \mathbb{Z}_{\ell_n} \oplus \mathbb{Z}_{\ell_n},$$

where ℓ_n is the n^{th} Lucas number.

As we discussed in Section 2, the matrix N controls the behavior of the critical group $K(G) = \mathrm{coker}(A)$. More specifically, Corollary 2.2 states that for $p \neq 2$,

$$\mathrm{Syl}_p(H) \cong \mathrm{Syl}_p(\mathrm{coker}(N + I)),$$

where $\mathrm{Syl}_p(G)$ denote the p-primary component of an abelian group G. The next example demonstrates that it is necessary to exclude $p = 2$ in this corollary.

EXAMPLE 4.3. The double 5-wheel is a central reflex formed by attaching another pentagon to the outside rim of the 5-wheel (see Figure 2). Computing the Smith normal forms of A and $N + I$ gives

$$\mathrm{coker}(A) = (\mathbb{Z}_4)^4 \oplus (\mathbb{Z}_{11})^2 \text{ and thus } H = (\mathbb{Z}_4)^2 \oplus \mathbb{Z}_{11},$$

while

$$\mathrm{coker}(N + I) = (\mathbb{Z}_2)^4 \oplus \mathbb{Z}_{11}.$$

[1]Using the presentation of the critical group $K(G) = \mathrm{coker}\overline{L(G)}$, where $\overline{L(G)}$ denotes the reduced Laplacian matrix, we see that $K(G) = K(G')$. This follows from the fact that deleting a loop has no effect on the Laplacian $L(G)$, while contracting an isthmus corresponds to performing elementary row and column operations on $L(G)$.

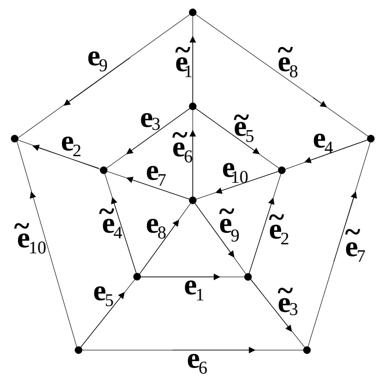


Figure 2. A double 5-wheel.

4.2. Simplicial Complexes. Let $\triangle$ denote the $(2k+1)$-dimensional simplex on the vertex set $V = \{v_0, \ldots, v_{2k+2}\}$. We identify the boundary of $\triangle$ with the sphere $\mathbb{S}^{2k}$ in the following way. We first identify $\triangle$ with the $(2k+1)$-simplex in $\mathbb{R}^{2k+1}$ which has vertices

$$v_0 = (0, 0, \ldots, 0), \quad v_1 = (1, 0, \ldots, 0), \quad v_2 = (0, 1, \ldots, 0), \quad \ldots, \quad v_{2k+2} = (0, 0, \ldots, 1).$$

We translate $\triangle$ so that its barycenter $\hat{\triangle}$ is at the origin, remove the interior of $\triangle$ and divide the points in the boundary of $\triangle$ by their lengths. Then, for each face F of X, the antipodal map sends $D(F)$ to F^c, i.e. $\tilde{F} = F^c$ when viewed as unoriented cells. When k is odd, X and $D(X)$ can be oriented in such a way that $\tilde{\tilde{F}} = F$ and X is an antipodally self-dual cell complex.

Let $\mathcal{T}(n, k)$ be the set of all simplicial complexes T on the vertex set $\{1, 2, \ldots, n\} = [n]$ such that

(1) T contains the complete $(k-1)$-skeleton,
(2) T has exactly $\binom{n-1}{k}$ k-faces,
(3) $H_k(T) = 0$.

Let X be the complete $2k$-dimensional complex on the vertex set $[2k+2]$ embedded on $\mathbb{S}^{2k}$. By [**5**, Proposition 2], we see that the definition of $\mathcal{T}(2k+2, k)$ agrees with the definition of $\mathcal{T}_k(X)$. The *blocker* or *Alexander dual* of a simplicial complex C is defined by $C^\vee := \{S \subseteq V : S^c \notin C\}$. A complex $T \in \mathcal{T}(2k+2, k)$ is said to be *self-dual* if $T^\vee = T$. Since

$$\widetilde{X \backslash T} = \{\widetilde{F} : F \notin T\} = \{F^c : F \notin T\} = \{F : F^c \notin T\} = T^\vee,$$

we see that this definition of self-dual complexes agrees with the definition of self-dual trees in Section 3.

Let C be the collection of all k-faces of X that contain vertex 1. We use the fact that vertex 1 is a cone point of $\overline{C}$ to prove the next lemma.

Lemma 4.4. *Let $r := \binom{2k+1}{k} = \binom{2k+1}{k+1}$ and let $C := \{F_1, \ldots, F_r\}$ be all of the k-faces of X that contain vertex* 1*. Then $\overline{C}$ is an acyclic, self-dual spanning tree in $\mathcal{T}_k(X)$.*

Sketch Proof of Theorem 1.9. Combining Lemma 4.4 and Theorem 1.6 gives the proof of the first assertion in Theorem 1.9. We now sketch a proof of the second assertion. Let $\overline{C}$ be the self-dual, acyclic spanning tree from Lemma 4.4. Kalai [**5**, page 342] shows that the reduced incidence matrix $I_r^k(X)$ can be formed from $I^k(X)$ by deleting the rows that correspond to $(k-1)$-faces containing vertex 1. Then $I_r^k(X)$ has rows indexed by the $(k-1)$-faces that don't contain vertex 1 and columns indexed by the k-faces not in $\overline{C}$ followed by the k-faces in $\overline{C}$ and is of the form $[\ N \mid I\ ]$. Also, $I_r^{k+1}(X)^T$ has rows indexed by the $(k+1)$-faces that do contain 1 and columns indexed by the k-faces not in $\overline{C}$ followed by the k-faces in $\overline{C}$ and is of the form $[\ I \mid N\ ]$.

Let A be the concatenated matrix

$$A = \left[\frac{I_r^k(X)}{-I_r^{k+1}(X)^T} \right] = \begin{bmatrix} N & I \\ -I & -N \end{bmatrix}.$$

Since $\partial_k \, \partial_{k+1} = 0$, $\text{Rowspace}(I_r^{k+1}(X)^T) = \text{Rowspace}(I_r^k(X))^\perp$. Thus N and hence A is skew-symmetric.

Example 4.5. For $k = 1$, we have

$$A = \begin{array}{c} \\ 2 \\ 3 \\ 4 \\ \hline +134 \\ -124 \\ +123 \end{array} \begin{array}{c} \begin{array}{ccc|ccc} -34 & +24 & -23 & +12 & +13 & +14 \end{array} \\ \left[\begin{array}{ccc|ccc} & -1 & +1 & +1 & & \\ +1 & & -1 & & +1 & \\ -1 & +1 & & & & +1 \\ \hline -1 & & & & +1 & -1 \\ & -1 & & -1 & & +1 \\ & & -1 & +1 & -1 & \end{array}\right] \end{array}.$$

We associate to each vertex i a weight x_i, and we form a weighted version $A(x)$ of our matrix A by setting

$$A(x)(\tau, \sigma) = A(\tau, \sigma) \cdot x_{\tau \triangle \sigma}.$$

Since $F_j \triangle (F_j \backslash \{1\}) = (\{1\} \cup F_j^c) \triangle F_j^c = 1$, the top right and bottom left blocks of $A(x)$ are $x_1 I$ and $-x_1 I$ respectively. If $n_{ij} \neq 0$, then $F_i \backslash \{1\} \subseteq F_j^c$ and

$$(F_i \backslash \{1\}) \triangle F_j^c = (F_j \backslash \{1\}) \triangle F_i^c.$$

Thus

$$n_{ij} \cdot x_{(F_i \backslash \{1\}) \triangle F_j^c} = -n_{ji} \cdot x_{(F_j \backslash \{1\}) \triangle F_i^c}$$

and it follows that $A(x)$ is skew-symmetric.

Example 4.6. For $k = 1$, we have

$$A(x) = \begin{array}{c} \\ 2 \\ 3 \\ 4 \\ \hline +134 \\ -124 \\ +123 \end{array} \begin{array}{c} \begin{array}{ccc|ccc} -34 & +24 & -23 & +12 & +13 & +14 \end{array} \\ \left[\begin{array}{ccc|ccc} & -x_4 & +x_3 & +x_1 & & \\ +x_4 & & -x_2 & & +x_1 & \\ -x_3 & +x_2 & & & & +x_1 \\ \hline -x_1 & & & & +x_4 & -x_3 \\ & -x_1 & & -x_4 & & +x_2 \\ & & -x_1 & +x_3 & -x_2 & \end{array}\right] \end{array}$$

The ideas of the rest of the proof are very similar to those in Theorem 1.2.

□

In Section 4.1 we discussed the critical groups of graphs. For the complete graph K_n, the critical group has the structure

$$K(K_n) \cong (\mathbb{Z}_n)^{n-2}$$

(see [**3**, Section 8]). The next proposition gives an analogous result for simplicial complexes.

Proposition 4.2. *Let K be the complete k-dimensional simplicial complex on $[n]$ and let* $A = \left[\frac{I_r^k(K)}{-I_r^{k+1}(K)^T}\right]$. *Then*

$$\operatorname{coker}(A) \cong (\mathbb{Z}_n)^{\binom{n-2}{k}}.$$

The proof of this proposition is divided into three steps:

(1) Prove that $\operatorname{coker}(A)$ is all n-torsion.
(2) Prove that $\operatorname{coker}(A)$ has a generating set of cardinality $\binom{n-2}{k}$.
(3) Finish the proof by using Kalai's result that $\det(A) = \det\left(I_r^k(K) I_r^k(K)^T\right) = n^{\binom{n-2}{k}}$.

We note that in the special case when $n = 2k+2$, Theorem 1.4 takes on the form

$$\operatorname{coker}(A) \cong H \oplus H,$$

where $H = (\mathbb{Z}_{2k+2})^{\binom{2k-1}{k}}$.

References

[1] H. Bai, On the Critical Group of the n-cube, 2003. Linear Algebra Appl., **369**, 251-261.
[2] N. Biggs, Algebraic potential theory on graphs, 1997. Bull. London Math. Soc., **29**, no. 6, 641-682.
[3] N. Biggs, Chip-Firing and the Critical Group of a Graph, 1999. Journal of Algebraic Combinatorics, **9**, 25-45.
[4] B. Jacobson, A. Neidermaier, V. Reiner, Critical groups for complete multipartite graphs and Cartesian products of complete graphs, 2003. J. Graph Theory, **44**, no. 3, 231-250.
[5] G. Kalai, Enumeration of $\mathbb{Q}$-Acyclic Simplicial Complexes, 1983. Israel Journal of Mathematics, **45**, 337-351.
[6] G. Kuperberg, Kasteleyn Cokernels, 2002. Electronic Journal of Combinatorics, **9**, no. 1, R29.
[7] S. Lang, Algebra, Springer-Verlag, 3^{rd} Edition, 2002.
[8] A. Lundell and S. Weingram, The Topology of CW Complexes, Litton Educational Publishing, Inc., 1969.
[9] J. Munkres, Elements of Algebraic Topology, Perseus Books Publishing, L.L.C., 1984.
[10] V. Reiner, Lectures on Matroids and Oriented Matroids, 2005. Available at `http://www.math.umn.edu/~reiner/Talks/Vienna05/index.html`.
[11] C.A.B. Smith and W.T. Tutte, A Class of Self-Dual Maps, 1950. Canadian Journal of Mathematics, **2**, 179-196.
[12] W.T. Tutte, On the Spanning Trees of Self-Dual Maps, 1979. Annals New York Academy of Sciences, 540-548.

School of Mathematics, 206 Church St. S.E., University of Minnesota, Minneapolis, MN 55455
E-mail address: maxwell@math.umn.edu

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Counting unrooted hypermaps on closed orientable surface

Alexander Mednykh and Roman Nedela

Abstract. In this paper we derive an enumeration formula for the number of hypermaps of given genus g and given number of darts n in terms of the numbers of rooted hypermaps of genus $\gamma \leq g$ with m darts , where $m|\, n$.

Résumé. Dans ce travail on denombre les hypergraphes d'un genus g donné, et un nombr e de fleches n, selon le nombre de hyper cartes de genus $\gamma \leq g$, et av ec m fleches, ou $m|\, n$.

1. Introduction

In this paper we derive an enumeration formula for the number of hypermaps of given genus g and given number of darts n in terms of the numbers of rooted hypermaps of genus $\gamma \leq g$ with m darts , where $m|\, n$. Explicit expressions for the number of rooted hypermaps of genus g with n darts were derived by Walsh [**32**] for $g = 0$, and by Arques [**2**] for $g = 1$. We apply our general counting formula to derive explicit expressions for the number of unrooted spherical and toroidal hypermaps with given number of darts.

Oriented map is 2-cell decomposition of a closed orientable surface with a fixed global orientation. Generally, maps can be described combinatorially via graph embeddings. Oriented hypermaps are generalisations of oriented maps. While maps are 2-cell embeddings of graphs, hypermaps can be viewed as embeddings of hypermaps into closed orientable surfaces. Such a model was investigated by Walsh in [**32**], where the underlying hypergraph is described via the corresponding 2-coloured bipartite graph B, and the hypermap itself is determined by a 2-cell embedding $B \to S$.

Beginnings of the enumerative theory of maps are closely related with the enumeration of plane trees considered in 60-th by Tutte [**28**], Harary, Prins and Tutte [**6**], see [**7, 22**] as well. Later a lot of other distinguished classes of maps including triangulations, outerplanar, cubic, Eulerian, nonseparable, simple, looples, two-face maps and others were considered. Enumeration of maps on surfaces has attracted a lot of attention last decades [**23**]. Although there are more than 100 published papers on map enumeration most of them deal with the enumeration of rooted maps of given property. In particular, there is a lack of results on enumeration of unrooted maps of genus ≥ 1. Most of the results on map enumeration in the unrooted case restrict to planar maps [**17, 18, 33, 34, 20**]. A recent paper [**25**] presents a breakthrough in the enumeration problem for unrooted maps on closed oriented surface. In the presented paper we apply the methods employed in [**24**] and [**25**] to solve an analogous problem for hypermaps.

2. Hypermaps on surfaces and orbifolds

Hypermaps on surfaces. An *oriented combinatorial hypermap* is a triple $\mathcal{H} = (D; R, L)$, where D is a finite set of darts (called brins, blades, bits as well) and R, L are permutations of D such that $\langle R, L\rangle$ is transitive on D. Orbits of R are called hypervertices, orbits of L are called hyperedges and orbits of RL are called hyperfaces. The degree of a hypervertex (hyperedge, hyperface) is the size of the respective orbit.

Key words and phrases. Enumeration, Map, Surface, Orbifold, Rooted hypermap, Unrooted hypermap, Fuchsian group.

Let $|D| = n$. Denote by v, e and f the numbers of hypervertices, hyperedges and hyperfaces. Then genus g of $\mathcal{H}$ is given by Euler-Poincare formula as follows

$$v + e + f - n = 2 - 2g.$$

Given hypermaps $\mathcal{H}_i = (D_i; R_i, L_i)$, $i = 1, 2$ a mapping $\psi : D_1 \to D_2$ such that $R_2\psi = \psi R_1$ and $L_2\psi = \psi L_1$ is called a morphism (or a covering) $\mathcal{H}_1 \to \mathcal{H}_2$. Note that each morphism between hypermaps is by definition an epimorphism. If $\psi : \mathcal{H}_1 \to \mathcal{H}_2$ is a bijection, ψ is an isomorphism. Isomorphisms $\mathcal{H} \to \mathcal{H}$ form a group $\mathrm{Aut}(\mathcal{H})$ of automorphisms of $\mathcal{H}$. It is easily seen that $\mathrm{Aut}(\mathcal{H})$ acts semiregularly on D, equivalently, the stabiliser of a dart is trivial. A hypermap $\mathcal{H}$ is called *rooted* if one element x of D is chosen to be a root. Morphisms between rooted hypermaps take roots onto roots. It follows that a rooted hypermap admits no non-trivial automorphisms.

By *a surface* we mean a connected, orientable surface without boundary. A *topological map* is a 2-cell decomposition of a surface. Standardly, maps on surfaces are described as 2-cell embeddings of graphs. Oriented combinatorial maps are hypermaps $(D; R, L)$ such that L is a fixed-point-free involution. Walsh observed that oriented hypermaps can be viewed as particular maps. Namely, he proved a one-to-one correspondence [**32**, Lemma 1] between hypermaps and the set of (oriented) 2-coloured bipartite maps. That means that one of the two global orientations of the underlying surface is fixed, and moreover, we assume that a colouring of vertices, say by black and white colours, is preserved by morphisms between maps. The correspondence is given as follows. Let $\mathcal{M}$ 2-coloured bipartite map on an orientable surface S with a fixed global orientation. We set D to be the set of edges of $\mathcal{M}$. The orientation of S induces at each black vertex v of $\mathcal{M}$ a cyclic permutation R_v of edges incident with v. This way a permutation $R = \prod R_v$ of D is defined. Similarly, the orientation of S determines at each white vertex u a cyclic permutation L_u. Set $L = \prod L_u$. Hence we have a unique hypermap $(D; R, L)$ corresponding to $\mathcal{M}$. Conversely, given hypermap $(D; R, L)$ we first define a bipartite 2-colored graph X whose edges are elements of D, black vertices are orbits of R and white vertices are orbits of L. An edge $x \in D$ is incident to a (black or white) vertex u if $x \in u$. The permutation R and L induce local rotations of arcs outgoing from black and white vertices, respectively. It is well known (see Gross and Tucker [**5**, Section 3.2]) that the system of rotations determines a 2-cell embedding of X into an orientable surface.

Similarly as above, an oriented 2-coloured bipartite map is called *rooted* if one of the edges is selected to be a root. Morphisms between rooted 2-coloured bipartite maps take a root onto a root.

There is yet another way to describe hypermaps. Let $\mathcal{H} = (D; R, L)$ be a hypermap. Clearly, the permutation group $\langle R, L\rangle$ is an epimorphic image of the free product $\Delta^+ = C * C \cong \langle \rho \rangle * \langle \lambda \rangle$ of two infinite cyclic groups. The group Δ^+ acts on D via epimorphism taking $\rho \mapsto R$ and $\lambda \mapsto L$. Thus using some standard considerations in permutation group theory each hypermap can be described by a subgroup $F \le \Delta^+$ [**13, 30, 31, 9**]. The subgroup F, called a *hypermap subgroup*, can be identified with a stabiliser of a dart in the action of Δ^+ on D. Since the action of Δ^+ on D is transitive, the number of darts $|D| = n$ coincides with index $[\Delta^+ : F]$ of F in Δ^+. Given $F \le \Delta^+$ the corresponding hypermap can be constructed as an *algebraic hypermap* $\mathcal{H}(\Delta^+/F) = (D; R, L)$, where $D = \{xF | x \in \Delta^+\}$ is the set of left cosets, and the action of R, L on D is defined by $R(xF) = (\rho x)F$, $L(xF) = (\lambda x)F$. Note that the group Δ^+ is sometimes called a universal oriented triangle group. More precisely, Δ^+ is identified with the triangle group $T(\infty, \infty, \infty) =< x, y, z : x\,y\,z = 1 >$ acting on the hyperbolic plane $\mathbf{H}^2$ by orientation preserving isometries (see G.Jones, D.Singerman [**13**]). In this case $\mathbf{H}^2/\Delta^+$ is a trice punctured sphere and $\mathbf{H}^2/F$ is a punctured orientable surface, whose genus g coincides with the genus of the corresponding hypermap.

We summarise the above discussion in the following propositions.

Proposition 2.1. *The following objects are in one-to-one correspondence:*

(1) *rooted 2-coloured bipartite maps of genus g with n edges,*
(2) *rooted hypermaps $(D; R, L)$ of genus g with $|D| = n$,*
(3) *subgroups of the group $\Delta^+ = T(\infty, \infty, \infty)$ of index n and genus g .*

Part (1) $\Leftrightarrow$ (2) follows from Walsh [**32**]. Part (2) $\Leftrightarrow$ (3) is in ([**13, 4**]).

By definition isomorphic hypermaps have conjugated hypermap subgroups. Hence isomorphism classes of hypermaps correspond to conjugacy classes of subgroups.

Proposition 2.2. *The following objects are in one-to-one correspondence:*

(1) *isomorphism classes of 2-coloured bipartite maps of genus g with n edges,*

(2) *isomorphism classes of hypermaps* $(D; R, L)$ *of genus* g *with* $|D| = n$,
(3) *conjugacy classes subgroups of index* n *and genus* g *of the group* $\Delta^+ = T(\infty, \infty, \infty)$.

Regular coverings. Let $\psi : \mathcal{H}_1 \to \mathcal{H}_2$ be a covering of hypermaps. The covering transformation group consists of automorphisms α of $\mathcal{H}_1$ satisfying the condition $\psi = \psi \circ \alpha$. A covering $\psi : \mathcal{H}_1 \to \mathcal{H}_2$ will be called *regular* if the covering transformation group acts transitively on a fibre $\psi^{-1}(x)$ over a dart x of $\mathcal{H}_2$. Regular coverings can be constructed by taking a subgroup $G \leq Aut(\mathcal{H}_1)$, $\mathcal{H}_1 = (D; R, L)$, and setting $\bar{D}$ to be the set of orbits of G, $\bar{R}[x] = [Rx]$, $\bar{L}[x] = [Lx]$. Then the natural projection $x \mapsto [x]$ defines a regular covering $M \to N$, where $\mathcal{H}_2 = (\bar{D}, \bar{R}, \bar{L})$.

Maps and hypermaps on orbifolds. Given regular covering $\psi : \mathcal{H} \to \mathcal{K}$, let be x be a hypervertex, hyperface or a hyperedge of $\mathcal{K}$. Let $\mathcal{H}$ be of genus g, $\mathcal{K}$ be of genus γ and let $G \leq \mathrm{Aut}(\mathcal{H})$ be a covering transformation group. The ratio of degrees $b(x) = deg(\tilde{x})/deg(x)$, where $\tilde{x} \in \psi^{-1}(x)$ is a lift of x along ψ, will be called a *branch index* of x. By transitivity of the action of the group of covering transformations a branch index is a well-defined positive integer not depending on the choice of the lift $\tilde{x}$. Hence b is a well defined integer function defined on the union $V(\mathcal{K}) \cup E(\mathcal{K}) \cup F(\mathcal{K})$. Writing all the values $b(x)$, $b(x) \geq 2$ in a non-decreasing order we get an integer sequence $m_1, m_2, \ldots, m_r$. This way an orbifold S_g/G with signature $[\gamma; m_1, m_2, \ldots, m_r]$ is defined.

For our purposes we define a topological 2-dimensional orbifold $O = O[\gamma; m_1, \ldots, m_r]$ to be a closed orientable surface of genus γ with a distinguished set of points $\mathcal{B}$, called branch points, and an integer function assigning to each $x \in \mathcal{B}$ an integer $b(x) \geq 2$. A 2-coloured bipartite map of genus γ is a map on O provided the following two conditions are satisfied:

(1) no branch point $x \in \mathcal{B}$ lies on an edge,
(2) each face contains at most one branch point $x \in \mathcal{B}$.

The operation associating a 2-coloured bipartite map to a hypermap is functorial. In particular the signature of an orbifold associated with a regular covering of hypermaps coincides with the signature of an orbifold determined by the corresponding regular covering of Walsh 2-coloured bipartite maps. Note also that a regular covering $\psi : \mathcal{H} \to \mathcal{K}$, extends (uniquely) to a regular covering $S_g \to S_g/G$, where g is genus of $\mathcal{H}$ and G is the group of covering transformations.

Let O be an orbifold with signature $[\gamma; m_1, m_2, \ldots, m_r]$. The *orbifold fundamental group* $\pi_1(O)$ is an F-group

$$\pi_1(M, \sigma) = F[\gamma; m_1, m_2, \ldots, m_r] =$$
$$\langle a_1, b_1, a_2, b_2, \ldots, a_\gamma, b_\gamma, e_1, \ldots, e_r | \prod_{i=1}^{\gamma}[a_i, b_i] \prod_{j=1}^{r} e_j = 1, e_1^{m_1} = \ldots e_r^{m_r} = 1\rangle. \tag{2.1}$$

Let $\mathcal{H} \to \mathcal{H}/G = \mathcal{K}$ be a regular covering between hypermaps with a covering transformation group G, let $\mathcal{H}$ be finite. Let the the signarure of the orbifold $\mathcal{K} = \mathcal{H}/G$ be $[\gamma; m_1, m_2, \ldots, m_r]$. Then the Euler characteristic of the underlying surface of $\mathcal{H}$ is given by the Riemann-Hurwitz equation:

$$\chi = |G|\big(2 - 2\gamma - \sum_{i=1}^{r}(1 - \frac{1}{m_i})\big). \tag{2.2}$$

3. General counting formula.

The following theorem is the main result of [**24**].

Theorem 3.1. *Let* Γ *be a finitely generated group. Then the number of conjugacy classes of subgroups of index* n *in the group* Γ *is given by the formula*

$$N_\Gamma(n) = \frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{\substack{K < \Gamma \\ [\Gamma:K] = m}} Epi(K, Z_\ell).$$

In fact, a little modification of the proof allows us to generalise the above statement to subsets of subgroups of given index closed under conjugacy. Let $\mathcal{P}$ be a set of subgroups of a finitely generated group Γ closed under conjugation. By $Epi_{\mathcal{P}}(K, Z_\ell)$ we denote the number of epimorphisms $K \to Z_\ell$ with the kernel in $\mathcal{P}$.

Hence we have the following

Theorem 3.2. *Let Γ be a finitely generated group and $\mathcal{P}$ is a set of subgroups of Γ closed under conjugation. Then the number of conjugacy classes of subgroups of index n in $\mathcal{P}$ is given by the formula*

$$N_{\Gamma}^{\mathcal{P}}(n) = \frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{\substack{K < \Gamma \\ [\Gamma : K] = m}} Epi_{\mathcal{P}}(K, Z_\ell).$$

A group epimorphism is called *order preserving* if it preserves the orders of elements of finite order. Given closed orientable surface S_g of genus g and a cyclic orbifold $O = S_g/Z_\ell$ we denote by $Epi_0(\pi_1(O), Z_\ell)$ the number of order preserving epimorphisms $\pi_1(O) \to Z_\ell$.

The following result is the main tool to calculate the number of unrooted hypermaps on a closed oriented surface.

Theorem 3.3. *Let S_g be a closed orientable surface of genus g. Denote by $h_O(m)$ be the number of rooted hypermaps with m darts on a cyclic orbifold $O = S_g/Z_\ell$.*

Then the number of unrooted hypermaps of genus g having n darts is

$$H_g(n) = \frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{O \in Orb(S/Z_\ell)} h_O(m) Epi_0(\pi_1(O), Z_\ell),$$

where the second sum runs through all admissible cyclic orbifolds S_g/Z_ℓ.

Proof. Given $S = S_g$ let $\mathcal{P} = \mathcal{P}_g$ be the set subgroups of genus g of $\Delta^+ = T(\infty, \infty, \infty)$. By Propositions 2.1 and 2.2 rooted hypermaps on S correspond subgroups in $\mathcal{P}$, and isomorphism classes of unrooted hypermaps on S correspond to conjugacy classes of subgroups in $\mathcal{P}$. Setting $\Gamma = \Delta^+$ in Theorem 3.2 we get

$$H_g(n) = N_{\Delta^+}^{\mathcal{P}}(n) = \frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{\substack{K < \Delta^+ \\ [\Delta^+ : K] = m}} Epi_{\mathcal{P}}(K, Z_\ell).$$

Given epimorphism $\psi : K \to Z_\ell$ with kernel $H \in \mathcal{P}$ determines a regular covering of algebraic hypermaps $\psi^* : \mathcal{H}(\Delta^+/H) \to \mathcal{H}(\Delta^+/K)$ induced by $H \trianglelefteq K$ with the group of covering transformations isomorphic to Z_ℓ. Let σ be the signature of the orbifold $O = O(\sigma) = S_g/Z_\ell$ determined by the covering of hypermaps. Hence the set of epimorphisms $\psi : K \to Z_\ell$ with $Ker(\psi) = H \in \mathcal{P}$ split into classes characterised by the signatures of the cyclic orbifolds $O = S/Z_\ell$. Denote by $Epi_\sigma(K, Z_\ell)$ the number of epimorphisms $K \to Z_\ell$ with kernel $H \in \mathcal{P}$ and quotient orbifold $O = S/Z_\ell$ with signature σ. We set $\mathcal{P}_\sigma = \{K | K < \Delta^+, Epi_\sigma(K, Z_\ell) \neq 0\}$.

It is well known that the group Δ^+ acts on the universal covering surface $\mathcal{H}^2$ as a discontinuous group of conformal automorphisms. This allows us to introduce the structure of Riemann surface (as well as the orbifold structure) on the hypermaps $\mathcal{H}(\Delta^+/H)$, $\mathcal{H}(\Delta^+/K)$, respectively. A regular covering of hypermaps $\psi : \mathcal{H}(\Delta^+/H) \to \mathcal{H}(\Delta^+/K)$ extends to a branched regular covering $S \to O$ of the orbifold $O = O(\sigma)$ by the closed surface S. By the Riemann Extension Theorem there is a one-to-one correspondence between coverings $\mathcal{H}^2/H \to \mathcal{H}^2/K$ and coverings of the compactified quotient spaces $S = \overline{\mathcal{H}^2/H} \to O = \overline{\mathcal{H}^2/K}$ (see [**12**] for a more detailed explanation). We want to show $Epi_\sigma(K, Z_\ell) = Epi_0(\Gamma(\sigma), Z_\ell)$. Given $K \in \mathcal{P}_\sigma$ we calculate the number of regular Z_ℓ-coverings $\mathcal{H}^2/H \to \mathcal{H}^2/K$ with $H \trianglelefteq K$ and $H \in \mathcal{P}$. By G. Jones [**11**] there are $Epi_\sigma(K, Z_\ell)/\varphi(\ell)$ such coverings. On the other hand, we have $Epi_0(\Gamma(\sigma), Z_\ell)/\varphi(\ell)$ of regular Z_ℓ-coverings $S = \overline{\mathcal{H}^2/H} \to O = \overline{\mathcal{H}^2/K}$ over the orbifold $O = O(\sigma)$ with the signature σ [**11**]. By virtue of the one-to-one correspondence these numbers coincide. Hence, we have $Epi_\sigma(K, Z_\ell) = Epi_0(\Gamma(\sigma), Z_\ell)$ as it was required. Given m, ℓ and σ denote by $\nu_\sigma(m)$ the number of subgroups $K < \Delta^+$ in $\mathcal{P}(\sigma)$ and by $Sign(S_g/Z_\ell)$ the set of signatures of cyclic g-admissible orbifolds. We have

$$H_g(n) = \frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{\substack{K < \Delta^+ \\ [\Delta^+ : K] = m}} Epi_{\mathcal{P}}(K, Z_\ell) = \frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{\sigma \in Sign(S_g/Z_\ell)} \nu_\sigma(m) Epi_\sigma(K, Z_\ell) =$$

$$\frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{\sigma \in Sign(S_g/Z_\ell)} \nu_\sigma(m) Epi_0(\Gamma(\sigma), Z_\ell).$$

Taking into the account the correspondence between groups in $\mathcal{P}_\sigma$ and rooted hypermaps on the orbifold $O = O(\sigma)$ we get $\nu_\sigma(m) = h_O(m)$ and the proof is complete.

In what follows we derive a formula enumerating numbers of rooted hypermaps on orbifolds in terms of numbers of rooted hypermaps on surfaces. Let $\mathcal{H}$ be a rooted hypermap on an orbifold O such that $\mathcal{H} = \tilde{\mathcal{H}}/Z_\ell = (D; R, L)$ is a quotient of an ordinary finite map $\tilde{\mathcal{H}}$ on a surface S_g. Thus $O = S_g/G$ where $G \cong Z_\ell$ is a cyclic group of orientation preserving symmetries of S_g of order ℓ. It follows that each branch index of the branched covering $S_g \to O$ is a divisor of ℓ and can write $O = O[\gamma; 2^{q_2}, \ldots, \ell^{q_\ell}]$, where $q_i \geq 0$ denotes the number of branch points of index i, for $i = 2, \ldots, \ell$. In this case, genera γ and g are related by the Riemann-Hurwitz equation $2 - 2g = \ell(2 - 2\gamma - \sum_{j=2}^{\ell} q_j(1 - 1/j))$. We use the convention $h_\gamma(m) = \nu_{[\gamma;\,\emptyset]}(m)$ denoting the number of rooted hypermaps with m darts on a surface of genus g. Clearly, the exponential notation $O = O[\gamma; 2^{q_2}, \ldots, \ell^{q_\ell}]$ can be used for any oriented orbifold (not necessarily cyclic) provided the indexes of branch points are bounded by ℓ.

Given integers $x_1, x_2, \ldots, x_q$ and $y \geq x_1 + x_2 + \cdots + x_q$ we denote by

$$\binom{y}{x_1, x_2, \ldots, x_q} = \frac{y!}{x_1! x_2! \ldots x_q! (y - \sum_{j=1}^{q} x_j)!},$$

the multinomial coefficient.

Now we are able to determine the number of rooted hypermaps on an arbitrary orbifold.

Proposition 3.4. *The number of rooted hypermaps on an orbifold $O = O[\gamma; 2^{q_2}, \ldots, \ell^{q_\ell}]$ with m darts is*

$$h_O(m) = \binom{m+2-2\gamma}{q_2, q_3, \ldots, q_\ell} h_\gamma(m). \tag{5.1}$$

Proof. Let $\mathcal{H}$ be a rooted hypermap on S_γ with v hypervetices, e hyperedges and f hyperfaces. Then $\mathcal{H}$ gives rise to as many rooted hypermaps as is the number of partitions of the set $V(\mathcal{H}) \cup E(\mathcal{H}) \cup F(\mathcal{H})$ of cardinality $v + e + f = m + 2 - 2\gamma$ into disjoint subsets of cardinalities $q_1, q_2, \ldots, q_\ell$. This is exactly the number

$$\binom{m + 2 - 2\gamma}{q_2, q_3, \ldots, q_\ell}.$$

Combining Proposition 3.4 and Theorem 3.3 we get our main theorem.

Theorem 3.5. *The number of unrooted hypermaps on a closed surface S_g of genus g with n darts is given by*

$$H_g(n) = \frac{1}{n} \sum_{\substack{\ell | n \\ \ell m = n}} \sum_{\substack{O \in Orb(S/Z_\ell) \\ O = O[\gamma; 2^{q_2}, 3^{q_3}, \ldots, \ell^{q_\ell}]}} Epi_0(\pi_1(O), Z_\ell) \binom{m+2-2\gamma}{q_2, q_3, \ldots, q_\ell} h_\gamma(m),$$

where the second sum runs through all cyclic orbifolds S_g/Z_ℓ.

Note that the numbers $Epi_0(\pi_1(O), Z_\ell)$ were computed by the authors in [**25**] in terms of some standard arithmetical functions. The following section surveys results on $Epi_0(\pi_1(O), Z_\ell)$.

4. Number of epimorphisms from an F-group onto a cyclic group

As one can see in Theorems 3.3 and 3.5 to derive an explicit formula for the number of unrooted hypermaps with given genus and given number of darts one needs to deal with the numbers $Epi_0(\pi_1(O), \mathsf{Z}_\ell)$ of order preserving epimorphisms from an F-group Γ onto a cyclic group Z_ℓ. These numbers are counted using some number theoretical machinery in [**25**]. In what follows we recall some relevant results used in later computations.

Denote by $\mu(n)$, $\phi(n)$ and $\Phi(x, n)$ the Möbius, Euler and von Sterneck functions, respectively. The relationship between them is given by the formula

$$\Phi(x, n) = \frac{\phi(n)}{\phi(\frac{n}{(x,n)})} \, \mu\left(\frac{n}{(x, n)}\right),$$

where (x, n) is the greatest common divisor of x and n. It was shown by O. Hölder that $\Phi(x, n)$ coincides with the Ramanujan sum $\sum_{\substack{1 \leq k \leq n \\ (k,\, n)=1}} \exp(\frac{2\,ikx}{n})$. For the proof, see Apolstol [**1**, p.164] and [**26**]. An arithmetic function, called by Liskovets *orbicyclic arithmetic function* [**21**], is a multivariate integer function defined by

$$E(m_1, m_2, \ldots, m_r) = \frac{1}{m} \sum_{k=1}^{m} \Phi(k,\, m_1) \cdot \Phi(k,\, m_2) \ldots \Phi(k,\, m_r).$$

Recall that the Jordan multiplicative function $\phi_k(n)$ of order k can be defined as follows:

$$\phi_k(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) d^k.$$

The following proposition is proved in [**25**].

Proposition 4.1. *Let $\Gamma = F[g; m_1, \ldots, m_r]$ be an $F-$group of signature $[g; m_1, \ldots, m_r]$. Denote by $m = lcm\,(m_1, \ldots, m_r)$ the least common multiple of $m_1, \ldots, m_r$ and let $m|\ell$. Then the number of order-preserving epimorphisms of the group Γ onto a cyclic group Z_ℓ is given by the formula*

$$Epi_0(\Gamma, Z_\ell) = m^{2g} \phi_{2g}(\ell/m) E(m_1, m_2, \ldots, m_r).$$

In particular, if $\Gamma = F[g; \emptyset] = F[g; 1]$ is a surface group of genus g we have

$$Epi_0(\Gamma, Z_\ell) = \phi_{2g}(\ell).$$

Let us note that the condition $m|\ell$ in the above proposition gives no principal restriction, since $Epi_0(\Gamma, \mathsf{Z}_\ell) = 0$ by the definition provided m does not divide ℓ. An orbifold $O = O[g; m_1, \ldots, m_r]$ will be called γ*-admissible* if it can be represented in the form $O = S_\gamma/\mathsf{Z}_\ell$, where S_γ is an orientable surface of genus γ surface and Z_ℓ is a cyclic group of automorphisms of S_γ. There is an orbifold $O = S_\gamma/\mathsf{Z}_\ell$ with signature $[g; m_1, m_2, \ldots, m_r]$ if and only if there exists ℓ such that the number $Epi_0(\pi_1(O), Z_\ell) \neq 0$ and the numbers γ, g, $m_1, \ldots, m_r$ and ℓ are related by the Riemann-Hurwitz equation $2 - 2\gamma = \ell(2 - 2g - \sum_{i=1}^{r}(1 - 1/m_i))$. The Wiman theorem makes us sure that $1 \leq \ell \leq 4\gamma + 2$ for $\gamma > 1$.

Using Proposition 4.1 and result by Harvey [**8**]we derive the following lists of γ-admissible orbifolds, for $\gamma = 0, 1$.

Corollary 4.2. 0*-admissible orbifolds are $O = O[0; \ell^2]$, with $Epi_0(\pi_1(O), Z_\ell) = \phi(\ell)$ for any positive integer ℓ.*

Corollary 4.3. *Let $O = O[g; m_1, m_2, \ldots, m_r] = S_1/Z_\ell$ be a 1-admissible orbifold. Then one of the following cases happens:*

$O = O[1; \emptyset]$, *with* $Epi_0(\pi_1(O), Z_\ell) = \sum\limits_{k|\ell} \mu(\ell/k)k^2 = \phi_2(\ell)$ *for any* ℓ,
$\ell = 2$ *and* $O = O[0; 2^4]$, *with* $Epi_0(\pi_1(O), Z_\ell) = 1$,
$\ell = 3$ *and* $O = O[0; 3^3]$, *with* $Epi_0(\pi_1(O), Z_\ell) = 2$,
$\ell = 4$ *and* $O = O[0; 4^2, 2]$, *with* $Epi_0(\pi_1(O), Z_\ell) = 2$,
$\ell = 6$ *and* $O = O[0; 6, 3, 2]$, *with* $Epi_0(\pi_1(O), Z_\ell) = 2$.

The lists of $2-$ and $3-$admissible orbifolds can be found in [**25**].

5. Counting unrooted hypermaps on the sphere and torus

In this section we apply the above results to calculate the number of unrooted hypermaps with given number of darts on the sphere and torus.

Theorem 5.1. *The number of spherical unrooted hypermaps with n darts is given by the formula*

$$H_0(n) = \frac{1}{n}\left(\frac{3 \cdot 2^{n-1}}{(n+1)(n+2)} \binom{2n}{n} + \sum_{\substack{\ell|n,\, \ell>1 \\ \ell m = n}} 3 \cdot 2^{m-2} \binom{2m}{m} \phi(\ell) \right)$$

Proof. For $\ell > 1$ there is only one possible action of cyclic group Z_ℓ on the sphere S. The corresponding orbifold O has a signature $[0; \ell,\, \ell]$ and by Corollary 4.2 we have $Epi_0(\pi_1(O), Z_\ell) = \phi(\ell)$. By Theorem 3.5 we obtain

$$H_0(n) = \frac{1}{n}\big(h_0(n) + \sum_{\substack{\ell|n,\, \ell>1 \\ \ell m = n}} \phi(\ell) \binom{m+2}{2} h_0(m)\big).$$

To finish the proof we note that by T. Walsh [**32**]

$$h_0(m) = \frac{3 \cdot 2^{m-1}}{(m+1)(m+2)}\binom{2m}{m}. \tag{5.1}$$

The numbers of rooted and unrooted spherical hypermaps up to 30 darts is given in Table 1.
Table 1. Numbers of rooted and unrooted hypermaps on the sphere with at most 30 darts
No. of darts, rooted hypermaps, unrooted hypermaps

01, 1, 1
02, 3, 3
03, 12, 6
04, 56, 20
05, 288, 60
06, 1584, 291
07, 9152, 1310
08, 54912, 6975
09, 339456, 37746
10, 2149888, 215602
11, 13891584, 1262874
12, 91287552, 7611156
13, 608583680, 46814132
14, 4107939840, 293447817
15, 28030648320, 1868710728
16, 193100021760, 12068905911
17, 1341536993280, 78913940784
18, 9390758952960, 521709872895
19, 66182491668480, 3483289035186
20, 469294031831040, 23464708686960
21, 3346270487838720, 159346213738020
22, 23981605162844160, 1090073011199451
23, 172667557172477952, 7507285094455566
24, 1248519259554840576, 52021636161126702
25, 9063324995286990848, 362532999811480604
26, 66032796394233790464, 2539722940697502966
27, 482722511571640123392, 17878611539691757938
28, 3539965084858694238208, 126427324476844560112
29, 26035872237025235042304, 897788697828456380772
30, 192014557748061108436992, 6400485258395785352796

We note that the numbers $H_0(n)$ was determined in terms of unrooted planar 2-constellations formed by n polygons by M. Bosquet-Melon and G. Schaeffer [**3**].

Now we derive an explicit formula for counting unrooted maps on torus. Rooted toroidal maps were enumerated by D. Arquès in [**2**]. He proved that

$$h_1(n) = \frac{1}{3}\sum_{k=0}^{n-3} 2^k(4^{n-2-k}-1)\binom{n+k}{k}. \tag{5.2}$$

Theorem 5.2. *The number of unrooted toroidal hypermaps $H_1(n)$ with n darts is equal to*

$$\tfrac{1}{n}\Big(\tbinom{\frac{n}{2}+2}{4}h_0\left(\tfrac{n}{2}\right) + 2\tbinom{\frac{n}{3}+2}{3}h_0\left(\tfrac{n}{3}\right) + 6\tbinom{\frac{n}{4}+2}{3}h_0\left(\tfrac{n}{4}\right) + 12\tbinom{\frac{n}{6}+2}{3}h_0\left(\tfrac{n}{6}\right) + \sum_{\substack{\ell \mid n\\ \ell m=n}} \phi_2(\ell)h_1(m)\Big),$$

where ϕ_2 is the Jordan function, and functions h_0 and h_1 are given by (5.1) and (5.2), respectively.

Proof. Following Theorem 3.5 and Corollary 4.3 we have

$$H_1(n) = \frac{1}{n}\big(h_{[0;2^4]}(n/2) + 2h_{[0;3^3]}(n/3) + 2h_{[0;2,4^2]}(n/4) +$$

$$2h_{[0;2,3,6]}(n/6) + \sum_{\substack{\ell|n \\ \ell m=n}} \sum_{k|\ell} \mu(\ell/k)k^2 h_1(n/\ell)\big). \tag{5.3}$$

It remains to calculate the numbers of rooted hypermaps on orbifolds $O[0;2^4]$, $O[0;3^3]$, $0[2;4^2]$ and $O[0;2,3,6]$.

By Proposition 3.4 we obtain

$$h_{[0;2^4]}(m) = \binom{m}{4} h_0(m),\ h_{[0;3^3]}(m) = \binom{m+2}{3} h_0(m),$$

$$h_{[0;2,3,6]}(m) = \binom{m+2}{1,1,1} h_0(m) = 6\binom{m+2}{3} h_0(m),$$

$$h_{[0;2,4^2]}(m) = \binom{m+2}{1,2} h_0(m) = 3\binom{m+2}{3} h_0(m).$$

Inserting the above numbers into (5.3) we get the theorem.

The following list containing the numbers of rooted and oriented unrooted maps of genus 1 up to 30 edges follows.

Table 2. Numbers of rooted and unrooted hypermaps on the torus with at most 30 darts

No. of darts, rooted hypermaps, unrooted hypermaps

03, 1, 1
04, 15, 6
05, 165, 33
06, 1611, 285
07, 14805, 2115
08, 131307, 16533
09, 1138261, 126501
10, 9713835, 972441
11, 81968469, 7451679
12, 685888171, 57167260
13, 5702382933, 438644841
14, 47168678571, 3369276867
15, 388580070741, 25905339483
16, 3190523226795, 199408447446
17, 26124382262613, 1536728368389
18, 213415462218411, 11856420991413
19, 1740019150443861, 91579955286519
20, 14162920013474475, 708146055343668
21, 115112250539595093, 5481535740059577
22, 934419385591442091, 42473608898628639
23, 7576722323539318101, 329422709719100787
24, 61375749135369153195, 2557322884534185500
25, 496747833856061953365, 19869913354242478293
26, 4017349254284543961771, 154513432889706455145
27, 32467023775647069984085, 1202482362061007078175
28, 262225359776626483309227, 9365191420865873023026
29, 2116714406654571321840981, 72990151953605907649689
30, 17077642118698511054318251, 569254737292213025378571

The above tables were computed using MATHEMATICA, Ver. 5. The input numbers of rooted maps come from [**2**].

Acknowledgements

We would like to thank V. Liskovets and T.R.S. Walsh for helpful discussions during preparation of this paper. The research of the first author was partially supported by the Russian Foundation for Basic Research (grant 03-01-00104), INTAS (grant 03-51-3663), and Fondecyt (grants 7050189 and 1060378). The research of the second author was partially supported by the Ministry of Education of the Slovak Republic, grant APVT-51-012502.

References

[1] T.M. Apostol, "Introduction to analytical number theory", Springer,Berlin- New York, 1976.
[2] D. Arquès, Hypercartes poitées sur le tore: Décompositions et dénombrrements, *J. Combinatorial Theory B* **43**:3 (1987),275–286.
[3] M. Bousquet-Mélou and G. Schaeffer, Enumeretaion of planar constellations, *Adv. Appl. Math.* **24** (2000), 297– 329.
[4] R. Cori and A. Machì, Maps, hypermaps and their automorphisms: a survey I, II, III', *Expositiones Math.* **10** (1992), 403–427, 429–447, 449–467.
[5] J.L. Gross, T. W. Tucker, Topological graph theory, Dover Publications, New York, 2001.
[6] F. Harary, G. Prins, W.T. Tutte, The number of plane trees, *Indag. Math.* **26** (1964), 319-329.
[7] F. Harary, W.T. Tutte, The number of plane trees with given partition, *Mathematika* **11** (1964), 99-101.
[8] W.J. Harvey, Cyclic groups of automorphisms of a compact Riemann surface, *Quart. J. Math. Oxford* **17** (1966), 86-97.
[9] D. Garbe, Über die regulären Zerlegungen orientierbarer Flächen, *J. Reine Angew. Math.* **237** (1969), 39–55.
[10] D.M. Jackson, T.I. Visentin, A Combinatorial Relationship between Eulerian Maps and Hypermaps in Orientable Surfaces *J. Combinatorial Theory A* **87** (1999), 120– 150.
[11] G.A. Jones, Enumeration of homomorphisms and surface-coverings, *Quart. J. Math. Oxford* **46**:2 (1995), 485-507.
[12] G.A. Jones and D. Singerman, Complex functions: An algebraic and geometric viewpoint, Cambridge Univ. Press, Cambridge, 1987.
[13] G.A. Jones and D. Singerman, Theory of maps on orientable surfaces, *Proc. London Math. Soc.* **37** (1978), 273-307.
[14] C. Jordan, Traité des substitutions et des équationes algébriques, Réimpression du orig. 1870. Paris: Éditions J. Gabay, xvi, 669 p. (1989).
[15] J.H. Kwak, J. Lee, Enumeration of connected graph coverings, *J. Graph Theory* **23** (1996), 105–109.
[16] V. A. Liskovets, On the enumeration of subgroups of a free group. *Dokl. Akad. Nauk BSSR* **15** (1971), 6–9.
[17] V.A. Liskovets, Enumeration of nonisomorphic planar maps, *Selecta Math. Sovietica*, **4** (1985), 303-323.
[18] V.A. Liskovets, A reductive technique for enumerating non-isomorphic planar maps. *Discrete Math.* **156** (1996), 197–217.
[19] V.A. Liskovets, T.R.S. Walsh, Enumeration of Eulerian and unicursal planar maps *Discrete Math.* **282** (2004), 209–221.
[20] V.A. Liskovets, Enumerative formulae for unrooted planar maps: A pattern, *Electronic J. Comb.* **11** (2004), 14 pages.
[21] V.A. Liskovets, The orbicyclic arithmetic function, talk at workshop Graph embeddings and maps on surfaces, Stara Lesna 2005.
[22] Y.P. Liu, Some enumerating problems of maps with vertex partition, *Kexue Tongbao Sci. Bul.*, English Ed., **31** (1986), 1009-1014.
[23] Y.P. Liu, "Enumerative theory of maps", Kluwer Academic Publishers, London 1999.
[24] A. D. Mednykh, Counting non-eqivalent coverings and non-isomorphic maps for Riemann surfaces, *Proceedings of the 17th Annual International Conference on Formal Power Series and Algebraic Combinatorics, FPSAC'05, June 20-25, 2005–Taormina, Italy* (to appear)
[25] A. D. Mednykh, R. Nedela, Enumeration of unrooted maps with given genus, Submitted.
[26] C.A. Nicol, H.S. Vandiver, A von Sterneck arithmetical function and restricted partitions with respect to modulus, *Proc. Nat. Acad. Sci. U. S. A.* **40** (1954), 825-835.
[27] R.P. Stanley, "Enumerative combinatorics", Vol. 2, Cambidge Univ. Press, Cambridge 2004.
[28] W.T. Tutte, The number of planted plane trees with a given partition, *Amer. Math. Monthly* **71** (1964), 64-74.
[29] W.T. Tutte, On the enumeration of planar maps, *Bull. Amer. Math. Soc.* **74** (1968), 64-74.
[30] A. Vince, Combinatorial maps, *J. Combin. Theory* B **34** (1983), 1–21.
[31] A. Vince, Regular combinatorial maps, *J. Combin. Theory* B **35** (1983), 256–277.
[32] T.R.S. Walsh, Hypermaps versus bipartite maps , *J. Combinatorial Theory B* **18**: 2(1975), 155–163.
[33] N.C. Wormald, Counting unrooted planar maps, *Discrete Math.* **36** (1981), 205-225.
[34] N.C. Wormald, On the number of planar maps, *Canad. J. Math.* **33** (1) (1981), 1-11.

Sobolev Institute of Mathematics, Novosibirsk State University, 630090 Novosibirsk, Russia
E-mail address: `mednykh@math.nsc.ru`

Institute of Mathematics, Slovak Academy of Sciences, Severná 5, 975 49 Banská Bystrica, Slovakia
E-mail address: `nedela@savbb.sk`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

General Augmented Rook Boards & Product Formulas

Brian K. Miceli

Abstract. There are a number of so-called factorization theorems for rook polynomials that have appeared in the literature. For example, Goldman, Joichi, and White [**6**] showed that for any Ferrers board $B = F(b_1, b_2, \dots, b_n)$,

$$\prod_{i=1}^{n}(x + b_i - (i-1)) = \sum_{k=0}^{n} r_k(B)(x)\downarrow_{(n-k)}$$

where $r_k(B)$ is the k-th rook number of B and $(x)\downarrow_k = x(x-1)\cdots(x-(k-1))$ is the usual falling factorial polynomial. Similar formulas where $r_k(B)$ is replaced by some appropriate generalization of rook numbers and $(x)\downarrow_k$ is replaced by polynomials like $(x)\uparrow_{k,j} = x(x+j)\cdots(x+j(k-1))$ or $(x)\downarrow_{k,j} = x(x-j)\cdots(x-j(k-1))$ can be found in the work of Goldman and Haglund [**5**], Remmel and Wachs [**11**], Haglund and Remmel [**7**], and Briggs and Remmel [**3**]. We shall call such formulas generalized product formulas. The main goal of this paper is to develop a new rook theory setting where we can give a uniform combinatorial proof of a generalized product formula which includes all the cases referred to above. That is, given any two sequences of non-negative integers, $\mathcal{B} = (b_1, \dots, b_n)$ and $\mathcal{A} = (a_1, \dots, a_n)$, and two sign functions $sgn, \overline{sgn} : \{1, \dots, n\} \to \{-1, 1\}$, we shall define a rook theory setting and appropriate generalization of rook numbers $r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn})$ such that

$$\prod_{i=1}^{n}(x + sgn(i)b_i) = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}) \prod_{j=1}^{n-k}(x + (\sum_{s=1}^{j} \overline{sgn}(s)a_s)).$$

Thus, for example, we obtain a combinatorial interpretations of the connection coefficients between any two bases of the polynomial ring $Q[x]$ of the form $\{(x)\downarrow_{k,j}\}_{k\geq 0}$ or $\{(x)\uparrow_{k,j}\}_{k\geq 0}$. We also find q-analogues and (p,q)-analogues of the above formulas.

Résumé.

Le but principal de cet article est de développer une nouvelle théorie rook dans laquelle nous pouvons fournir des preuves combinatoires uniformes d'une formule de produit généralisée qui inclut toutes les cas cités ci-dessus. C'est-à-dire, se donnant deux suites quelconques de nombres entiers positifs, $\mathcal{B} = (b_1, \dots, b_n)$ et $\mathcal{A} = (a_1, \dots, a_n)$, et deux fonctions de signes $sgn, \overline{sgn} : \{1, \dots, n\} \to \{-1, 1\}$, nous définissons une théorie rook ainsi qu'une généralisation appropriée des nombres rook $r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn})$ tel que

$$\prod_{i=1}^{n}(x + sgn(i)b_i) = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}) \prod_{j=1}^{n-k}(x + (\sum_{s=1}^{j} \overline{sgn}(s)a_s)).$$

Donc, par exemple, nous obtenons une interprétation combinatoire des coefficients de connexion entre deux bases de l'anneau des polynômes $Q[x]$ de la forme $\{(x)\downarrow_{k,j}\}_{k\geq 0}$ ou $\{(x)\uparrow_{k,j}\}_{k\geq 0}$. Nous trouvons aussi des q-analogues et des (p,q)-analogues de ces formules.

1. Introduction

Let $\mathbb{N} = \{1, 2, 3, \dots\}$ denote the set of natural numbers. For any positive integer a, we will set $[a] := \{1, 2, \dots, a\}$. We will say that $\mathcal{B}_n = [n] \times [n]$ is an n by n array of squares (like a chess board), which we

2000 *Mathematics Subject Classification.* Primary 05A10; Secondary 05A15.

Key words and phrases. algebraic combinatorics, rook theory, enumeration, inverses.

This paper adapted from a section of the doctoral thesis of the author with thanks to the direction and assistance provided by the thesis advisor, Jeffrey Remmel.

call *cells*. The cells of $\mathcal{B}_n$ will be numbered from left to right and bottom to top with the numbers from $[n]$, and we will refer to the cell in the i^{th} row and j^{th} column of $\mathcal{B}_n$ as the (i,j) cell of $\mathcal{B}_n$. Any subset of $\mathcal{B}_n$ is called a *rook board*. If B is a board in $\mathcal{B}_n$ with column heights $b_1, b_2, \ldots, b_n$ reading from left to right, with $0 \leq b_i \leq n$ for each i, then we will write $B = F(b_1, b_2, \ldots, b_n) \subseteq \mathcal{B}_n$. In the special case that $0 \leq b_1 \leq b_2 \leq \cdots \leq b_n \leq n$, we will say that $B = F(b_1, b_2, \ldots, b_n)$ is a *Ferrers board*.

Given a board $B = F(b_1, b_2, \ldots, b_n)$, there are three sets of numbers we can associate with B, namely, the rook, file, and hit numbers of B. The rook number, $r_k(B)$, is the number of placements of k rooks in the board B so that no two rooks lie in the same row or column. The file number, $f_k(B)$, is the number of placements of k rooks in the board B so that no two rooks lie in the same column but where we allow any given row to contain more than one rook. Given a permutation $\sigma = \sigma_1\sigma_2\ldots\sigma_n$ in the symmetric group S_n, we shall identify σ with the placement $\mathbb{P}_\sigma = \{(1,\sigma_1), (2,\sigma_2), \ldots, (n,\sigma_n)\}$. Then the hit number, $h_k(B)$, is the number of $\sigma \in S_n$ such that the placement $\mathbb{P}_\sigma$ intersects the board in exactly k cells.

All of these numbers have been studied extensively by combinatorialists. Here are three fundamental identities involving these numbers. Define $(x)\downarrow_m = x(x-1)\cdots(x-(m-1))$ and $(x)\uparrow_m = x(x+1)\cdots(x+(m-1))$. Then

$$\sum_{k=0}^{n} h_k(B)x^k = \sum_{k=0}^{n} r_k(B)(n-k)!(x-1)^k, \tag{1.1}$$

$$\prod_{i=1}^{n}(x + b_i - (i-1)) = \sum_{k=0}^{n} r_{n-k}(B)(x)\downarrow_k, \text{ and} \tag{1.2}$$

$$\prod_{i=1}^{n}(x + b_i) = \sum_{k=0}^{n} f_{n-k}(B)x^k. \tag{1.3}$$

Identity (1.1) is due to Kaplansky and Riordan [**8**] and holds for any board $B \subseteq \mathcal{B}_n$. Identity (1.2) holds for all Ferrers boards $B = F(b_1, \ldots, b_n)$ and is due to Goldman, Joichi, and White [**6**]. Identity (1.3) is due to Garsia and Remmel [**4**] and holds for all boards of the form $B = F(b_1, \ldots, b_n)$. Formulas (1.2) and (1.3) are examples of what we shall call *product formulas* in rook theory.

We note that in the special case where $B = \mathbf{B}_n := F(0, 1, 2, \ldots, n-1)$, Equations (1.2) and (1.3) become

$$x^n = \sum_{k=0}^{n} r_{n-k}(\mathbf{B}_n)(x)\downarrow_k \text{ and} \tag{1.4}$$

$$(x)\uparrow_n = \sum_{k=0}^{n} f_{n-k}(\mathbf{B}_n)x^k. \tag{1.5}$$

This shows that $r_{n-k}(\mathbf{B}_n) = S_{n,k}$, where $S_{n,k}$ is the Stirling number of the second kind, and $(-1)^{n-k}f_{n-k}(\mathbf{B}_n) = s_{n,k}$, where $s_{n,k}$ is the Stirling number of the first kind, and thus, we obtain rook theory interpretations for the Stirling numbers of the first and second kind.

There are natural q-analogues of formulas (1.1), (1.2), and (1.3). That is, define $[n]_q = 1+q+\cdots+q^{n-1} = \dfrac{1-q^n}{1-q}$. We then define q-analogues of the factorials and falling factorials by $[n]_q! = [n]_q[n-1]_q\cdots[2]_q[1]_q$ and $[x]_q\downarrow_m = [x]_q[x-1]_q\cdots[x-(m-1)]_q$, Garsia and Remmel [**4**] defined q-analogues of the hit numbers, $h_k(B,q)$, q-analogues of the rook numbers, $r_k(B,q)$, and q-analogues of file numbers, $f_k(B,q)$, for Ferrers boards B so that the following hold:

$$\sum_{k=0}^{n} h_k(B,q)x^{n-k} = \sum_{k=0}^{n} r_{n-k}(B,q)[k]_q!x^k(1-xq^{k+1})\cdots(1-xq^n), \tag{1.6}$$

$$\prod_{i=1}^{n}[x + b_i - (i-1)]_q = \sum_{k=0}^{n} r_{n-k}(B,q)[x]_q\downarrow_k, \quad \text{and} \tag{1.7}$$

$$\prod_{i=1}^{n}[x + b_i]_q = \sum_{k=0}^{n} f_{n-k}(B,q)([x]_q)^k. \tag{1.8}$$

Finally, we should mention that there are also (p, q)-analogues of such formulas (see Wachs and White [**12**], Briggs and Remmel [**2**], and Briggs [**1**]).

In recent years, a number of researchers have developed new rook theory models which give rise to new classes of product formulas. For example, Haglund and Remmel [**7**] developed a rook theory model where the analogue of the the rook number $m_k(B)$ counts partial matchings in the complete graph $\mathcal{K}_n$. They defined an analogue of a Ferrers board $\tilde{F}(a_1, \ldots a_{2n-1})$ where $2n-1 \geq a_1 \geq \cdots \geq a_{2n-1} \geq 0$ and where the nonzero entries in $(a_1, \ldots, a_{2n-1})$ are strictly decreasing, and, in their setting, they proved the following identity,

$$\prod_{i=1}^{2n-1} (x + a_{2n-i} - 2i + 2) = \sum_{k=0}^{2n-1} m_k(F) x(x-2)(x-4)\cdots(x - 2(n-(k-1))). \tag{1.9}$$

Remmel and Wachs [**11**] defined a more restricted class of rook numbers, $\tilde{r}_k^j(B)$, in their *j-attacking* rook model and proved that for Ferrers boards $B = F(b_1, \ldots, b_n)$, where $b_{i+1} - b_i \geq j - 1$ if $b_i \neq 0$,

$$\prod_{i=1}^{n} (x + b_i - j(i-1)) = \sum_{k=0}^{n} \tilde{r}_{n-k}^j(B) x(x-j)(x-2j)\cdots(x-(k-1)j). \tag{1.10}$$

Goldman and Haglund [**5**] developed an *i-creation rook theory model* and proved that for Ferrers boards one has the following identity,

$$\prod_{j=1}^{n} (x + b_i + j(i-1)) = \sum_{k=0}^{n} r_{n-k}^{(i)}(B) x(x+(i-1))\cdots(x+(k-1)(i-1)). \tag{1.11}$$

In all of these new models, the authors proved q-analogues and or (p, q)-analogues of their product formulas.

2. A General Product Formula

Suppose we are given any two sequences of natural numbers: $\mathcal{B} = \{b_i\}_{i=1}^n, \mathcal{A} = \{a_i\}_{i=1}^n \in \mathbb{N}^n$. Define $A_i = a_1 + a_2 + \cdots + a_i$, the ith partial sum of the a_i's, and let $B = F(b_1, b_2, \ldots, b_n)$ be a rook board. We will also define two functions, sgn and $\overline{sgn}$, such that $sgn, \overline{sgn} : [n] \to \{-1, +1\}$. Our goal is to define a rook theory model with an appropriate notion of the rook numbers $r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn})$ such that the following product formula holds:

$$\prod_{i=1}^{n} (x + sgn(i)(b_i)) = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}) \prod_{j=1}^{n-k} (x + \sum_{s \leq j} \overline{sgn}(s)(a_s)). \tag{2.1}$$

We will refer to Equation (2.1) as the *general product formula* and the number $r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn})$ as the k^{th} *augmented rook number of* $\mathcal{B}$ *with respect to* $\mathcal{A}$, sgn, *and* $\overline{sgn}$.

2.1. Special Cases of the General Product Formula.

We first wish to consider the case where $sgn(i) = +1$ and $\overline{sgn}(i) = -1$ for every $1 \leq i \leq n$. In this case we will set

$$r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}) = r_k^{\mathcal{A}}(\mathcal{B}).$$

Thus, we want to prove Equation (2.2):

$$\prod_{i=1}^{n} (x + b_i) = \sum_{k=o}^{n} r_k^{\mathcal{A}}(\mathcal{B})(x - A_1)(x - A_2)\cdots(x - A_{n-k}). \tag{2.2}$$

To do this, we first construct an *augmented rook board*, $B^{\mathcal{A}} = F(b_1 + A_1, b_2 + A_2, \ldots, b_n + A_n)$. In $B^{\mathcal{A}}$, the cells in the i-th column are $(1, i), \ldots, (b_i + a_1 + \cdots + a_i, i)$ reading from bottom to top. We shall refer to the cells $(1, i), \ldots, (b_i, i)$ as the b_i part of column i, the cells $(b_i + 1, i), \ldots, (b_i + A_i, i)$ as the A_i part of column i, and, for each $s \leq i$, the cells $(b_i + a_1 + \cdots a_{s-1} + 1, i), \ldots, (b_i + a_1 + \cdots + a_s, i)$ as the a_s part of column i where by convention $a_{-1} = 0$. We call the part of the board $B^{\mathcal{A}}$ which corresponds to the A_i's the *augmented part of* $B^{\mathcal{A}}$. We now consider rook placements in $B^{\mathcal{A}}$ with at most one rook in each column. We define the following cancellation rule: a rook r placed in column j of $B^{\mathcal{A}}$ will cancel, in each column to its right, all of the cells which lie in the a_i part of that column where i is the highest subscript j such that the a_j part of that column has not been canceled by a rook to the left of r. For example, in Figure 1,

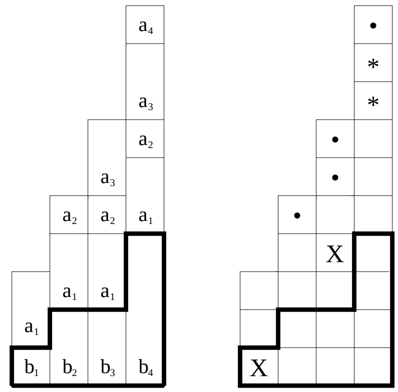


Figure 1. $B^{\mathcal{A}}$, with $B = F(1,2,2,4)$ and $\mathcal{A} = (2,1,2,1)$, and a placement of two rook in $B^{\mathcal{A}}$.

where $B = F(1,2,2,4)$ and $\mathcal{A} = (2,1,2,1)$, the rook in the first column cancels the cells in the a_2 part of the second column, the a_3 part of the third column, and the a_4 part of the fourth column (those cells which contain a "•"). The rook in the third column cancels the cells in the a_3 part the fourth column (those cells which contain a "$*$"). We then define $r_k^{\mathcal{A}}(\mathcal{B})$ to be the number of ways of placing k such rooks in $B^{\mathcal{A}}$ so that no rook lies in a cell which is canceled by a rook to its left.

We can now construct a *general augmented rook board*, $B_x^{\mathcal{A}}$, defined by the sequences $\mathcal{B} = \{b_i\}_{i=1}^n$ and $\mathcal{A} = \{a_i\}_{i=1}^n$ and some nonnegative integer x. The board $B_x^{\mathcal{A}}$ will be the board $B^{\mathcal{A}}$ (the augmented part of $B^{\mathcal{A}}$ will here be referred to as the *upper augmented part of* $B_x^{\mathcal{A}}$), with x rows appended below, called the *x-part* and then a "mirror image" of the augmented part of $B^{\mathcal{A}}$ below that, called the *lower augmented part of* $B_x^{\mathcal{A}}$. In the lower augmented part, we number the cells in i-th column with $(1,i),\ldots,(b_i + A_i, i)$ reading from top to bottom and we define the a_s part of the i-th column of the lower augmented board to consist of the cells $(a_1 + \cdots + a_{s-1} + 1, i), \ldots, (a_1 + \cdots + a_s, i)$. We say that the board $B^{\mathcal{A}}$ is separated from the x-part by the *high bar* and the x-part is separated from the lower augmented part by the *low bar.* An illustration of this type of board with $B = F(1,2,2,4)$, $\mathcal{A} = (2,1,2,1)$, and $x = 4$ can be seen in the left side of Figure 2.

In order to define a proper rook placement in the board $B_x^{\mathcal{A}}$, we make the rule that exactly one rook must be placed in every column of $B_x^{\mathcal{A}}$. When placing rooks in $B_x^{\mathcal{A}}$, we will define the following cancellation rules:

(1) A rook placed above the high bar in the j^{th} column of $B_x^{\mathcal{A}}$ will cancel all of the cells in columns $j+1, j+2, \ldots, n$, both in the upper and lower augmented parts, which belong to the a_i part of the column where i is largest j such that cells in the a_j part of the column are not canceled by a rook to their left.
(2) Rooks placed below the high bar do not cancel any cells.

An example of a rook placement in these boards can be seen in the right side of Figure 2. In this placement, the rook placed in the first column is placed above the high bar, thus it cancels in the columns to its right those cells contained in the a_i part of highest subscript in both the upper and lower augmented parts (denoted by a "•"). The rook placed in the second column is placed below the the high bar so that it cancels nothing. The rook placed in the third column is again placed above the high bar so that it cancels as does the rook placed in the first column (denoted by a "$*$"), and the last rook may be placed in any available cell.

We will now prove two lemmas in order to prove Equation (2.2).

Lemma 2.1. *If there are $b_j + A_m$ cells to place a rook above the high bar in column j, then there are A_m cells below the low bar to place a rook in column j.*

Proof: By how we define our cancellation, a block of cells from a_i gets canceled above the high bar if and only if a block of cells from a_i gets canceled below the low bar.

Lemma 2.2. *If k rooks are placed above the high bar in $B_x^{\mathcal{A}}$, then the column heights of the uncanceled cells in the lower augmented part of $B_x^{\mathcal{A}}$, when read from left to right, are $A_1, A_2, \ldots, A_{n-k}$.*

Proof: Suppose the first rook above the high bar is placed in the j^{th} column. The columns below the low bar which lie to the left of column j have heights $A_1, A_2, \ldots, A_{j-1}$. Now, the rook that was placed in the

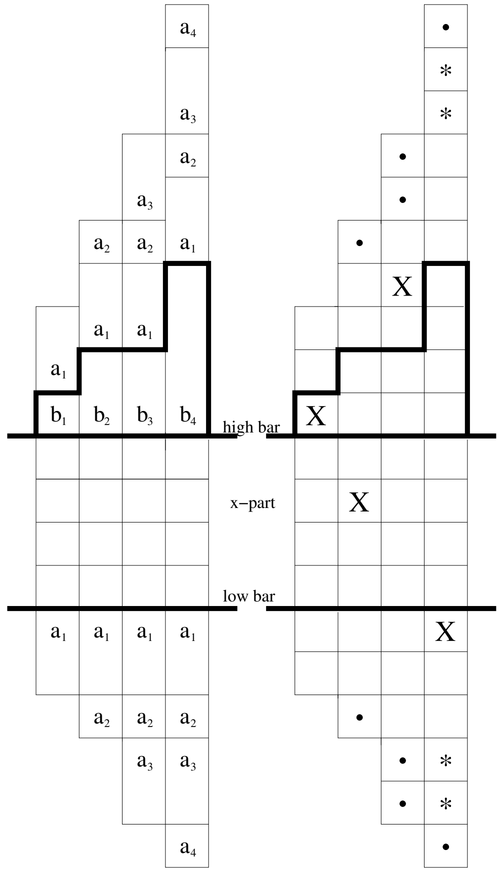


FIGURE 2. $B_x^{\mathcal{A}}$, with $B = F(1,2,2,4), \mathcal{A} = (2,1,2,1)$, and $x = 4$, and a placement of rooks in $B_x^{\mathcal{A}}$.

the j^{th} column will cancel all the cells in the a_{j+1} part of the $(j+1)^{st}$ column, all the cells in the a_{j+2} part of the $(j+2)^{nd}$ column, etc.. Thus after this cancellation, the heights of the columns below the low bar into which a rook may be placed are $A_1, A_2, \ldots, A_{j-1}, A_j, \ldots A_{n-1}$. Now suppose that the leftmost rook to the right to column j is in column k. Then the rook in column k will cancel all the cells in a_k part of the $(k+1)^{st}$ column, all the cells in the a_{k+1} part of the $(k+2)^{nd}$ column, etc.. Thus after this second cancellation, the heights of the columns below the low bar into which a rook may be placed are $A_1, A_2, \ldots, A_{j-1}, A_j, \ldots A_{k-1}, A_k, \ldots, A_{n-2}$. We can continue this type of reasoning to show that if there are k rooks are placed above the high bar in $B_x^{\mathcal{A}}$, then the column heights of the uncanceled cells in the lower augmented part of $B_x^{\mathcal{A}}$, when read from left to right, are $A_1, A_2, \ldots, A_{n-k}$.

We are now in position to prove (2.2). We shall show that (2.2) is the result of computing the sum S of the weights of all placements of n rooks in $B_x^{\mathcal{A}}$ in two different ways, where we define the weight of the rooks placed above the low bar to be "+1", the weight of the rooks placed below the low bar to be "−1", and the weight of any placement to be the product of the weights of the rooks in the placement.

If we first place the rooks starting with the leftmost column and working to the right, then we can see that in the first column there are exactly $x + b_1 + 2a_1$ cells in which to place the first rook, where the "$2a_1$" corresponds to placing the rook in either the upper or lower augmented part of the 1^{st} column. Since all of the rooks above the high bar have weight "+1" and all the rooks placed below the low bar have weight "−1", it is easy to see that the possible placements of rooks in the first column contributes a factor of $x + b_1 + a_1 + (-a_1) = x + b_1$ to S. When we consider the possible placements of a rook in the second column, we have two cases.

<u>Case I:</u> Suppose the rook that the 1^{st} column was placed below the high bar. Then nothing was canceled in the second column so we can place a rook in any cell of the second column. Thus there are a total

$x+b_2+2(a_1+a_2)$ ways to place the rook in the second column in this case. However, given our weighting of the rooks, we see that the possible placements of rooks in the second column contributes a factor of $x+b_2+(a_1+a_2)+(-a_1-a_2)=x+b_2$ to S.

Case II: If the rook in the first column was placed above the high bar, then the cells corresponding to a_2 part in both the upper and lower augmented parts of the 2^{nd} column are canceled. Thus in this case, there are $x+b_2+2a_1$ cells left to place the rook in the second column. However, given our weighting of rooks, the possible placements of rooks in the second column contributes a factor of $x+b_2+(a_1)+-a_1=x+b_2$ to S in this case.

In general, suppose we are placing a rook in the j^{th} column that does not have a rook above the high bar reading from left to right. Assume that we have placed s rooks above the high bar and t rooks below the high bar in the first $j-1$ columns. Then by Lemma 2.1, we have, $x+b_j+2(A_{j-s})$ choices as to where to place the rook in that column. Again, due to our weighting, it is easy to see that possible placement of rooks contributes a factor of $x+b_j+A_{j-s}+(-A_{j-s})=x+b_j$ to S. It follows that $S=\prod_{i=1}^{n}(x+b_i)$.

The second way of counting this sum of the weights of all the rook placements in $B_x^{\mathcal{A}}$ is to organize the placements by how many rooks lie above the high bar. Suppose that we place k rooks above the high bar and then wish to extend that to a placement in the entire board. The number of ways of placing the k rooks above the high bar is given by $r_k^{\mathcal{A}}(\mathcal{B})$. For any such placement of k rooks above the high bar, we are left with $n-k$ columns in which to place rooks below the high bar. We consider the placement of the remaining rooks in these available columns starting with the leftmost one and working right. By Lemma 2.2, the number of ways we can do this will be $(x+A_1)(x+A_2)\cdots(x+A_{n-k}))$. However, these placements come with a weighting of $(x+(-A_1)(x+(-A_2)\cdots(x+(-A_{n-k}))$ since the cells below the low bar have weight $''-1''$. Thus the sum of the weights of the set of placements in $B_x^{\mathcal{A}}$ with k rooks above the high bar is $r_k^{\mathcal{A}}(\mathcal{B})(x-A_1)(x-A_2)\cdots(x-A_{n-k}))$. Summing over all possible k gives us the RHS of (2.2).

□

Now suppose we change the weights which are assigned to rooks in $B_x^{\mathcal{A}}$ by declaring that the weight of a rook placed in the upper augmented part is "-1" and all other rooks have weight "$+1$". Again the weight of the placement is the product of the weights of the rooks in the placement. This weighting corresponds to the case where $sgn(i)=+1$ and $\overline{sgn}(i)=+1$ for every $1\le i\le n$. We will define $\tilde{r}_k^{\mathcal{A}}(\mathcal{B})$ to be the weighting of all placements of k rooks in $B^{\mathcal{A}}$ with this newly assigned weight, and this yields an equation which is analogous to Equation (2.2), namely,

$$\prod_{i=1}^{n}(x+b_i)=\sum_{k=o}^{n}\tilde{r}_k^{\mathcal{A}}(\mathcal{B})(x+A_1)(x+A_2)\cdots(x+A_{n-k}). \tag{2.3}$$

Proof of Equation (2.3): This proof follows exactly the proof of Equation (2.2) with the weights from the upper and lower augmented parts switched.

□

We can see that these two special cases encapsulate all of the product formulas stated in the Introduction. Next we sketch a proof for the general product formula (2.1).

2.2. The General Product Formula. We have now shown how to generate our general product formula in the special cases where the functions sgn and $\overline{sgn}$ are certain constant functions; however, the proofs of Equations (2.2) and (2.3) do not depend on sgn and $\overline{sgn}$ being constant. Rather, the proofs depend only on the condition that, for each column j, if the cells corresponding to the a_i part of the upper augmented part in column j are weighted with $\omega(a_i)$, then the cells corresponding to the a_i part in the lower augmented part in column j must be weighted with -$\omega(a_i)$. Moreover, the proofs do not depend on the weighting of rooks placed in the cells of the cells in the b_i part of column i in $B^{\mathcal{A}}$. Thus, if we define $r_k^{\mathcal{A}}(\mathcal{B},sgn,\overline{sgn})$ to be the weight of all placements of k rooks in the board $B^{\mathcal{A}}$, with each rook in the b_i part of column i having weight $sgn(i)$ and each rook in the a_i part of any column below the low bar having weight $\overline{sgn}(i)$, and the weight of any placement to be the product of the rooks in that placement, then we can show that Equation (2.1) is the result of computing the sum S of the weights of all placements of n rooks in $B_x^{\mathcal{A}}$ exactly as in the proofs of Equations (2.2) and (2.3).

3. Q-Analogues of General Product Formulas

In this section, we shall describe how one can derive q-analogues of some of the general product formulas described in Section 2. We do this by q-counting rook placements considered in Section 2. To simplify our notation, we shall use the convention that for any negative integer x, $[x]_q := -[|x|]_q$. If we set $\overline{A}_k = \sum_{i=1}^k \overline{sgn}(i)a_i$, then we can prove the following q-analogue of Equation 2.1:

$$\prod_{i=1}^{n}([x]_q + sgn(i)[b_i]_q) = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}, q) \prod_{s=1}^{n-k}([x]_q + [\overline{A}_s]_q). \tag{3.1}$$

For each cell c in the board $B^{\mathcal{A}}$, we let $below_{B^{\mathcal{A}}}(c)$ denote the number of cells that lie directly below c in its part. That is, if c is a cell in the augmented part of $B^{\mathcal{A}}$, then $below_{B^{\mathcal{A}}}(c)$ is the number of cells below c in the augmented part of $B^{\mathcal{A}}$ and if c is not in the augmented part of $B^{\mathcal{A}}$, then $below_{B^{\mathcal{A}}}(c)$ is just the number of cells below c in B. We may then extend this definition to the board $B_x^{\mathcal{A}}$ by defining $below_{B_x^{\mathcal{A}}}(c)$ to be the number of cells below a given cell c in $B_x^{\mathcal{A}}$ in its part. To each cell c in the board $B_x^{\mathcal{A}}$ we will assign a q-weight, $\omega_q(c)$. Given a placement P in $B_x^{\mathcal{A}}$, we will define the q-weight of that placement to be $\omega_q(P) = \prod_{r\in P}\omega_q(r)$, where $\omega_q(r) = \omega_q(c)$ if the rook r is in cell c. First, we define $\omega_q(c) = q^{below_{B_x^{\mathcal{A}}}(c)}$ if c is in the x-part of the board. Next we set $\omega_q(c) = sgn(i)q^{below_{B_x^{\mathcal{A}}}(c)}$ if c is in the i^{th} column of the board B. For the lower augmented part of the board, the definition of $\omega_q(c)$ is slightly more involved. Suppose we are at the k^{th} column of the lower augmented part of $B_x^{\mathcal{A}}$, which has column height $a_1 + a_2 + \cdots + a_k$. Recall that we labeled the cells in k-th column of the lower augmented board from top to bottom with the pairs $(1,k),(2,k),\ldots,(a_1+\cdots+a_k,k)$. Then, for $i \le a_1$, we set $\omega_q((i,k)) = \overline{sgn}(1)q^{i-1}$. Now, assume by induction that we have assigned weights to the cells $(1,k),(2,k),\ldots,(a_1+\cdots+a_i,k)$ so that $\sum_{j=1}^{a_1+\cdots+a_i}\omega_q((j,k)) = [\overline{A}_i]_q$. Then we will label the cells $(a_1+\cdots+a_i+1,k),\ldots,(a_1+\cdots+a_i+a_{i+1},k)$ in the following manner:

1. Case I: $\overline{A}_i \ge 0$
   - (a) If $\overline{A}_i \le \overline{A}_{i+1}$, then we assign the q-weight of the cells $(a_1+\cdots+a_i+1,k),\ldots,(a_1+\cdots+a_i+a_{i+1},k)$ to be $q^{\overline{A}_i}, q^{\overline{A}_i+1},\ldots,q^{\overline{A}_{i+1}-1}$, respectively.
   - (b) If $0 \le \overline{A}_{i+1} \le \overline{A}_i$, then we assign the q-weight of the cells $(a_1+\cdots+a_i+1,k),\ldots,(a_1+\cdots+a_i+a_{i+1},k)$ to be $-q^{\overline{A}_i-1}, -q^{\overline{A}_i-2},\ldots,-q^{\overline{A}_{i+1}}$, respectively.
   - (c) If $\overline{A}_{i+1} < 0$, then we assign the q-weight of the cells $(a_1+\cdots+a_i+1,k),\ldots,(a_1+\cdots+a_i+a_{i+1},k)$ to be $-q^{\overline{A}_i-1}, -q^{\overline{A}_i-2},\ldots,-1,-1,-q,-q^2,\ldots,-q^{|\overline{A}_{i+1}|-1}$, respectively.
2. Case II: $\overline{A}_i < 0$
   - (a) If $\overline{A}_i \ge \overline{A}_{i+1}$, then we assign the q-weight of the cells $(a_1+\cdots+a_i+1,k),\ldots,(a_1+\cdots+a_i+a_{i+1},k)$ to be $-q^{|\overline{A}_i|}, -q^{|\overline{A}_i|+1},\ldots,-q^{|\overline{A}_{i+1}|-1}$, respectively.
   - (b) If $0 \ge \overline{A}_{i+1} \ge \overline{A}_i$, then we assign the q-weight of the cells $(a_1+\cdots+a_i+1,k),\ldots,(a_1+\cdots+a_i+a_{i+1},k)$ to be $q^{|\overline{A}_i|-1}, q^{|\overline{A}_i|-2},\ldots,-q^{|\overline{A}_{i+1}|}$, respectively .
   - (c) If $\overline{A}_{i+1} > 0$, then we assign the q-weight of the cells $(a_1+\cdots+a_i+1,k),\ldots,(a_1+\cdots+a_i+a_{i+1},k)$ to be $q^{|\overline{A}_i|-1}, q^{|\overline{A}_i|-2},\ldots,1,1,q,q^2,\ldots,q^{\overline{A}_{i+1}-1}$, respectively.

Finally, in order to assign the q-weights to the k^{th} column of the upper augmented part of $B_x^{\mathcal{A}}$, we will simply take the weights that we assigned to the lower augmented part of the k^{th} column, flip them upside down and multiply them all by "-1". An example of this weighting can be seen in Figure 3, where the q-number displayed in each cell of the diagram corresponds to the q-weight a rook placed in that cell would be given. For example, we can see that the q-weights assigned to the lower augmented part of the fourth column, read from top to bottom are: $1, q, -q, -1, -1, 1$. The weights in the upper augmented part of the same column are, when read from bottom to top: $-1, -q, q, 1, 1, -1$, which is the previous sequence with every element multiplied by "-1".

Now we can prove Equation 3.1 similar to the way we proved Equation 2.1 in the previous section. That is, Equation 3.1 results by computing the sum S_q of the q-weights of all placements of n rooks in $B_x^{\mathcal{A}}$ in two different ways.

3.1. Special Cases of the General Q-Analogue Formula.

Now consider the special cases where sgn and $\overline{sgn}$ are the constant functions -1 or $+1$. In this case, it is easy to see that

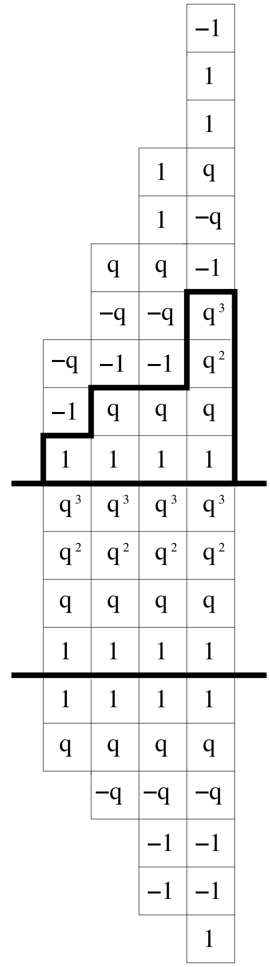


Figure 3. A q-analogue of the rook placement in Figure 2 with $sgn(i) = +1$ for $i = 1, 2, 3, 4$ and $\overline{sgn}(i) = +1$ for $i = 1, 3, 4$ $\overline{sgn}(i) = -1$ for $i = 2$. Here each cell is labeled with the q-weight that a rook placed in that cell would be given.

(i) if a rook in is the b_i part of column i of $B^{\mathcal{A}}$, then its q-weight will be $sgn(i)q^{below_{B_x^{\mathcal{A}}}(c)}$,
(ii) if a rook in is the A_i part of column i of $B^{\mathcal{A}}$, then its q-weight will be $-\overline{sgn}(i)q^{below_{B_x^{\mathcal{A}}}(c)}$,
(iii) if rook in is the x part of column i of $B^{\mathcal{A}}$, then its q-weight will be $q^{below_{B_x^{\mathcal{A}}}(c)}$, and
(iv) the q-weights of a cell in the lower augmented part of the board is just the q-weight of its mirror image in the upper augmented part of the board multiplied by $'' - 1''$.
In this case (3.1) becomes

$$\prod_{i=1}^{n}([x]_q + sgn(i)[b_i]_q) = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}, q) \prod_{s=1}^{n-k}([x]_q + \overline{sgn}(s)[A_s]_q). \tag{3.2}$$

It turns out that by slightly modifying our q-counting of rook placements, we can prove analogues of (3.2) where we replace $[x] - [c]$ by $[x - c]$ or $[x] + [c]$ by $[x + c]$.

3.1.1. *Case I: $sgn(i) = \overline{sgn}(i) = -1$.* For $x, c \in \mathbb{N}$ with $x > c$, we have that $[x]_q - [c]_q = q^c[x - c]_q$. Thus (3.2) becomes

$$\prod_{i=1}^{n} q^{b_i}[x - b_i]_q = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}, q) \prod_{s=1}^{n-k} q^{A_s}[x - A_s]. \tag{3.3}$$

It is then easy to see that if we replace $r_k^{\mathcal{A}}(\mathcal{B},q)$ with $\hat{r}_k^{\mathcal{A}}(\mathcal{B},q)$ by

$$\hat{r}_k^{\mathcal{A}}(\mathcal{B},q) := q^{(A_1+A_2+\cdots+A_{n-k})-(b_1+\cdots+b_n)} r_k^{\mathcal{A}}(\mathcal{B},q),$$

we obtain the following form of Formula 3.2:

$$\prod_{i=1}^{n}([x-b_i]_q) = \sum_{k=o}^{n} \hat{r}_k^{\mathcal{A}}(\mathcal{B},q)([x-A_1]_q)([x-A_2]_q)\cdots([x-A_{n-k}]_q). \tag{3.4}$$

3.1.2. *Case II: $sgn(i) = -1, \overline{sgn}(i) = +1$.* For $x, c \in \mathbb{N}$ we have that $[x]_q + q^x[c]_q = [x+c]_q$. Thus if we want to replace $[x]_q + [A_i]_q$ by $[x+A_i]_q = [x]_q + q^x[A_i]$, then we should weight each rook that lies in upper augmented part of $B^{\mathcal{A}}$ by an extra factor of q^x. This means that when we consider placements in $B_x^{\mathcal{A}}$, then we must also weight each rook that lies in the lower augmented part of $B_x^{\mathcal{A}}$ with an extra factor of q^x so that for any given column the weights of possible placements in the lower and upper augmented parts cancel each other as in the proofs in Section 2. Thus we define $\hat{\hat{r}}_k^{\mathcal{A}}(\mathcal{B},q)$ to be the sum of the q-weight over all placements of k rooks in $B^{\mathcal{A}}$ where each rook placed in the augmented part receiving an extra factor of q^x. Then it is easy to see that (3.2) becomes

$$\prod_{i=1}^{n}([x]_q - [b_i]_q) = \sum_{k=0}^{n} \hat{\hat{r}}_k^{\mathcal{A}}(\mathcal{B},q)[x+A_1]\cdots[x+A_{n-k}]. \tag{3.5}$$

Finally we replace $\hat{\hat{r}}_k^{\mathcal{A}}(\mathcal{B},q)$ by a new q-rook number, $\tilde{\tilde{r}}_k^{\mathcal{A}}(\mathcal{B},q)$ where $\tilde{\tilde{r}}_k^{\mathcal{A}}(\mathcal{B},q) := q^{-(b_1+\cdots+b_n)} r_k^{\mathcal{A}}(\mathcal{B},q)$. In doing this, we obtain the following formula:

$$\prod_{i=1}^{n}([x-b_i]_q) = \sum_{k=0}^{n} \tilde{\tilde{r}}_k^{\mathcal{A}}(\mathcal{B},q)([x+A_1]_q)([x+A_2]_q)\cdots([x+A_{n-k}]_q). \tag{3.6}$$

We can also use methods similar to the ones used in Cases I and II, to prove the following product formulas for appropriate choices of $\overline{r}_k^{\mathcal{A}}(\mathcal{B},q)$ and $\overline{\overline{r}}_k^{\mathcal{A}}(\mathcal{B},q)$.

3.1.3. *Case III: $sgn(i) = +1, \overline{sgn}(i) = -1$.*

$$\prod_{i=1}^{n}([x+b_i]_q) = \sum_{k=o}^{n} \overline{r}_k^{\mathcal{A}}(\mathcal{B},q)([x-A_1]_q)([x-A_2]_q)\cdots([x-A_{n-k}]_q) \tag{3.7}$$

3.1.4. *Case IV: $sgn(i) = \overline{sgn}(i) = +1$.*

$$\prod_{i=1}^{n}([x+b_i]_q) = \sum_{k=o}^{n} \overline{\overline{r}}_k^{\mathcal{A}}(\mathcal{B},q)([x+A_1]_q)([x+A_2]_q)\cdots([x+A_{n-k}]_q) \tag{3.8}$$

4. (P,Q)-Analogues of General Product Formulas

For any $n \in \mathbb{N}$ we define $[n]_{p,q} = p^{n-1} + qp^{n-2} + \cdots + q^{n-2}p + q^{n-1}$, and we again use the convention that for a negative integer x, $[x]_{p,q} := -[|x|]_{p,q}$. Then we can give a combinatorial interpretation of the following (p,q)-analogue formula:

$$\prod_{i=1}^{n}([x]_{p,q} + sgn(i)[b_i]_{p,q}) = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}, p, q) \prod_{s=1}^{n-k}([x]_{p,q} + [\overline{A}_s]_{p,q}). \tag{4.1}$$

Again, we will assign a weight to each cell c of the board $B_x^{\mathcal{A}}$, which we will call the (p,q)-weight of c, and this will be denoted by $\omega_{p,q}(c)$. We will also define the statistic $above_{B_x^{\mathcal{A}}}(c)$, for any cell c in the board $B_x^{\mathcal{A}}$, to be the number of cells that lie above c in its part, that is, if c is in the x-part of the board, then $above_{B_x^{\mathcal{A}}}(c)$ is the number of cells that lie above c in the x-part. For a placement P of rooks in $B_x^{\mathcal{A}}$, we will let the (p,q)-weight of P be $\omega_{p,q}(P) = \prod_{r \in P} \omega_{p,q}(r)$, where $\omega_{p,q}(r) = \omega_{p,q}(c)$ if the rook r is placed in cell c. Now, we can (p,q)-weight the cells of $B_x^{\mathcal{A}}$ in the following manner:

(1) If c is in the x-part of the board, then $\omega_{p,q}(c) = p^{above_{B_x^{\mathcal{A}}}(c)} q^{below_{B_x^{\mathcal{A}}}(c)}$.
(2) If c is in the i^{th} column of the board B, then $\omega_{p,q}(c) = sgn(i) p^{above_{B_x^{\mathcal{A}}}(c)} q^{below_{B_x^{\mathcal{A}}}(c)}$.

(3) If c is in the k^{th} column of the lower augmented part of the board, then we will set $\omega(1,k) = [\overline{A}_1]_{p,q}$. We will then set $\omega(a_1 + \cdots + a_i + 1, k) = [\overline{A}_{i+1}]_{p,q} - [\overline{A}_i]_{p,q}$ and $\omega(j,k) = 0$ otherwise.

(4) If c is in the k^{th} column of the upper augmented part of the board, then weights will be assigned, from bottom to top, as they were in the lower augmented part, with all of the weights multiplied by "-1".

We note that this type of weighting is more complicated than our q-weighting since now a cell can receive a (p,q)-weight which is a polynomial in p and q rather than just a plus or minus a power of q. Moreover, there are many other choices we could make for the weights, but none of them reduce to the q-weight when $p = 1$. However, in certain special cases, we can assign a more natural (p,q)-weight which is consistent with some of the (p,q)-analogues of product formulas that have appeared in the literature, but we shall not consider these types of results in this paper.

We can now prove Equation 4.1 in the exact same way that we proved Equation 3.1.

5. Conclusion and Perspectives

We have given a rook theory interpretation of the product formula

$$\prod_{i=1}^{n}(x + sgn(i)b_i) = \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn}) \prod_{j=1}^{n-k}(x + (\sum_{s=1}^{j} \overline{sgn}(s)a_s)),$$

and this interpretation can be used to obtain identities studied by Goldman and Haglund [**5**], Remmel and Wachs [**11**], Haglund and Remmel [**7**], and Briggs and Remmel [**3**]. We also have q- and (p,q)- analogues of this general product formula.

One application of this new theory is in finding the inverses of connection coefficients for different bases of $\mathbb{Q}[x]$ [**9**]. If we define the functions $(x)\uparrow_{k,a}= x(x+a)(x+2a)\cdots(x+(k-1)a)$ and $(x)\downarrow_{k,b}= x(x-b)(x-2b)\cdots(x-(k-1)b)$, then for any $a \in \mathbb{N}$, the sets $\{(x)\uparrow_{n,a}\}_{n\geq 0}$ and $\{(x)\downarrow_{n,a}\}_{n\geq 0}$ will both form bases of $\mathbb{Q}[x]$. Thus, there exist numbers $C_{n,k}(b\downarrow, a\uparrow)$ and $C_{n,k}(a\uparrow, b\downarrow)$ such that

$$(x)\downarrow_{n,b}= \sum_{k=0}^{n} C_{n,k}(b\downarrow, a\uparrow)(x)\uparrow_{k,a} \tag{5.1}$$

and

$$(x)\uparrow_{n,a}= \sum_{k=0}^{n} C_{n,k}(a\uparrow, b\downarrow)(x)\downarrow_{k,b}\ . \tag{5.2}$$

From linear algebra it is known that $||C_{n,k}(b\downarrow, a\uparrow)||^{-1} = ||C_{n,k}(a\uparrow, b\downarrow)||$, that is to say,

$$\sum_{j=k}^{n} C_{n,j}(a\uparrow, b\downarrow)C_{j,k}(b\downarrow, a\uparrow) = \chi(n=k). \tag{5.3}$$

However, this result may be obtained from our rook theory model. Given the numbers $a, b \in \mathbb{N}$ we will define $\mathcal{B} = (0, b, 2b, \ldots, (n-1)b)$ and $\mathcal{A} = (0, a, a, \ldots, a)$. By now defining $sgn(i) = -1$ and $\overline{sgn}(i) = +1$ we see that $C_{n,k}(b\downarrow, a\uparrow) = r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn})$ and $C_{n,k}(a\uparrow, b\downarrow) = r_k^{\mathcal{B}}(\mathcal{A}, \overline{sgn}, sgn)$. We can now write equations (5.1) and (5.2) as

$$(x)\downarrow_{n,b}= \sum_{k=0}^{n} r_k^{\mathcal{A}}(\mathcal{B}, sgn, \overline{sgn})(x)\uparrow_{k,a} \tag{5.4}$$

and

$$(x)\uparrow_{n,a}= \sum_{k=0}^{n} r_k^{\mathcal{B}}(\mathcal{A}, \overline{sgn}, sgn)(x)\downarrow_{k,b}\ . \tag{5.5}$$

In particular, we now have that

$$\sum_{j=k}^{n} r^{\mathcal{A}}_{n-k}(\mathcal{B}, sgn, \overline{sgn}) r^{\mathcal{B}}_{j-k}(\mathcal{A}, \overline{sgn}, sgn) = \chi(n = k), \tag{5.6}$$

and we have a completely combinatorial proof of this fact based solely on involutions on rook placements in an augmented rook board setting. We can give similar combinatorial proofs for all the possible choice of $\uparrow$ and $\downarrow$ in the coefficient $C_{n,k}(b \downarrow, a \uparrow)$. For example, we can find combinatorial interpretations of the inverses of the numbers $C_{n,k}(a \uparrow, b \uparrow)$ and $C_{n,k}(a \downarrow, b \downarrow)$ which satisfy the equations

$$(x) \uparrow_{n,a} = \sum_{k=0}^{n} C_{n,k}(a \uparrow, b \uparrow)(x) \uparrow_{k,b} \tag{5.7}$$

and

$$(x) \downarrow_{n,a} = \sum_{k=0}^{n} C_{n,k}(a \downarrow, b \downarrow)(x) \downarrow_{k,b} . \tag{5.8}$$

Another application of our rook theory model relates to the numbers $S^{p(x)}_{n,k}$ defined in [**10**] by

$$S^{p(x)}_{n+1,k} = S^{p(x)}_{n,k-1} + p(k) S^{p(x)}_{n,k}, \tag{5.9}$$

where $p(x)$ is any polynomial with nonnegative integer coefficients and with initial conditions $S^{p(x)}_{0,0} = 1$ and $S^{p(x)}_{n,k} = 0$ whenever $n < 0, k < 0$, or $n < k$. We call such numbers *poly-Stirling numbers of the second kind* [**10**]. Then, for example, in the special case where $p(x) = x^m$, we can use an extension of the theory of general augmented rook boards to give a combinatorial proof of the formula

$$(x^n)^m = \sum_{k=0}^{n} S^{x^m}_{n,k} \prod_{j-1}^{k} (x^m - (j-1)^m). \tag{5.10}$$

Finally, we should note that a theory of hit numbers corresponding to the rook theory for our generalized product formulas has yet to be developed.

References

[1] K. S. Briggs, q-Analogues and p,q-Analogues of Rook Numbers and Hit Numbers and Their Extensions, *Thesis, UC, San Diego* (2005)
[2] K. S. Briggs and J. B. Remmel, A p,q-analogue of a Formula of Frobenius, *Electronic Journal of Combinatorics* **10(1)** (2003), #R9.
[3] K. S. Briggs and J. B. Remmel, m-Rook Numbers and a Generalization of a Formula of Frobenius for $C_m \S S_n$, *Electronic Journal of Combinatorics*, submitted.
[4] A. M. Garsia and J. B. Remmel, Q-Counting Rook Configurations and a Formula of Frobenius, *J. Combin. Theory Ser. A* **41** (1986), 246-275.
[5] J. Goldman and J. Haglund Generalized Rook Polynomials, *J. Combin. Theory Ser. A* **91** (2000), 509-530.
[6] J. R. Goldman, J. T. Joichi, and D. E. White, Rook Theory I. Rook equivalence of Ferrers boards, *Proc. Amer. Math Soc.* **52** (1975), 485-492.
[7] J. Haglund and J. B. Remmel, Rook Theory for Perfect Matchings, *Advances in Applied Mathematics* **27** (2001), 438-481.
[8] I. Kaplansky and J. Riordan, The problem of rooks and its applications, *Duke Math. J.* **13** (1946), 259-268.
[9] J. Liese, B. K. Miceli, and J. B. Remmel, The Combinatorics of the Connection Coefficients of Generalized Rising and Falling Factorial Polynomials, *In Preparation*.
[10] B. K. Miceli, TBD, *Thesis, UC, San Diego* (2006)
[11] J. B. Remmel and M. Wachs, Generalized p,q-Stirling numbers, *private communication*.
[12] M. Wachs and D. White, p,q-Stirling numbers and set partition statistics, *J. Combin. Theory Ser. A* **56** (1991), 27-46.

Department of Mathematics, University of California, San Diego, La Jolla, CA, USA, 92093-0112
E-mail address: bmiceli@math.ucsd.edu
URL: http://www.math.ucsd.edu/~bmiceli

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

The Chromatic Symmetric Function of Symmetric Caterpillars and Near-Symmetric Caterpillars

Matthew Morin

Abstract. For every proper coloring κ of a graph with vertex set $\{v_1, v_2, \ldots, v_n\}$, one obtains a monomial of degree n defined by $\mathbf{x}^\kappa = x_{\kappa(v_1)} x_{\kappa(v_2)} \ldots x_{\kappa(v_n)}$. Summing these monomial terms over all proper colorings of a given graph G gives the chromatic symmetric function $X_G(\mathbf{x})$. Using Stanley's expansion of the chromatic symmetric function in the power sum basis $\{p_\lambda(\mathbf{x})\}_{\lambda \vdash n}$ of the space Λ^n of homogeneous symmetric functions of degree n, we identify properties of our graph as various coefficients of the $p_\lambda(\mathbf{x})$ in this expansion for $X_G(\mathbf{x})$.

We focus on caterpillars, that is, those trees which becomes a path when all of its vertices of degree one, are deleted. This path is known as the spine of the caterpillar. A caterpillar C is said to be symmetric if there is an isomorphism that exchanges the endpoints of the spine, and is called near-symmetric if the caterpillar becomes symmetric upon shifting a single edge of C into the spine.

We use the coefficients of $p_\lambda(\mathbf{x})$ in the expansion of $X_C(\mathbf{x})$, for λ being a partition with two parts, to show that the chromatic symmetric function distinguishes symmetric and near-symmetric caterpillars from all other caterpillars. We also show that if two trees have a different number of leaves, then they also have different chromatic symmetric functions.

Résumé. Nous sommes intéressés dans le problème de si la fonction symétrique chromatique $X_G(\mathbf{x})$ distingue les arbres nonisomorphe. En utilisant l'expansion de Stanley de la fonction symétrique chromatique dans la base $\{p_\lambda(\mathbf{x})\}_{\lambda \vdash n}$ de l'espace Λ^n des fonctions symétriques homogènes de degré n, nous identifions des propriétés de notre graphique comme divers coefficients de $p_\lambda(\mathbf{x})$ dans cette expansion pour $X_G(\mathbf{x})$.

Nous concentrons sur chenilles, c'est-à-dire, ces arbres qui devient un chemin quand tous ses sommets du degré un sont supprimés. Ce chemin est connu comme épine de la chenille. Une chenille C s'appelle symétrique s'il y a un isomorphisme qui échange les sommets finaux de l'épine, et s'appelle proche-symétrique si la chenille devient symétrique par l'insertion d'un arc de C en l'épine.

Nous employons les coefficients de $p_\lambda(\mathbf{x})$ dans l'expansion de $X_C(\mathbf{x})$, pour λ étant une cloison avec deux parts, pour prouver que la fonction symétrique chromatique distingue les chenilles symétriques et chenilles proche-symétriques de tous autres chenilles. Aussi, nous prouvons que la fonction symétrique chromatique distingue les arbres qui ont un nombre différent de sommets du degré un.

1. Introduction

In this paper we shall only consider the case of *simple* graphs, that is, those with no loops or multiple edges. Let $\mathbf{x} = x_1, x_2, \ldots$ be a countable sequence of commutative inderterminates and G be a graph with vertex set V and edge set E. Given a coloring κ of G, that is a map $\kappa : V \to \mathbb{N}$, we write $\mathbf{x}^\kappa$ for the monomial term of degree $n = |V|$ defined by

$$\mathbf{x}^\kappa = \prod_{v \in V} x_{\kappa(v)}.$$

The *chromatic symmetric function* $X_G(\mathbf{x})$ is then defined by taking

$$X_G(\mathbf{x}) = \sum_\kappa \mathbf{x}^\kappa, \tag{1}$$

2000 *Mathematics Subject Classification.* Primary 05C15; Secondary 05C05.
Key words and phrases. chromatic symmetric function, trees, caterpillars.

where the sum is over all *proper* colorings κ, i.e. those colorings for which $\kappa(u) \neq \kappa(v)$ for every edge uv of the graph G.

Any coloring of a graph partitions the vertex set into a finite number of color classes, and given a proper coloring of G, permuting these color classes yields another proper coloring of G. Thus $X_G(\mathbf{x})$ is symmetric in the inderterminates $x_1, x_2, x_3, \ldots$. For convenience, we shall drop reference to the variables and write X_G in place of $X_G(\mathbf{x})$.

The *chromatic polynomial* $\chi(G, k)$ of a graph G gives the number of proper colorings using only k colors. We note that $X_G(1^k) = \chi(G, k)$, where 1^k denotes setting $x_1 = x_2 = \ldots = x_k = 1$ and $x_{k+1} = x_{k+2} = \ldots = 0$, since then a monomial survives if, and only if, it comes from a proper coloring using the colors $\{1, 2, \ldots, k\}$, in which case the contribution to the sum is 1. It is easy to see that the chromatic polynomial of any n-vertex tree T is given by $\chi(T, k) = k(k-1)^{n-1}$.

We are interested in the following question of Stanley [**Stanley, 1995**].

Problem 1.1. *Does the chromatic symmetric function distinguish every pair of nonisomorphic trees? That is, given trees T_1 and T_2, do we have $X_{T_1} = X_{T_2}$ if, and only if, $T_1 \cong T_2$?*

The rest of the paper is structured as follows. In the next section we derive some straightforward results for graphs. In Section 3 we look at a labelling procedure for caterpillars, and discuss its relation to symmetric caterpillars. Section 4 uses this labelling procedure to solve Problem 1.1 in the case of symmetric and near-symmetric caterpillars. In Section 5 we turn to counting the number of n-vertex symmetric caterpillars. Finally, in Section 6, we conclude by collecting our results, showing the existence of certain families of graphs.

1.1. Acknowledgements. The author is extremely grateful to Stephanie van Willigenburg for directing us to Problem 1.1, for always having fresh suggestions, and for doing a thorough job of editing. In addition, many thanks must go to Richard Stanley who not only inspired this work, but also passed comments on it back our way. Particularly, the proposition in Section 6 came about from his suggestion to combine our various results. Further thanks go out to Jeremy Martin for taking an interest in this problem, and for spending some time reading my work and offering further suggestions.

2. Definitions and General Results

If $\{p_\lambda(\mathbf{x})\}_{\lambda \vdash n} = \{p_\lambda\}_{\lambda \vdash n}$ is the power sum basis of Λ^n, the space of homogeneous symmetric functions of degree n, then we have the following.

Theorem 2.1. [**Stanley, 1995**, Theorem 2.5] *For an n-vertex graph G*

$$X_G = \sum_{F \subseteq E} (-1)^{|F|} p_{\lambda(F)},$$

where $\lambda(F)$ is the partition of n whose parts correspond to the sizes of the connected components in the spanning subgraph of G with edge set F.

From its definition, it is clear that X_G is homogeneous of degree n. Hence graphs with a different number of vertices have different chromatic symmetric functions.

We shall use the notation $[p_\lambda]X_G$ to denote the coefficient of p_λ in the expansion of X_G in terms of the basis $\{p_\lambda\}_{\lambda \vdash n}$ of Λ^n. From Theorem 2.1 we have

$$[p_\lambda]X_G = \sum_{\substack{F \subseteq E \\ \lambda(F) = \lambda}} (-1)^{|F|}. \tag{2}$$

The only way to obtain the partition $\lambda(F) = (1^n)$ is for F to include no edges of G, so the only contribution to the coefficient of $p_{(1^n)}$ comes from $F = \emptyset$. Hence, for each graph G,

$$[p_{(1^n)}]X_G = 1. \tag{3}$$

The only way to obtain the partition $\lambda(F) = (2, 1^{n-2})$ is for F to include a single edge of G. Hence the only contributions to the coefficient of $p_{(2,1^{n-2})}$ is from the sets F with $|F| = 1$. Thus Equation 2 gives

$$[p_{(2,1^{n-2})}]X_G = -|E|, \tag{4}$$

for every graph with edge set E. Similarly

$$[p_{(2^k,1^{n-2k})}]X_G = (-1)^k \mu_k(G), \tag{5}$$

where $\mu_k(G)$ is the number of ways of selecting k vertex-disjoint edges in G, that is, the *number of matchings in G of size k.*

In the interest of Problem 1.1, we turn to the case where our graph G is a tree T. We restrict to the case of $n \geq 3$ vertices. Then the partition $\lambda(F) = (k, 1^{n-k})$ arises precisely when the edge set F determines a k-vertex subtree of T, requiring exactly $k-1$ edges. Thus Equation 2 gives

$$[p_{(k,1^{n-k})}]X_G = (-1)^{k-1} T_k, \tag{6}$$

where T_k is the number of k-vertex subtrees of T. More generally, if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_j)$, we can show that

$$[p_\lambda]X_T = (-1)^{n-j} T_\lambda, \tag{7}$$

where T_λ is the number of partitions of T into disjoint subtrees of size $\lambda_1, \lambda_2, \ldots, \lambda_j$.

Within the context of trees, it is common to refer to a vertex of degree one as a *leaf*. Vertices of degree larger than one are called *internal* vertices. We say that an edge of a graph is *internal* if both of its endpoints are internal. Otherwise at least one endpoint of the edge is a leaf, and we call the edge *external*.

Every edge that is removed from a tree T increases the number of connected components by one, so to obtain a partition $\lambda(F)$ with two parts requires F to be of the form $E - \{e\}$, where e is an edge of T. In the next few sections we inspect the partitions obtained by removing internal edges, and show how the coefficients of these partitions in X_T help attack Problem 1.1 in the case of caterpillars. Before moving in that direction, we inspect the simpler case of partitions obtained by removing an external edge from a tree.

Proposition 2.2. *If T is an n-vertex tree with $n \geq 3$, then*

$$[p_{(n-1,1)}]X_T = (-1)^n L(T),$$

where $L(T)$ is the number of leaves of T.

Thus the chromatic symmetric function distinguishes trees with a different number of leaves.

Proof. Every leaf is the endpoint of some external edge of T, and since there are at least three vertices in T, no edge of T has a leaf as both of its endpoints. Thus the number of leaves in T is the same as the number of external edges in T.

To obtain the partition $\lambda(F) = (n-1, 1)$ in Equation 2, the edge subset F must isolate a single vertex of T. This can be accomplished when the set $F \subseteq E$ excludes a single external edge of T, and this is the only way this partition can arise. Since there are $n-1$ edges in T, these F have $|F| = n-2$, and hence

$$[p_{(n-1,1)}]X_T = (-1)^{n-2} L(T), \tag{8}$$

where $L(T)$ is the number of leaves of T.

□

3. Caterpillars, Spine Sets, and Symmetry

A *caterpillar* C is a tree which contains a path consisting of internal vertices of C such that every vertex of C that is not on the path is adjacent to a vertex on the path. This path is called the *spine* of the caterpillar. With our definitions, the spine of a caterpillar is the unique subgraph induced by the set of internal vertices of the caterpillar. If we do *not* make the requirement that the vertices of the spine be internal vertices of C, which may prove convenient in some instances, then the spine is no longer unique.

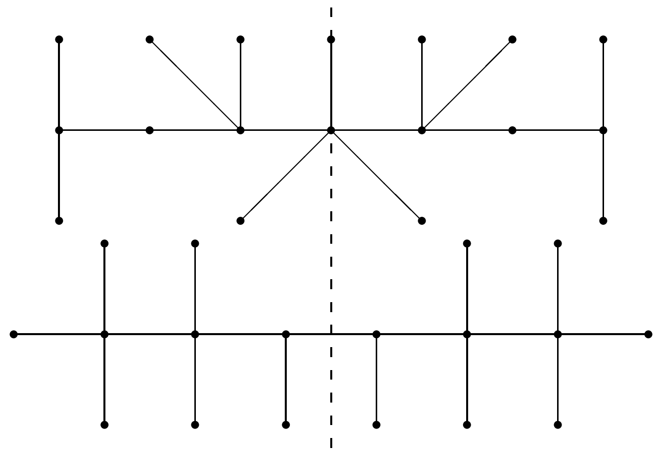
If the spine of a caterpillar consists of the path of vertices $x_1, x_2, \ldots, x_k$, then we call $\delta = (deg(x_1), deg(x_2), \ldots, deg(x_k))$ a *degree sequence* of the spine; the other degree sequence being $(deg(x_k), deg(x_{k-1}), \ldots, deg(x_1))$. We call the caterpillar *symmetric* if the degree sequence δ is palindromic, that is, when the two possible degree sequences of the spine are equal. Equivalently, a caterpillar C is symmetric if there is an automorphism of C that switches endpoints of the spine. Visually, suitably drawn, one half of the caterpillar is the mirror image of the other.

Example:

Here we see two symmetric caterpillars.

$\delta = (3, 2, 4, 5, 4, 2, 3)$

$\delta = (4, 4, 3, 3, 4, 4)$



Given an n-vertex caterpillar, we shall create a labelling of its edges with the numbers $1, 2, 3, \ldots, n-1$ as follows.

First take an endpoint of the spine and mark it. Now starting at the marked vertex, we iterate:

(1) Let u be the vertex of the spine which has just been marked.
(2) Label the unlabelled external edges incident to u with the smallest unused labels among $1, 2, 3, \ldots n-1$.
(3) If there is an unlabelled internal edge incident with u, say $e = uv$, then label e with the smallest unused label among $1, 2, 3, \ldots n-1$, mark vertex v, and proceed back to 1. If there is no unlabelled internal edge incident to u, then the labelling is complete.

Collecting the labels of the internal edges of a caterpillar C gives rise to a set $\mathscr{S}_C \subseteq \{2, 3, \ldots, n-2\}$ called a *spine set* of C. Note that two spine sets of a given caterpillar are possible, since either endpoint of the spine could have been chosen to be initially marked in the labelling procedure. If a degree sequence of the spine is $\delta = (\delta_1, \delta_2, \ldots, \delta_k)$ with δ_1 corresponding to the degree of vertex v, then the spine set of C one obtains by initially marking v is

$$\mathscr{S}_C = \{\delta_1, \delta_1 + \delta_2 - 1, \delta_1 + \delta_2 + \delta_3 - 2, \ldots, \delta_1 + \delta_2 + \ldots + \delta_{k-1} - k + 2\}.$$

Conversely, given any set $S \subseteq \{2, 3, \ldots, n-2\}$, say $S = \{x_1, x_2, \ldots, x_k\}$ where $x_1 < x_2 < \ldots < x_k$, we can associate to S the n-vertex caterpillar $\mathscr{C}_S$ that has spine set S by using the caterpillar whose spine has the degree sequence given by

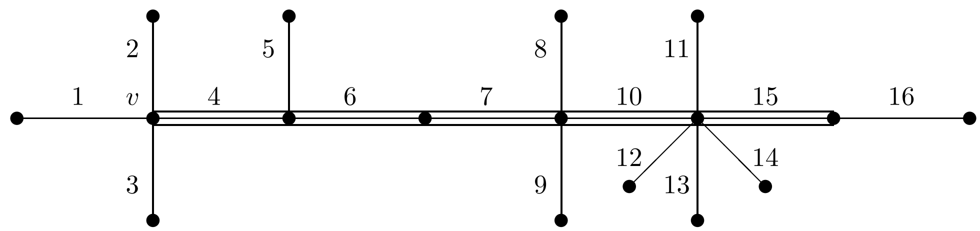
$$\delta = (x_1, x_2 - x_1 + 1, x_3 - x_2 + 1, \ldots, x_k - x_{k-1} + 1, n - x_k).$$

Then for each n-vertex caterpillar C with k internal edges, we have

$$\mathscr{C}_{\mathscr{S}_C} \cong C. \tag{9}$$

Example:

Below we see a caterpillar that has had its edges labelled as described by the above procedure, where the vertex labelled v is the one that was initially marked. The spine of the caterpillar has been highlighted.



From this we obtain the spine set $\mathscr{S}_C = \{4, 6, 7, 10, 15\}$.

Conversely, given the set $S = \{4, 6, 7, 10, 15\}$, we can construct the 17 vertex caterpillar $\mathscr{C}_S$ with $|S| = 5$ internal edges by taking a caterpillar with spine degree sequence given by

$$\begin{aligned}\delta &= (4, 6-4+1, 7-6+1, 10-7+1, 15-10+1, 17-15) \\ &= (4, 3, 2, 4, 6, 2).\end{aligned}$$

This is exactly the caterpillar we began with.

Given a set $S \subseteq \{2, 3, \ldots, n-2\}$, we call the set $S' = \{n - i | \ i \in A\}$ the *reflection* of S. If the set S satisfies $S = S'$ we shall call S a *symmetric* subset. The following result shows that the two spine sets one can obtain from a caterpillar are reflections of one another.

LEMMA 3.1. *For each $S \subseteq \{2, 3, \ldots, n-2\}$ we have $\mathscr{S}_{\mathscr{C}_S} \in \{S, S'\}$.*

PROOF. If $S \subseteq \{2, 3, \ldots, n-2\}$, then there are n vertices in $C = \mathscr{C}_S$ and $n-1$ edges. Suppose an endpoint of the spine of C is chosen to be initially marked, and the labelling procedure has been completed, producing a spine set T.

When the internal edge e of C was labelled k there must have been $n-1-k$ edges left to label. Further, starting the labelling procedure from the opposite endpoint of the spine, when we reach the point of labelling e, these $n-k-1$ edges are exactly the edges that have been labelled. Thus e will be labelled $n-k$, as required. $\square$

Since this result shows $\mathscr{S}'_C$ is also a spine set for the caterpillar C, Equation 9 yields

$$\mathscr{C}_{\mathscr{S}'_C} \cong C. \tag{10}$$

COROLLARY 3.2. *If C_1 and C_2 are n-vertex caterpillars, then $C_1 \cong C_2$ if, and only if, either $\mathscr{S}_{C_1} = \mathscr{S}_{C_2}$ or $\mathscr{S}_{C_1} = \mathscr{S}'_{C_2}$.*

PROOF. If $C_1 \cong C_2$, then there is an isomorphism between the two which takes the spine of one caterpillar onto the spine of the other. If we perform the labelling procedure on each caterpillar by starting at the ends of the spine which correspond through the isomorphism, we will produce the same spine set for each caterpillar; that is, $\mathscr{S}_{C_1} = \mathscr{S}_{C_2}$. If we had started the labelling procedure from the opposite end of one of the spines, then the proof of Lemma 3.1 shows that we obtain the reflected spine set. In which case $\mathscr{S}_{C_1} = \mathscr{S}'_{C_2}$.

Conversely, if either $\mathscr{S}_{C_1} = \mathscr{S}_{C_2}$ or $\mathscr{S}_{C_1} = \mathscr{S}'_{C_2}$, then by using either Equation 9 or Equation 10 we obtain $C_1 \cong C_2$, as desired. $\square$

PROPOSITION 3.3. *An n-vertex caterpillar C is symmetric if, and only if, its spine set $\mathscr{S}_C$ is a symmetric subset of $\{2, 3, \ldots, n-2\}$.*

PROOF. If a caterpillar is symmetric, that is, the degree sequence of the spine is palindromic, then the labelling procedure would produce the *same* spine set $\mathscr{S}_C$ from either end. Since we know from the proof of Lemma 3.1 that labelling from the opposite end should give the *reflected* spine set, this shows that if the caterpillar C is symmetric, then so is its spine set $\mathscr{S}_C$.

Conversely, if the degree sequence of the spine is not palindromic, we can easily check that the corresponding spine set is not symmetric. Suppose the degree sequence for C is $\delta = (\delta_1, \delta_2, \ldots, \delta_k)$ where

$$\delta_1 = \delta_k,\ \delta_2 = \delta_{k-1},\ \ldots,\ \text{and } \delta_t = \delta_{k-t+1}, \tag{11}$$

but

$$\delta_{t+1} \neq \delta_{k-t}. \tag{12}$$

Then, labelling the caterpillar from the one end of the spine gives the spine set

$$\{\delta_1, \delta_1 + \delta_2 - 1, \ldots, \delta_1 + \delta_2 + \ldots + \delta_{k-1} - k + 2\}$$

while labelling the caterpillar from the opposite end of the spine gives the spine set

$$\{\delta_k, \delta_k + \delta_{k-1} - 1, \ldots, \delta_k + \delta_{k-1} + \ldots + \delta_2 - k + 2\}.$$

The two spine sets above are written with elements shown in increasing order of size. Thus to check if they are the same sets, we need only check that, in the order shown, the j-th element of one matches the j-th element of the other, for each j. The first t elements of these sets (in the order shown) are the same by Equation 11, but by using both Equations 11 and 12, we find that the $t+1$-th elements, $\delta_1 + \delta_2 + \ldots + \delta_{t+1} - k + 2$ and $\delta_k + \delta_{k-1} + \ldots + \delta_{k-t} - k + 2$ respectively, differ. □

Corollary 3.4. *If C_1 and C_2 are n-vertex caterpillars and at least one of them is symmetric, then $C_1 \cong C_2$ if, and only if, $\mathscr{S}_{C_1} = \mathscr{S}_{C_2}$.*

Proof. Without loss of generality, let C_2 be symmetric. Proposition 3.3 gives $\mathscr{S}_{C_2} = \mathscr{S}'_{C_2}$. Now Corollary 3.2 gives the desired result. □

4. Results on X_C

4.1. A Bound on Coefficients. For each $i \in \mathscr{S}_C$, i corresponds to some internal edge e_i of C, and the graph obtained by removing the edge e_i from C consists of two disjoint caterpillars with $i-1$ and $n-i-1$ edges respectively. Hence the set $F = E - \{e_i\}$ induces the partition

$$\lambda(F) = (i, n-i). \tag{13}$$

Whenever λ is a partition with two parts and C is a caterpillar there is a straightforward bound on the coefficient of p_λ, namely

Proposition 4.1. *Let C be an n-vertex caterpillar and λ have two parts. Then either*

1. *$(-1)^n[p_\lambda]X_C = L(C)$, if $\lambda = (n-1, 1)$, or*
2. *$0 \leq (-1)^n[p_\lambda]X_C \leq 2$ otherwise.*

Proof. From Proposition 2.2, we have $[p_\lambda]X_T = (-1)^n L(T)$ in the case of $\lambda = (n-1, 1)$. Any other partition λ with two parts can only arise as $\lambda(E - \{e_i\})$ for some $i \in \mathscr{S}_C$. We show that any such λ can arise at most twice.

We are looking for occurrences of $\lambda = (j, n-j)$, and λ can only arise from the edges, if there are any, which would correspond to the potential elements j and $n-j$ of $\mathscr{S}_C$. Thus the magnitude of the coefficient of p_λ could be at most 2, if both j, $n-j \in \mathscr{S}_C$. □

From the proof of Proposition 4.1, we have the following fact.

Corollary 4.2. *If $\lambda = (j, n-j)$, $1 < j < n$, is a partition of n into two parts and C is a n-vertex caterpillar, then $[p_\lambda]X_C = (-1)^n|\{j, n-j\} \cap \mathscr{S}_C|$.*

4.2. Symmetric and Near-Symmetric Caterpillars.

THEOREM 4.3. *The chromatic symmetric function distinguishes the symmetric caterpillars from the nonsymmetric caterpillars. Further, it distinguishes among the symmetric caterpillars.*

PROOF. Let C be a symmetric caterpillar. We saw in Proposition 3.3 that a given n-vertex caterpillar C was symmetric if, and only if, $\mathscr{S}_C$ is symmetric. That is, $n-j \in \mathscr{S}_C$ if, and only if, $j \in \mathscr{S}_C$.

In the case when n is even and $j = \frac{n}{2}$ we have $j = n-j$, but otherwise we have $j \neq n-j$. Thus, for a symmetric caterpillar, Corollary 4.2 gives

$$(-1)^n[p_{(j,n-j)}]X_C = 0 \ \text{ if } \ j \notin \mathscr{S}_C, \tag{14}$$

$$(-1)^n[p_{(j,n-j)}]X_C = 2 \ \text{ if } \ j \in \mathscr{S}_C \text{ and } j \neq \frac{n}{2}, \tag{15}$$

and if n is even, then

$$(-1)^n[p_{(\frac{n}{2},\frac{n}{2})}]X_C = 1 \ \text{ if } \ \frac{n}{2} \in \mathscr{S}_C. \tag{16}$$

We have shown all symmetric caterpillars satisfy Equations 14, 15, and 16. Conversely, if a caterpillar satisfies Equations 14, 15, and 16, we shall show it is symmetric. Let a caterpillar C satisfy Equations 14, 15, and 16 and let j be a member of $\mathscr{S}_C$. To show C is symmetric, we need only show that $n-j \in \mathscr{S}_C$. If $j = \frac{n}{2}$, then $n-j = j$, so immediately $n-j \in \mathscr{S}_C$. If $j \neq \frac{n}{2}$, then by Equation 15 and Corollary 4.2 we find $n-j \in \mathscr{S}_C$, as required.

Hence we can use the chromatic symmetric function to distinguish the symmetric caterpillars from those that are nonsymmetric. Further, by Equations 14, 15, 16, and Corollary 4.2, the spine set of the caterpillar can be determined from its chromatic symmetric function. From Equation 9 we know that the spine set of a caterpillar determines the caterpillar. Thus chromatic symmetric function distinguishes the symmetric caterpillars from one another.

□

We can now make a slight perturbation of Theorem 4.3. Towards this end, we shall say that a nonsymmetric caterpillar C is *near-symmetric* if $\mathscr{S}_C \cup \{i\}$ is a symmetric subset for some number $i \in \{2, 3, \ldots n-2\}$.

Example: The caterpillar C with 11 vertices whose spine set is $\mathscr{S}_C = \{3, 4, 8\}$ is near-symmetric, as $\{3, 4, 7, 8\}$ is a symmetric subset of $\{2, 3, \ldots, 9\}$.

THEOREM 4.4. *The chromatic symmetric function distinguishes the near-symmetric caterpillars from those caterpillars which are not near-symmetric. Further, it distinguishes among the near-symmetric caterpillars.*

PROOF. Let C be a near-symmertric caterpillar, say with $\mathscr{S}_C \cup \{n-i\}$ being a symmetric subset. Then necessarily $i \in \mathscr{S}_C$. Looking at the coefficients of p_λ in X_C for partitions into two parts gives

$$(-1)^n[p_{(j,n-j)}]X_C = 0 \ \text{ if } \ j \notin \mathscr{S}_C, \tag{17}$$

$$(-1)^n[p_{(j,n-j)}]X_C = 2 \ \text{ if } \ j \in \mathscr{S}_C \text{ and } j \neq \frac{n}{2}, \ i, \tag{18}$$

$$(-1)^n[p_{(i,n-i)}]X_C = 1, \tag{19}$$

and if n is even, then

$$(-1)^n[p_{(j,n-j)}]X_C = 1 \ \text{ if } \ j = \frac{n}{2} \in \mathscr{S}_C. \tag{20}$$

Conversely, any caterpillar C which satisfies Equations 17, 18, 19, and 20 for some value i is found to be near-symmetric upon considering Corollary 4.2, as adding $n-i$ to $\mathscr{S}_C$ creates a symmetric subset.

As before, from Equations 17, 18, 19, and 20 and Corollary 4.2 we see that the chromatic symmetric function of a near-symmetric caterpillar determines the spine set $\mathscr{S}_C$ of the caterpillar. Then by Equation 9, we can recover C from $\mathscr{S}_C$. □

Combining Theorems 4.3 and 4.4 we obtain the following result.

Theorem 4.5. *Let $\mathcal{C}$ be the set of caterpillars and $\mathcal{S}$ be the set of caterpillars that are either symmetric or near-symmetric. Then if $C_1 \in \mathcal{C}$ and $C_2 \in \mathcal{S}$, we have $X_{C_1} = X_{C_2}$ if, and only if, $C_1 \cong C_2$.*

5. Counting Symmetric Caterpillars

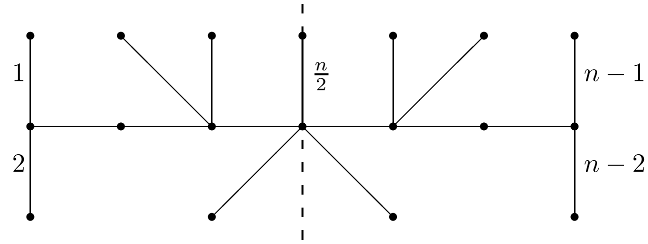
Proposition 5.1. *Let $\mathcal{S}(n,k)$ denote the number of nonisomorphic n-vertex symmetric caterpillars with k internal edges.*

(1) *If k is even, $\mathcal{S}(n,k) = \binom{\lfloor \frac{n-3}{2} \rfloor}{\frac{k}{2}}$.*

(2) *If k is odd, then*
   (a) *$\mathcal{S}(n,k) = \binom{\frac{n}{2}-2}{\frac{k-1}{2}}$ when n is even, and*
   (b) *$\mathcal{S}(n,k) = 0$ when n is odd.*

Proof. Suppose k, the number of edges in the spine, is even. Then visually the line of symmetry of C crosses the spine at the vertex in the center of the spine.

If n is even, then $n-1$, the number of edges, is odd. Hence one of the edges is forced to be along the line of symmetry of C. Under our labelling procedure, and by redrawing if necessary, we can assume the edge along the line of symmetry is labelled $\frac{n}{2}$. For example:

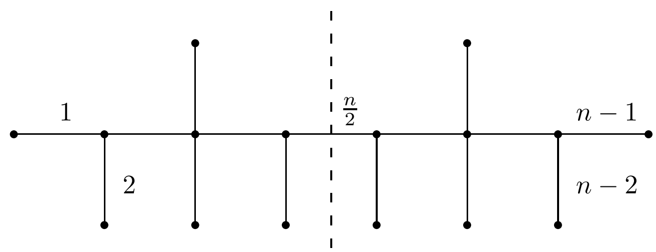


From the symmetry of the caterpillar, the rest of the caterpillar is determined once we know which $\frac{k}{2}$ of the first $\frac{n}{2}-1$ edges are internal edges. Thus we count the number of sets $S \subseteq \{2, 3, \ldots, \frac{n}{2}-1\}$ with $\frac{k}{2}$ elements, giving $\binom{\frac{n}{2}-2}{\frac{k}{2}} = \binom{\frac{n-4}{2}}{\frac{k}{2}} = \binom{\lfloor\frac{n-3}{2}\rfloor}{\frac{k}{2}}$ symmetric caterpillars.

If n is odd, then the number of edges is even and, by redrawing if necessary, $\frac{n-1}{2}$ of the $n-1$ edges lie on each side of the line of symmetry. Further, knowing which of the first $\frac{n-1}{2}$ edges are internal edges determines the caterpillar. Thus we count sets of the form $S \subseteq \{2, 3, \ldots, \frac{n-1}{2}\}$ containing $\frac{k}{2}$ elements, obtaining $\binom{\frac{n-1}{2}-1}{\frac{k}{2}} = \binom{\frac{n-3}{2}}{\frac{k}{2}} = \binom{\lfloor\frac{n-3}{2}\rfloor}{\frac{k}{2}}$ symmetric caterpillars. This completes the proof of 1.

Now if k, the number of edges in the spine, is odd, then visually the line of symmetry of C bisects the central edge of the spine. Apart from this edge, every other edge is paired with its reflection across the line of symmetry. Thus the total number of edges is odd, forcing n to be even. This gives $\mathcal{S}(n,k) = 0$ for odd n.

If we assume n is even and C is symmetric, then the central edge of the spine of C is labelled $\frac{n}{2}$ by our labelling procedure, and, as before, knowing the internal edges of one side of the caterpillar determines the other.



Thus we seek to count all sets of the form $S \subseteq \{2, 3, \ldots, \frac{n}{2}-1\}$ containing $\frac{k-1}{2}$ elements. This gives $\binom{\frac{n}{2}-2}{\frac{k-1}{2}}$ symmetric caterpillars in this final case. $\square$

From this result, one can check that the total number of symmetric caterpillars with n vertices is $2^{\lfloor\frac{n-2}{2}\rfloor}$. A more direct approach can be found in [**Harary/ Schwenk, 1973**], where it is also shown that the total number of caterpillars with n vertices is $2^{n-4} + 2^{\lfloor\frac{n-4}{2}\rfloor}$.

6. Conclusions

In the majority of this paper we have remained within the context of caterpillars as opposed to trees in general. As previously noted, there are $2^{n-4} + 2^{\lfloor\frac{n-4}{2}\rfloor}$ n-vertex caterpillars. We have proved the result in the case of symmetric caterpillars, of which there are $2^{\lfloor\frac{n-2}{2}\rfloor}$, and also in the case of near-symmetric caterpillars [**Morin, 2005**].

By collecting various results, we find that we have proved the following.

Proposition 6.1. *There are collections $\mathcal{Q}_n$ of n-vertex graphs such that:*

(1) $\lim_{n\to\infty} |\mathcal{Q}_n| = \infty$,
(2) $\chi(G_1, k) = \chi(G_2, k)$ *for every pair of graphs* G_1, $G_2 \in \mathcal{Q}_n$, *and*
(3) *If* G_1, $G_2 \in \mathcal{Q}_n$ *and* $X_{G_1} = X_{G_2}$, *then* $G_1 = G_2$.

Proof. We look at the collection $\mathcal{Q}_n$ of symmetric caterpillars with n vertices. We have Property 1 by Proposition 5.1. Since all the caterpillars in $\mathcal{Q}_n$ has n vertices, we have $\chi(G,k) = k(k-1)^{n-1}$ for each $G \in \mathcal{Q}_n$. Finally Theorem 4.3 gives Property 3. □

We note that $\mathcal{Q}_n$ could have also been chosen to be the set of near-symmetric n-vertex caterpillars in the above proof.

References

[Harary/ Schwenk, 1973] Frank Harary and Allen J. Schwenk, "The Number of Caterpillars", *Discrete Mathematics* **6** (1973), 359-365.

[Morin, 2005] Matthew Morin, *Caterpillars, Ribbons, and The Chromatic Symmetric Function.* M.Sc. Thesis, University of British Columbia (2005).

[Stanley, 1995] Richard P. Stanley, "A Symmetric Function Generalization of a Chromatic Polynomial of a Graph," *Advances In Mathematics* **111** (1995), 166-194.

Department of Mathematics, University of British Columbia, Canada

E-mail address: mjmorin@interchange.ubc.ca

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Green polynomials at roots of unity and its application

Hideaki Morita

ABSTRACT. We consider Green polynomials at roots of unity. We obtain a recursive formula for Green polynomials at appropriate roots of unity, which is described in a combinatorial manner. The coefficients of the recursive formula are realized by the number of permutations satisfying a certain condition, which leads to interpretation of a combinatorial property of certain graded modules of the symmetric group in terms of representation theory.

RÉSUMÉ. Nous étudions les polynômes de Green évalués en les racines de l'unité. Nous obtenons une formule récursive pour ces polynômes en certaines racines de l'unité, que nous décrivons combinatoirement. Les coefficients de cette formule récursive énumèrent certaines permutations, ce qui permet d'interpréter une propriété combinatoire de certains modules du groupe symétrique, en termes de la théorie de la représentation.

1. Introduction

The Green polynomials $Q^{\mu}_{\rho}(q)$ at roots of unity are considered. We handle Green polynomials $Q^{\mu}_{\rho}(q)$ of type A for any partition μ, and consider the behavior of them at l-th roots of unity ζ_l, where l is not larger than the maximum multiplicity M_μ of μ. We describe a certain recursive formula of Green polynomials $Q^{\mu}_{\rho}(q)$ at $q = \zeta_l$ for the partition ρ satisfying $Q^{\mu}_{\rho}(\zeta_l) \neq 0$. The results of Lascoux-Leclerc-Thibon on Hall-Littlewood functions at roots of unity play an important role in the argument. Our result includes the result of Lascoux-Leclerc-Thibon on Green polynomials as a special case.

We also consider the recursive formula in terms of representation theory of the symmetric group S_n. It is known that the Green polynomials give the graded characters of a family of graded representations of the symmetric group, called the DeConcini-Procesi-Tanisaki algebras, which includes the coinvariant algebra as a special case. The DeConcini-Procesi-Tanisaki algebra R_μ was first introduced by C. DeConcini and C. Procesi [**DP**] as an algebraic model of the cohomology ring of a certain subvariety of the flag variety parametrized by a partition μ, and T. Tanisaki [**T**] gives simple generators of the defining ideal of the algebra, described by combinatorial information on the partition μ. The DeConcini-Procesi-Tanisaki algebra R_μ has a structure of graded S_n-modules, and the Green polynomial $Q^{\mu}_{\rho}(q)$ gives its graded character values at the conjugacy class of which cycle type is ρ. The recursive formula is equivalent to some representation theoretical interpretation of a certain combinatorial property on the Hilbert polynomial $\mathrm{Hilb}_{R_\mu}(q)$ of R_μ, that is, $\mathrm{Hilb}_{R_\mu}(q)$ has l-th roots of unity ζ_l^j $(j = 1, 2, \ldots, l-1)$ as its zeros for each positive integer l not larger than the maximum multiplicity M_μ of μ. This property of the Hilbert polynomial is equivalent to the fact that the direct sums $R_\mu(k;l)$ $(k = 0, 1, \ldots, l-1)$ of the homogeneous components of R_μ of which degrees are congruent to k modulo l, have the same dimension. The recursive formula shows that there exists a subgroup $H_\mu(l)$ of S_n and $H_\mu(l)$-modules $Z_\mu(k;l)$ of equal dimension such that each $R_\mu(k;l)$ is induced from the corresponding $H_\mu(l)$-modules $Z_\mu(k;l)$ for each $k = 0, 1, \ldots, l-1$, which could be regarded as a representation theoretical interpretation of the property 'coincidence of dimensions'. This work is a sequel of [Mt, MN1, MN2].

2000 *Mathematics Subject Classification.* Primary 20C30; Secondary 05E05.

Key words and phrases. Symmetric group, Green polynomial, Hall-Littlewood function, root of unity, Springer representation.

2. Preliminaries

We follow [M] for fundamental notation. Let n be a positive integer and μ a partition of n. Define M_μ to be the maximum multiplicity of the partition μ:

$$M_\mu := \max\{m_1(\mu), m_2(\mu), \cdots, m_n(\mu)\},$$

where $m_i = m_i(\mu)$ denotes the multiplicity of i in the sequence μ. Let μ and ρ be partitions and let q be an indeterminate. The Green polynomial $X^\mu_\rho(q)$ is defined to be the coefficients of the Hall-Littlewood function $P_\mu(x;q)$ in the linear expansion

$$p_\rho(x) = \sum_\mu X^\mu_\rho(q) P_\mu(x;q),$$

where $p_\rho(x)$ denotes the power-sum function corresponding to the partition ρ, and the sum is over partitions μ of the same size as ρ. We also define the polynomial $Q^\mu_\rho(q)$ for partitions μ and ρ of the same size by

$$Q^\mu_\rho(q) = q^{n(\mu)} X^\mu_\rho(q^{-1}),$$

where $n(\mu) = \sum_{i\geq 1}(i-1)\mu_i$ if $\mu = (\mu_1, \mu_2, \dots)$. The polynomial $Q^\mu_\rho(q)$ is also called the Green polynomial. The Green polynomial $Q^\mu_\rho(q)$ is a polynomial with integer coefficients whose degree is $n(\mu)$, which was introduced by J. A. Green [**Gr**] to describe irreducible character values of the general linear group $GL_n(\mathbf{F}_q)$ over a finite field $\mathbf{F}_q$.

Let $\varphi_r(q)$ be the polynomial $(1-q)(1-q^2)\cdots(1-q^r)$, and $b_\mu(q)$ the polynomial

$$b_\mu(q) = \prod_{i\geq 1} \varphi_{m_i(\mu)}(q),$$

where $m_i(\mu)$ is the multiplicity of i in the partition μ. Define

$$Q_\mu(x;q) = b_\mu(q) P_\mu(x;q),$$

which are referred to, as well as the P_μ, as Hall-Littlewood functions. If we replace the variables $x = (x_1, x_2, \dots)$ of $Q_\mu(x;q)$ by

$$x/(1-q) = (x_1, x_2, \dots; qx_1, qx_2, \dots; q^2x_1, q^2x_2, \dots),$$

then we obtain the *modified* Hall-Littlewood function, which is denoted by

$$Q'_\mu(x;q) \left(= Q_\mu\left(\frac{x}{1-q}; q\right)\right).$$

Equivalently, it is also defined by replacing $p_k(x)$ by $p_k(x)/(1-t^k)$ after expressing $Q_\mu(x;t)$ as a polynomial in $\{p_k(x) | k \geq 1\}$. It is known (see, e.g., [**DLT**]) that the Green polynomial $X^\mu_\rho(q)$ is obtained as the inner product value

$$X^\mu_\rho(x) = \langle Q'_\mu(x;q), p_\rho(x)\rangle$$

of the modified Hall-Littlewood function $Q'_\mu(x;q)$ and the power-sum function $p_\rho(x)$. The inner product $\langle\cdot,\cdot\rangle$ of the ring $\Lambda[q]$ is defined by $\langle s_\lambda, s_\mu\rangle = \delta_{\lambda\mu}$, where s_λ denotes the Schur function corresponding to the partition λ, and $\delta_{\lambda\mu}$ the Kronecker delta.

In the rest of this section, we recall results on (modified) Hall-Littlewood functions at roots of unity due to Lascoux-Leclerc-Thibon [**LLT**]. Let $\mu \vdash n$ be a partition, l an integer such that $2 \leq l \leq M_\mu$ be fixed, and $m_i(\mu) = lq_i + r_i$, $0 \leq r_i \leq l-1$, for each i. Set $q = q_1 + 2q_2 + \cdots + nq_n$ and $r = r_1 + 2r_2 + \cdots + nr_n$. Let $\tilde{\mu}(l)$ and $\bar{\mu}(l)$ be the partitions

$$\tilde{\mu}(l) := (1^{lq_1} 2^{lq_2} \cdots n^{lq_n})$$

and

$$\bar{\mu}(l) := (1^{r_1} 2^{r_2} \cdots n^{r_n}).$$

It is clear that the partition μ decomposes into the disjoint union $\mu = \tilde{\mu}(l) \cup \bar{\mu}(l)$. Also define

$$\tilde{\mu}(l)^{1/l} := (1^{q_1} 2^{q_2} \cdots n^{q_n}),$$

which is a partition of q.

Example 2.1. If $\mu = (3,3,3,2,2,1)$, then $M_\mu = 3$. Let $l = 2$ be fixed. Then $\tilde{\mu}(l) = (3,3,2,2)$, $\bar{\mu}(l) = (3,1)$, and $\mu = (3,3,2,2) \cup (3,1)$. Also the partition $\tilde{\mu}(l)^{1/l}$ is $(3,2)$.

Let μ be a partition, and l a positive integer such that $l \leq M_\mu$. The modified Hall-Littlewood function $Q'_\mu(x;q)$ at $q = \zeta_l$, a primitive l-th root of unity, is factorized in such a way that is consistent with the decomposition of the partition $\mu = \tilde{\mu}(l) \cup \bar{\mu}(l)$.

Proposition 2.1 ([LLT, Theorem 2.1.]). $Q'_\mu(x;\zeta_l) = Q'_{\bar{\mu}(l)}(x;\zeta_l) \prod_{i\geq 1} \left(Q'_{(i^l)}(x;\zeta_l)\right)^{q_i}$.

Example 2.2. Let $\mu = (3,3,3,2,1,1,1,1,1)$ and $l = 2$. Then $\bar{\mu}(l) = (3,2,1)$, and we have

$$Q'_{(3,3,3,2,1,1,1,1,1)}(x;\zeta_2) = Q'_{(3,2,1)}(x;\zeta_2) Q'_{(3^2)}(x;\zeta_2) \left(Q'_{(1^2)}(x;\zeta_2)\right)^2.$$

Proposition 2.2 ([LLT, Theorem 2.2.]). $Q'_{(i^l)}(x;\zeta_l) = (-1)^{(l-1)i}(p_l \circ h_i)(x)$, *where* $(p_l \circ h_i)(x)$ *denotes the plethysm.*

Remark 2.3. Note that

$$(p_l \circ h_i)(x) = \sum_{\lambda \vdash i} z_\lambda^{-1} p_{l\lambda}(x), \tag{2.1}$$

Thus we have for example $Q'_{(3^2)}(x;\zeta_2) = (-1)^{(2-1)3}(p_2 \circ h_3)(x) = -z_{(3)}^{-1} p_{(6)}(x) - z_{(2,1)}^{-1} p_{(4,2)} - z_{(1,1,1)}^{-1} p_{(2,2,2)}(x)$.

It follows from Proposition 2.1, Proposition 2.2 and (2.1) that the Green polynomial corresponding to a rectangular partition $\mu = (r^k)$ at a primitive k-th root of unity is described by a certain 'smaller' Green polynomial.

Proposition 2.3 ([LLT, Theorem 3.2.]). *Let* $\mu = (r^k)$ *be a rectangular partition,* ζ_k *a primitive* k*-th root of unity. If* $m_i(\mu) \geq 1$ *for some* $i \geq 1$ *divisible by* k*, then it holds that*

$$X^\mu_\rho(\zeta_k) = (-1)^{(k-1)j} k X^{((r-j)^k)}_{\rho\setminus\{i\}}(\zeta_k), \tag{2.2}$$

where $i = jk$.

If we rewrite the identity (2.2) in terms of the polynomial $Q^\mu_\rho(x)$, then the sign $(-1)^{(k-1)j}$ is vanished and we have [Mt, Lemma 7 or Proposition 5]

$$Q^\mu_\rho(\zeta_k) = k Q^{((r-j)^k)}_{\rho\setminus\{i\}}(\zeta_k).$$

Applying this identity repeatedly, we also have

$$Q^\mu_\rho(\zeta_k) = k^{l(\rho)},$$

if the partition ρ consists of multiples of k.

3. Roots of unity

Let μ be a partition of n l a positive integer such that $2 \leq l \leq M_\mu$ be fixed, and $m_i(\mu) = lq_i + r_i$, $0 \leq r_i \leq l-1$, for each i. Set $q = q_1 + 2q_2 + \cdots + nq_n$ and $r = r_1 + 2r_2 + \cdots + nr_n$. Let $\tilde{\mu}(l)$, $\bar{\mu}(l)$, and $\tilde{\mu}(l)^{1/l}$ be as in the previous section. We define 'partitions of a partition'as follows. Let $\nu = (\nu_1, \nu_2, \ldots, \nu_d)$ be a partition of n. A *partition* of the partition ν is by definition a sequence of partitions

$$\lambda = (\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(d)})$$

such that $\lambda^{(i)} \vdash \nu_i$ for each $i = 1, 2, \ldots, d$, which is denoted by $\lambda \vdash \nu$. We distinguish any nontrivial permutation of $\lambda = (\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(d)})$ from the original one. For example, we consider that the following two partitions $((2),(1,1)), ((1,1),(2))$ are different as partitions of $(2,2)$. The *length* $l(\lambda)$ of $\lambda \vdash \nu$ is defined by

$$l(\lambda) = \sum_{i=1}^{d} l(\lambda^{(i)}),$$

and the *size* $|\lambda|$ is defined by the sum of sizes of the components $\lambda^{(i)}$ of λ, which is equal to $n = |\nu|$. Also define

$$z_\lambda := \prod_{i\geq 1} z_{\lambda^{(i)}},$$

where z_π is defined by

$$z_\pi = 1^{m_1} m_1! 2^{m_2} m_2! \cdots n^{m_n} m_n!$$

for a partition $\pi = (1^{m_1} 2^{m_2} \cdots n^{m_n}) \vdash n$ of a positive integer as usual. Let $\nu = (\nu_i)$ be a partition of n and $\lambda = (\lambda^{(i)})$ a partition of ν. Let

$$m_k(\lambda) := \sum_{i=1}^{d} m_k(\lambda^{(i)})$$

for each possible $k \geq 1$. Then define

$$m_\lambda := \prod_{k \geq 1} \binom{m_k(\lambda)}{m_k(\lambda^{(1)}), m_k(\lambda^{(2)}), \ldots, m_k(\lambda^{(d)})}.$$

Also, for each positive integer l, let $l\lambda$ denotes the partition whose components are those of λ multiplied by l.

Example 3.1. Let $\nu = (4,2)$. Then the partitions λ of ν are $((4),(2))$, $((3,1),(2))$, $((2,2),(1,1))$, $((2,1,1),(1,1))$ and so on. Suppose that $\lambda = ((2,1,1),(2)) \vdash \nu$. Then m_λ is computed as follows: $m_{((2,1,1),(2))} = \binom{m_1(\lambda)}{m_1(\lambda^{(1)}), m_1(\lambda^{(2)})}\binom{m_2(\lambda)}{m_2(\lambda^{(1)}), m_2(\lambda^{(2)})} = \binom{2}{2,0}\binom{2}{1,1} = 2$. For the same λ, if $l = 2$ for example, the partition $l\lambda = 2\lambda$ is $(4,4,2,2)$.

Let ρ be a partition and ν a *subpartition* of ρ, i.e., $m_i(\nu) \leq m_i(\rho)$ for each possible $i \geq 1$. Then we define the *binomial coefficient* $\binom{\rho}{\nu}$ by

$$\binom{\rho}{\nu} := \prod_{i \geq 1} \binom{m_i(\rho)}{m_i(\nu)}.$$

Let μ be a partition, and l an integer such that $2 \leq l \leq M_\mu$ be fixed. For a partition ν of $|\tilde{\mu}(l)|$, define

$$C(\nu, \mu; l) := \sum_{\substack{\pi \vdash \tilde{\mu}(l)^{1/l} \\ l\pi = \nu}} m_\pi.$$

If there exists no $\pi \vdash \tilde{\mu}(l)^{1/l}$ such that $l\pi = \nu$, then $C(\nu, \mu; l) = 0$.

Example 3.2. Let $\mu = (5,4,4,2,2,1)$, and l such that $2 \leq l \leq M_\mu$ fixed, say $l = 2$. Then $\tilde{\mu}(l) = (4,4,2,2)$ and $\tilde{\mu}(l)^{1/2} = (4,2)$. Suppose that $\nu = (4,4,4) \vdash |\tilde{\mu}(l)|$. Then there exists only one $\pi \vdash \tilde{\mu}(l)^{1/2}$ such that $2\pi = \nu$, i.e., $\pi = ((2,2),(2))$. Hence $C(\nu, \mu; 2) = m_{((2,2),(2))} = \binom{3}{2,1} = 3$. On the other hand, if $\nu = (4,4,2,2)$, then there exist two $\pi \vdash (4,2)$ such that $2\pi = \nu$, i.e., $\pi = ((2,2),(1,1)), ((2,1,1),(2))$. Hence we have $C(\nu, \mu; 2) = m_{((2,2),(1,1))} + m_{((2,1,1),(2))} = \binom{2}{0,2}\binom{2}{2,0} + \binom{2}{2,0}\binom{2}{1,1} = 1 + 2 = 3$. On the other hand, in the case where $\tilde{\mu}(l)$ is given by $(4,4)$ for $l = 2$ and $\nu = (4,2,2)$, the partitions $\pi \vdash \tilde{\mu}(l)^{1/l}$ satisfying $l\pi = \nu$ are $\pi = ((2),(1,1)), ((1,1)(2))$. Since we distinguish these two partitions, $C(\nu, \mu; l)$ is obtained by $m_{((2),(1,1))} + m_{((1,1),(2))} = 1 + 1 = 2$.

Now we can state our main result, which retrieves LLT's result, Proposition 2.3, if we consider the case where μ is a rectangle and $l = M_\mu$. Proposition 2.1 and Proposition 2.2 play a crucial role in the proof.

Theorem 3.3. *Let $\mu = (1^{m_1} 2^{m_2} \cdots n^{m_n})$ be a partition of n, a positive integer $l = 1, 2, \ldots, M_\mu$ fixed, and ζ_l an l-th primitive root of unity. Let $m_i = lq_i + r_i$, $0 \leq r_i \leq l-1$, for each $i = 1, 2, \ldots, n$. Let $r = r_1 + 2r_2 + \cdots + nr_n$, and $\bar{\mu}(l) = (i^{r_i}) \vdash r$.*

Then we have:

1. $Q^\mu_\rho(\zeta_l) \neq 0 \Longrightarrow \rho = l\tilde{\rho} \cup \bar{\rho}$ *for some* $\tilde{\rho} \vdash \tilde{\mu}(l)^{1/l}$ *and* $\bar{\rho} \vdash r$.
2. *For such a partition $\rho = l\tilde{\rho} \cup \bar{\rho}$, it holds that:*

$$Q^\mu_\rho(\zeta_l) = \sum_{\substack{\nu \vdash |\tilde{\mu}(l)| \\ \nu \subset \rho}} \binom{\rho}{\nu} C(\nu, \mu; l) l^{l(\nu)} Q^{\bar{\mu}(l)}_{\rho \setminus \nu}(\zeta_l).$$

Example 3.4. Let $\mu = (5,4,4,2,2,1) \vdash 18$ and $l = 2$. In this case, we have $\widetilde{\mu(2)} = (4,4,2,2)$ and $\widetilde{\mu(2)}^{1/2} = (4,2)$. Suppose that $\rho = (4,4,2,2) \cup (4,2) = (4,4,4,2,2,2)$. Then subpartitions ν of ρ which satisfy $\nu \vdash |\widetilde{\mu(2)}| = 12$ are $\nu = (4,4,4), (4,4,2,2)$. Consider the case where $\nu = (4,4,4)$. Then

$\binom{\rho}{\nu} = \binom{3}{0}\binom{3}{3} = 1$. There exists only one $\lambda \vdash \tilde{\mu(2)}^{1/2} = (4,2)$ such that $2\lambda = (4,4,4)$, i.e., $\lambda = ((2,2),(2))$, and we have $m_\lambda = \binom{2+1}{2,1} = 3$. Thus $C(\nu,\mu;2) = 3$. If $\nu = (4,4,2,2)$, then $\binom{\rho}{\nu} = \binom{3}{2}\binom{3}{2} = 9$. The corresponding λ's satisfying $2\lambda = \nu$ are $\lambda = ((2,2),(1,1)),((2,1,1),(2))$, and $m_{((2,2),(1,1))} = \binom{2}{0,2}\binom{2}{2,0} = 1$, $m_{((2,1,1),(2))} = \binom{2}{2,0}\binom{2}{1,1} = 2$. Hence we have $C(\nu,\mu;2) = 3$ in this case. Thus we have $Q^{(5,4,4,2,2,1)}_{(4,4,4,2,2,2)}(\zeta_2) = \binom{\rho}{(4,4,4)}C((4,4,4),\mu;2)2^{l(4,4,4)}Q^{\tilde{\mu}(l)}_{\rho\setminus(4,4,4)}(\zeta_2) + \binom{\rho}{(4,4,2,2)}C((4,4,2,2),\mu;2)2^{l(4,4,2,2)}Q^{\tilde{\mu}(l)}_{\rho\setminus(4,4,2,2)}(\zeta_2) = 1 \times 3 \times 8Q^{(5,1)}_{(2,2,2)}(\zeta_2) + 9 \times 3 \times 16Q^{(5,1)}_{(4,2)}(\zeta_2)$.

4. Permutation enumeration

In this section, we shall give a combinatorial characterization of the coefficients

$$\binom{\rho}{\nu}C(\nu,\mu;l)l^{l(\nu)},$$

in the preceding formula. Let μ be a partition of a positive integer n, and an integer $l \in \{2,3,\dots,M_\mu\}$ fixed. We define a product of cyclic permutations $a = a_\mu(l)$ corresponding to μ and l as follows. To avoid abuse of notation, we shall see the definition by the following example. It is clear from the definition that the element $a_\mu(l)$ has the order l.

EXAMPLE 4.1 (Definition of $a_\mu(l)$). Let $\mu = (3,3,2,2,2,1)$ and $l = 2(\leq M_\mu = 3)$. We fix the numbering of the Young diagram of μ as follows:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & \\ 9 & 10 & \\ 11 & 12 & \\ 13 & & . \end{array}$$

Corresponding to the number $l = 2$, we extract subtableaux

$$\begin{array}{ccc} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}, \quad \begin{array}{cc} 7 & 8 \\ 9 & 10 \end{array}$$

Then the cyclic permutation product $a_\mu(2)$ is defined by using the letters corresponding to $\tilde{\mu}(l)$ as follows:

$$a_\mu(2) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 7 & 8 & 9 & 10 \\ 9 & 10 & 7 & 8 \end{pmatrix}$$

□

Let $n = ql + r$, $0 \leq r \leq l-1$. Recall that $\tilde{\mu}(l)$ is a partition of $n - r$. Let $S_{\tilde{\mu}(l)}$ be the Young subgroup which permutes the letters corresponding to $\tilde{\mu}(l)$ in the preceding tableau, and let S_r be the subgroups which permutes the remaining letters. It is obvious that elements of these groups commute with each other. In the preceding definition (Example 11), these groups are the following:

$$\begin{aligned} S_{\tilde{\mu}(l)} &= S_{\{1,2,3\}} \times S_{\{4,5,6\}} \times S_{\{7,8\}} \times S_{\{9,10\}}, \\ S_r &= S_{\{11,12,13\}}, \end{aligned}$$

where $\tilde{\mu}(l) = (3,3,2,2)$, $r = 3$ and $S_{\{i,j,\dots,k\}}$ denotes the symmetric group of the letters $\{i,j,\dots,k\}$. Consider the subgroup of S_n

$$H_\mu(l) := \left(S_{\tilde{\mu}(l)} \times S_r\right) \rtimes \langle a_\mu(l)\rangle = \left(S_{\tilde{\mu}(l)} \rtimes \langle a_\mu(l)\rangle\right) \times S_r.$$

The following lemma is proved by straightforward computation.

LEMMA 4.2. *The cycle types ρ of elements of the subgroup $H_\mu(l)$ are of the form*

$$\rho = l\tilde{\rho} \cup \bar{\rho},$$

where $\tilde{\rho} \vdash \tilde{\mu}(l)^{1/l}$ and $\bar{\rho} \vdash r$. Conversely, if ρ is a partition of such a form, then there exists an element of $H_\mu(l)$ whose cycle type is ρ.

Example 4.3. Consider the case $\mu = (3,3,2,2,2,1)$ and $l = 2$. Then the corresponding cyclic permutation product is $a_\mu(2) = (1,4)(2,5)(3,6)(7,9)(8,10)$. If we consider $w = (1,2)(7,8)a_\mu(2)(11,13) \in H_\mu(2)$, then $w = (1,4,2,5)(3,6)(7,9,8,10)(11,13)$ and its cycle type is $(4,4,2,2,1)$, which is the union of $(4,4,2)$ and $(2,1)$. The partition $(4,4,2)$ is written in the form $(4,4,2) = 2((2,1),(2))$ for $((2,1),(2)) \vdash (3,2) = \tilde{\mu}(l)^{1/2}$. Conversely, if we consider $\rho = 2((2,1),(1,1)) \cup (3) = (4,3,2,2,2)$, then choose $\tau_1 = (1,2) \in S_{\tilde{\mu}(l)}$ and $\tau_2 = (11,12,13) \in S_r$ for example. It is easy to see that the cycle type of $w = \tau_1\tau_2 a_\mu(2)$ coincides with ρ. $\square$

A direct but a little complicated enumeration shows the following proposition. Remark that $l(\lambda) = l(k\lambda)$ for any partition λ and any positive integer k.

Proposition 4.1. *Let $\mu \vdash n$ be a partition, $l = 2,3,\ldots,M_\mu$ fixed, and $a = a_\mu(l)$ the cyclic permutation product corresponding to μ and l. Let $\rho \vdash n$ be a partition of the form $\rho = l\tilde{\rho} \cup \bar{\rho}$ where $\tilde{\rho} \vdash \tilde{\mu}(l)^{1/l}$ and $\bar{\rho} \vdash r$. Suppose that $w \in S_n$ be a permutation whose cycle type is ρ. Then it follows that*

$$\binom{\rho}{l\tilde{\rho}} C(l\tilde{\rho},\mu;l) l^{l(\tilde{\rho})} = \sharp\{\sigma \in S_n/S_{\tilde{\mu}(l)} \times S_r | w\sigma a^{-1} \equiv \sigma \ mod\ S_{\tilde{\mu}(l)} \times S_r\}.$$

Example 4.4. Let $\mu = (2,2,2,2,2,1)$ and $l = 2,\ldots,M_\mu(=5)$ be fixed, say $l = 2$. Then the corresponding product of cyclic permutations is $a = (13)(24)(57)(68)$. The subgroups $S_{\tilde{\mu}(l)}$ and $S_r = S_3$ are $S_{\{1,2\}} \times S_{\{3,4\}} \times S_{\{5,6\}} \times S_{\{7,8\}}$ and $S_{\{9,10,11\}}$ respectively. Let us consider the case $w = (12)a(9,10) = (1324)(57)(68)(9,10)$ ($\tau_1 = (12)$, $\tau_2 = (9,10)$). The cycle type ρ of w is $\rho = (4,2,2,2,1)$. If we let $\tilde{\rho} = ((2),(1,1)) \vdash \tilde{\mu}(l)^{1/2} = (2,2)$ and $\bar{\rho} = (2,1) \vdash r = 3$, we have $\rho = 2\tilde{\rho} \cup \bar{\rho}$. Then it follows that

$$\sum_{\substack{\lambda \vdash \tilde{\mu}(l)^{1/2}=(2,2) \\ 2\lambda=(4,2,2)}} m_\lambda = m_{((2),(1,1))} + m_{((1,1),(2))} = 2,$$

and $\binom{\rho}{l\tilde{\rho}} = \binom{2+1}{2} = 3$. Thus we have

$$\sharp\{\sigma \in S_{11}/S_{(2^4)} \times S_3 | w\sigma a^{-1} \equiv \sigma \bmod S_{(2^4)} \times S_3\} = \binom{3}{2}\left(m_{((2),(1,1))} + m_{((1,1),(2))}\right)2^3 = 48.$$

5. Representation theory of the symmetric group

In this final section, we understand the main result in terms of representation theory of the symmetric group.

It is known that the Green polynomial $Q^\mu_\rho(q)$ gives the graded character value of a certain graded S_n-module, called the *DeConcini-Procesi-Tanisaki algebra* [DP]. The DeConcini-Procesi-Tanisaki algebras R_μ are defined for each partition μ of n, and afford a family of graded representations of S_n. We denote by

$$R_\mu = \bigoplus_{d\geq 0} R^d_\mu$$

its grading. Geometrically, the algebra R_μ is isomorphic to the cohomology ring

$$H^*(X_\mu, \mathbf{C})$$

of the fixed point subvariety X_μ of the flag variety, corresponding to the partition μ. In this point of view, the representation of S_n afforded by R_μ is called the Springer representation [S, L]. As an S_n-module, R_μ is isomorphic to the induced representation $\mathrm{Ind}^{S_n}_{S_\mu}1$.

The *graded character* $\mathrm{char}_q R_\mu$ of the graded module R_μ, evaluated on the conjugacy class corresponding to $\rho \vdash n$, is by definition a polynomial in q

$$\mathrm{char}_q R_\mu(\rho) = \sum_{d\geq 0} q^d \mathrm{char} R^d_\mu(\rho)$$

with integer coefficients. It is known that it coincides with the Green polynomial

$$Q^\mu_\rho(q) = \mathrm{char}_q R_\mu(\rho)$$

for each $\rho \vdash n$.

The aim of this section is to rephrase the recursive formula of the Green polynomials $Q^{\mu}_{\rho}(q)$ in the main theorem, in terms of the graded algebra R_μ. The formula gives a representation theoretical interpretation of a certain combinatorial property of the algebra R_μ. By considering behavior of the Hilbert polynomial

$$\mathrm{Hilb}_\mu(q) = \sum_{d\geq 0} q^d \dim R_\mu^d$$

of the graded module R_μ at roots of unity, we can show that R_μ has the following property. Let M_μ be the maximum multiplicity of μ, and let an integer $l \in \{2, 3, \dots, M_\mu\}$ be fixed. For each $k = 0, 1, \dots, l-1$, define

$$R_\mu(k;l) := \bigoplus_{d\equiv k \bmod l} R_\mu^d.$$

It is clear that these $R_\mu(k;l)$'s are S_n-submodules of R_μ. Then it follows that

Proposition 5.1. *The dimensions of the submodules* $R_\mu(k;l)$ $(k = 0, 1, \dots, l-1)$ *coincides with each other.*

This is a consequence of the fact that the Hilbert polynomial $\mathrm{Hilb}_\mu(q)$ has the roots of unity ζ_l^j for each $j = 1, 2, \dots, l-1$ as its zeros.

Our problem is to give an interpretation to this property "coincidence of dimensions" in terms of representation theory, that is, constructing a subgroup $H(l)$ and its modules $Z(k;l)$ $(k = 0, 1, \dots, l-1)$ of equal dimension such that

$$R_\mu(k;l) \cong_{S_n} \mathrm{Ind}_{H(l)}^{S_n} Z(k;l), \quad k = 0, 1, \dots, l-1.$$

Since the dimension of the induced representation $\mathrm{Ind}_{H(l)}^{S_n} Z(k;l)$ is $\dim Z(k;l)|S_n|/|H(l)|$, we can convince ourselves that these isomorphisms are representation theoretical interpretation of the coincidence of dimensions. Let $\mu \vdash n$ be a partition, $l \in \{2, 3, \dots, M_\mu\}$ fixed, $a = a_\mu(l)$ the cyclic permutation product corresponding to μ and l, and $C_l = \langle a \rangle$ the cyclic subgroup of S_n generated by a. Recall that the subgroup $H_\mu(l)$ is defined by $H_\mu(l) = \left(S_{\tilde{\mu}(l)} \rtimes C_l\right) \times S_r$. Consider, for each $k = 0, 1, \dots, l-1$, $H_\mu(l)$-modules $Z_\mu(k;l)$ defined as follows:

$$Z_\mu(k;l) = \bigoplus_{d=1}^{n(\bar{\mu}(l))} \varphi_l^{(k-d)} \otimes R^d_{\bar{\mu}(l)},$$

where $\varphi_l^{(r)}$ is the irreducible representation of the cyclic group $C_l = \langle a \rangle$ such that $a \longmapsto \zeta_l^r$. The Young subgroup $S_{\tilde{\mu}(l)}$ acts trivially on $Z_\mu(k;l)$. Since $\varphi_l^{(r)}$'s are one dimensional, the dimension of $Z_\mu(k;l)$ is equal to $\dim R_{\bar{\mu}(l)}$ for each k. We shall show that

$$R_\mu(k;l) \cong_{S_n} \mathrm{Ind}_{H_\mu(l)}^{S_n} Z_\mu(k;l), \quad k = 0, 1, \dots, l-1.$$

Actually, we shall show a certain $S_n \times C_l$-module isomorphism between R_μ and $\mathrm{Ind}_{S_{\tilde{\mu}(l)}\times S_r}^{S_n} R_{\bar{\mu}(l)}$, originally suggested by T. Shoji, which is equivalent to those isomorphisms.

We define $S_n \times C_l$-modules structures on R_μ and $\mathrm{Ind}_{S_{\tilde{\mu}(l)}\times S_r}^{S_n} R_{\bar{\mu}(l)}$ as follows. In both cases, the S_n-actions are natural ones. The action of C_l on R_μ is defined by

$$a.x = \zeta_l^d x, \quad x \in R_\mu^d.$$

Recall that the induced modules $\mathrm{Ind}_{S_{\tilde{\mu}(l)}\times S_r}^{S_n} R_{\bar{\mu}(l)}$ has the following realization:

$$\mathrm{Ind}_{S_{\tilde{\mu}(l)}\times S_r}^{S_n} R_{\bar{\mu}(l)} = \bigoplus_{\sigma\in S_n/S_{\tilde{\mu}(l)}\times S_r} \sigma \otimes R_{\bar{\mu}(l)}.$$

Then the C_l-action is defined by

$$a.\sigma \otimes x = \sigma a^{-1} \otimes a.x, \quad \sigma \in S_n/S_{\tilde{\mu}(l)} \times S_r,\ x \in R_{\bar{\mu}(l)}.$$

It is easy to see that the S_n-action and the C_l-action commute on each module. These two $S_n \times C_l$-modules are isomorphic, which is proved by comparing the characters of these modules.

Theorem 5.1. *Let μ be a partition of a positive integer n, and l an integer such that $2 \le l \le M_\mu$ fixed. Suppose that $n = ql + r$, $0 \le r \le l-1$, and let C_l be the cyclic group generated by the element $a = a_\mu(l)$. Then there exists an isomorphism of $S_n \times C_l$-modules*

$$R_\mu \cong \mathrm{Ind}_{S_{\tilde{\mu}(l)} \times S_r}^{S_n} R_{\bar{\mu}(l)}. \tag{5.1}$$

If we consider the eigenspace decomposition of the action of a in the $S_n \times C_l$-isomorphism (5.1), then we obtain a representation theoretical interpretation of the property, coincidence of dimension, of the algebra R_μ.

Proposition 5.2. *Let $\mu \vdash n$ be partition and an integer $l \in \{2, 3, \ldots, M_\mu\}$ fixed. Then there exist $H_\mu(l)$-modules $Z_\mu(k; l)$ $(k = 0, 1, \ldots, l-1)$ of equal dimension such that*

$$R_\mu(k; l) \cong_{S_n} \mathrm{Ind}_{H_\mu(l)}^{S_n} Z_\mu(k; l)$$

for each $k = 0, 1, \ldots, l-1$.

Example 5.2. Let $\mu = (5, 4, 4, 2, 2, 1)$ and $l = 2$. Then $\tilde{\mu}(2) = (4, 4, 2, 2)$, $\bar{\mu}(l) = (5, 1)$, and

$$a = a_\mu(2) = \begin{pmatrix} 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 10 & 11 & 12 & 13 & 6 & 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 14 & 15 & 16 & 17 \\ 16 & 17 & 14 & 15 \end{pmatrix}.$$

The dimensions of $R_\mu(k; 2)$, $k = 0, 1$, equals $\dim R_\mu / 2 = \binom{18}{5,4,4,2,2,1}/2 = 18!/5!4!4!2!2!1!2$. The subgroup $H_\mu(2)$ is defined by $H_\mu(2) = S_{\mu(2)} \rtimes \langle a \rangle \times S_6$, where $S_{\mu(2)} = S_{\{6,7,8,9\}} \times S_{\{10,11,12,13\}} \times S_{\{14,15\}} \times S_{\{16,17\}}$ and $S_r = S_{\{1,2,3,4,5,18\}}$ $(r = 3)$. Define $H_\mu(2)$-modules $Z_\mu(k; l)$ $(k = 0, 1)$ by $Z_\mu(k; 2) := \bigoplus_{d \equiv k \bmod 2} \varphi_2^{(k-d)} \otimes R_{\bar{\mu}(l)}^d$. These spaces are considered as $H_\mu(2)$-modules, where $S_{\mu\tilde{(2)}}$ acts on them trivially. The dimension of these modules are both equal to $\dim R_{\bar{\mu}(l)} = \binom{6}{5,1} = 6!/5!1!$. Then, for each $k = 0, 1$, we have an isomorphism of S_{18}-modules $R_\mu(k; 2) \cong \mathrm{Ind}_{(S_{(4,4,2,2)} \rtimes C_2) \times S_6}^{S_{18}} Z_\mu(k; 2)$. The induced modules are of dimension $18!/4!4!2!2!6!2 \times 6!/5!1! = 18!/5!4!4!2!2!1!2 = \dim R_\mu(k; 2)$ for each $k = 0, 1$.

Remark 5.3. Recently, the author was informed by T. Shoji that the problem considered in this section is given an affirmative answer in a largely generalized setting [**Sh**].

Acknowledgement.

The author would like to express his gratitude to Professor T. Shoji, Nagoya University, for his invaluable comments, which drive strongly the study presented in this article. The author also would like to thank Prof. T. Nakajima, Meikai University, for useful discussions and comments. Finally, the author would like to thank the referees for their comments which correct many errors of the original manuscript.

References

[DP] C. DeConcini and C. Procesi, Symmetric functions, conjugacy classes, and the flag variety, *Inv. Math.* **64** (1981), 203-230.

[DLT] J. Désarménien, B. Leclerc and J.-Y. Thibon, Hall-Littlewood Functions and Kostka-Foulkes Polynomials in Representation Theory, *Seminaire Lotharingien de Combinatoire* **32**, 1994.

[Gr] J. A. Green, The character of the finite general linear groups, *Trans. Amer. Math. Soc.*, **80** (1955), 402-447.

[L] G. Lusztig, Green polynomials and singularities of unipotent classes, *Adv. Math.* **42** (1981), 169-178.

[LLT] A. Lascoux, B. Leclerc and J. -Y. Thibon, Green polynomials and Hall-Littlewood functions at roots of unity, *Euro. J. Comb.* **15** (1994), 173-180.

[M] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, 2nd ed., Oxford University Press, 1995.

[Mt] H. Morita, Decomposition of Green polynomial of type A and DeConcini-Procesi-Tanisaki algebras of certain types, submitted.

[MN1] H. Morita and T. Nakajima, The coinvariant algebra of the symmetric group as a direct sum of induced modules, *Osaka J. Math.* **42** (2005), 217-231.

[MN2] H. Morita and T. Nakajima, A formula of Lascoux-Leclerc-Thibon and DeConcini-Procesi-Tanisaki algebras, to appear in *J. Alg. Combin.*

[S] T. A. Springer, Trigonometric sums, Green functions of finite groups and representations of Weyl groups, *Inv. Math.* **36** (1976), 173-207.

[Sh] T. Shoji, A variant of the induction theorem for Springer representations, preprint 2005.

[T] T. Tanisaki, Defining ideals of the closures of conjugacy classes and representations of the Weyl groups, *Tohoku J. Math.* **34** (1982), 575-585.

School of Science, Tokai University Hiratsuka 259-1292, Japan
E-mail address: `moritah@msg.biglobe.ne.jp`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Algebraic shifting of cyclic polytopes and stacked polytopes

Satoshi Murai

Abstract. Gil Kalai introduced the shifting-theoretic upper bound relation to characterize the f-vectors of Gorenstein* complexes (or homology spheres) by using algebraic shifting. In the present paper, we study the shifting-theoretic upper bound relation. First, we will study the relation between exterior algebraic shifting and combinatorial shifting. Second, by using the relation above, we will prove that the boundary complex of cyclic polytopes satisfies the shifting theoretic upper bound relation. We also prove that the boundary complex of stacked polytopes satisfies the shifting-theoretic upper bound relation.

Résumé. Gil Kalai a défini une relation "shifting-theoretic upper bound" pour caractériser les f-vecteurs des complexes de Gorenstein (sphères d'homologie) en termes de décalages algébriques. Dans cet article, nous étudions cette relation. Premièrement, nous étudions la relation entre le décalage algébrique exterieur et le décalage combinatoire. Ensuite, en utilisant cette relation, nous démontrons que le complexe des frontières des polytopes cycliques satisfait la relation "shifting-theoretic upper bound".

1. Introduction

Let Γ be a simplicial complex on $[n] = \{1, \ldots, n\}$. Thus Γ is a collection of subsets of $[n]$ such that (i) $\{j\} \in \Gamma$ for all $j \in [n]$ and (ii) if $\sigma \subset [n]$ and $\tau \in \Gamma$ with $\sigma \subset \tau$, then $\sigma \in \Gamma$. A *k-face* of Γ is an element $\sigma \in \Gamma$ with $|\sigma| = k+1$. The *k-skeleton* of Γ is a family of $(k+1)$-subset $\Gamma_k = \{\sigma \in \Gamma : |\sigma| = k+1\}$. Let $f_k(\Gamma) = |\Gamma_k|$ the numbers of k-faces of Γ. The vector $f(\Gamma) = (f_0(\Gamma), f_1(\Gamma), \ldots)$ is called the *f-vector* of Γ. If $\sigma = \{s_1, s_2, \ldots, s_r\}$ and $\tau = \{t_1, t_2, \ldots, t_r\}$ are r-subsets of $[n]$ with $s_j < s_{j+1}$ and $t_j < t_{j+1}$ for $j = 1, 2, \ldots, r-1$, write $\sigma \prec_p \tau$ if $s_j \leq t_j$ for all $1 \leq j \leq r$. A simplicial complex Γ is called *shifted* if $\tau \in \Gamma$ and $\sigma \prec_p \tau$ implies $\sigma \in \Gamma$.

The g-theorem gives a complete characterization of the f-vectors of boundary complexes of simplicial polytopes. (see [**10**, pp 75–78].) It has been conjectured that the characterization of g-theorem holds for all Gorenstein* complexes. In the present paper, we call this conjecture the g-conjecture. In [**5**], Kalai introduced the shifting-theoretic upper bound relation to solve the g-conjecture by using algebraic shifting. We recall shifting-theoretic upper bound relation.

Algebraic shifting is an operation which associates with each simplicial complex Γ another shifted simplicial complex $\Delta(\Gamma)$. There are two types of algebraic shifting, i.e., exterior algebraic shifting $\Gamma \to \Delta^e(\Gamma)$ and symmetric algebraic shifting $\Gamma \to \Delta^s(\Gamma)$.

For positive integers $i < j$, we write $[i,j] = \{i, i+1, \ldots, j-1, j\}$ and $[i] = \{1, 2, \ldots, i\}$. A d-subset σ is called *admissible* if $j \notin \sigma$ implies $[j+1, d-j+2] \subset \sigma$. Let $C(n,d)$ be the boundary complex of the cyclic d-polytope with n vertices. Kalai [**4**] proved that $\Delta^s(C(n,d))$ is pure and $\Delta^s(C(n,d))_{d-1}$ consists of

---

2000 *Mathematics Subject Classification.* Primary 13F55 ; Secondary 05E99.

Key words and phrases. exterior shifting, combinatorial shifting, cyclic polytope, stacked polytope.

The author is supported by JSPS Research Fellowships for Young Scientists.

all admissible d-subsets of $[n]$, in other words,

$$\begin{aligned}\Delta^s(C(n,d))_{d-1} &= \{[1,\lfloor\frac{d+1}{2}\rfloor]\cup\sigma : \sigma\subset[\lfloor\frac{d+1}{2}\rfloor+1,n],\ |\sigma| = d-\lfloor\frac{d+1}{2}\rfloor\}\\ &\bigcup_{1\le j\le\lfloor\frac{d+1}{2}\rfloor}\{([1,d-j+2]\setminus\{j\})\cup\sigma : \sigma\subset[d-j+3,n],\ |\sigma| = j-1\},\end{aligned}$$

where $\lfloor\frac{d+1}{2}\rfloor$ means the integer part of $\frac{d+1}{2}$. Furthermore, Kalai proved that the boundary complex P of every simplicial d-polytope with n vertices satisfies $\Delta^s(P)\subset\Delta^s(C(n,d))$ by using the Lefschetz property of P (see §1.2 for the Lefschetz property). Furthermore, Kalai noticed that if Γ is a $(d-1)$-dimensional Gorenstein* complex Γ on $[n]$ then the relation $\Delta^s(\Gamma)\subset\Delta^s(C(n,d))$ is equivalent to the Lefschetz property of Γ.

It is not hard to see that if every $(d-1)$-dimensional Gorenstein* complex Γ on $[n]$ satisfies $\Delta^e(\Gamma)\subset\Delta^s(C(n,d))$ then the g-conjecture is true. We say that a $(d-1)$-dimensional complex Γ on $[n]$ satisfies the *shifting-theoretic upper bound relation* if Γ satisfies $\Delta^e(\Gamma)\subset\Delta^s(C(n,d))$. Kalai and Sarkaria conjectured that if Γ is a simplicial complex on $[n]$ whose geometric realization can be embedded in S^{d-1} then Γ satisfies the shifting-theoretic upper bound relation. However, it is not known whether Γ satisfies the shifting-theoretic upper bound relation even if *Gamma* is the boundary complex of a simplicial polytope. In the present paper, we will show that $C(n,d)$ and the boundary complex of stacked polytopes satisfies the shifting-theoretic upper bound relation.

In general, the computation of exterior algebraic shifting is rather difficult. First, we will show that we can use combinatorial shifting to study shifting theoretic upper bound relation. Combinatorial shifting, which was introduced by Erdös, Ko and Rado [**3**], is also an operation which associates with each simplicial complex Γ another shifted simplicial complex $\Delta^c(\Gamma)$. Although combinatorial shifting may not be uniquely determined, it is easily computed by a simple combinatorial method. Regarding the relation between exterior algebraic shifting and combinatorial shifting, we have the following result.

Theorem 1.6. *Let Γ be a $(d-1)$-dimensional Cohen-Macaulay complex on $[n]$ with $h_d(\Gamma)\neq 0$ and with $h_i(\Gamma)=h_{d-i}(\Gamma)$ for $i=0,1,\dots,\lfloor\frac{d}{2}\rfloor$.*

(i) *If $\Delta^c(\Gamma)_{d-1}\subset\Delta^s(C(n,d))_{d-1}$, then this $\Delta^c(\Gamma)$ is pure.*
(ii) *If there is a combinatorial shifted complex $\Delta^c(\Gamma)$ of Γ with $\Delta^c(\Gamma)_{d-1}\subset\Delta^s(C(n,d))_{d-1}$, then one has $\Delta^e(\Gamma)\subset\Delta^s(C(n,d))$.*

Thus, we can use combinatorial shifting for the shifting-theoretic upper bound relation. Also, since combinatorial shifting is entirely a combinatorial operation, proving $\Delta^c(P)\subset\Delta^s(C(n,d))$ for the boundary complex P of a simplicial d-polytope without using the Lefschetz property would be interesting. By using Theorem 1.6, we compute the exterior algebraic shifted complex of the boundary complex of the cyclic d-polytope.

Theorem 2.1. *Let $C(n,d)$ be the boundary complex of the cyclic d-polytope with n vertices. Then there is a combinatorial shifted complex $\Delta^c(C(n,d))$ such that $\Delta^c(C(n,d))=\Delta^s(C(n,d))$. Thus, in particular, one has $\Delta^e(C(n,d))=\Delta^s(C(n,d))$.*

We also compute algebraic shifting of the boundary complex of a stacked d-polytope with n vertices.

Theorem 2.2. *Let $L(n,d)$ be the pure $(d-1)$-dimensional simplicial complex spanned by*

$$\{\{2,\dots,d+1\}\}\cup\{(\{1,\dots,d\}\setminus\{i\})\cup\{j\} : 1<i\le d,\ j>d \text{ or } j=i\}.$$

Let $P(n,d)$ be the boundary complex of a stacked d-polytope with n vertices. Then

(i) *One has $\Delta^e(P(n,d))=\Delta^s(P(n,d))=L(n,d)$.*
(ii) *If Γ is the boundary complex of a simplicial d-polytope with n vertices, then one has*

$$\Delta^s(P(n,d))\subset\Delta^s(\Gamma).$$

Note that $\Delta^s(P(n,d))=L(n,d)$ and (ii) easily follows from the relation $\Delta^s(P(n,d))\subset\Delta^s(C(n,d))$. To prove $\Delta^e(P(n,d))=L(n,d)$, we use the fact that the 1-skeleton of $P(n,d)$ is a chordal graph. However, we are not sure that $L(n,d)$ can be obtained by applying combinatorial shifting to $P(n,d)$, the boundary complex of an arbitrary stacked d-polytopes with n vertices.

1.1. algebraic shifting and combinatorial shifting. To define algebraic shifting, we need the theory of generic initial ideals in the exterior algebra.

Let K be an infinite field, V a vector space over K of dimension n with basis $e_1, \dots, e_n$ and $E = \bigoplus_{d=0}^n \bigwedge^d(V)$ the exterior algebra of V. In other words, E is a K-algebra which satisfies

(i) Each $\bigwedge^d(V)$ is a $\binom{n}{d}$ dimensional K-vector space with the canonical K-basis $\{e_{s_1} \wedge e_{s_2} \wedge \cdots \wedge e_{s_d} : 1 \leq s_1 < s_2 < \cdots < s_d \leq n\}$.

(ii) For any integers $i, j \in [n]$, one has $e_i \wedge e_j = -e_j \wedge e_i$.

For $\sigma = \{s_1, \dots, s_d\} \subset [n]$ with $s_1 < \cdots < s_d$, we call $e_\sigma = e_{s_1} \wedge \cdots \wedge e_{s_d} \in \bigwedge^d(V)$ a *monomial* of E of degree d. Fix a term order $<$. For every homogeneous element $f = \sum_{|\sigma|=d} \alpha_\sigma e_\sigma \in \bigwedge^d(V)$ with each $\alpha_\sigma \in K$, the monomial $\mathrm{in}_<(f) = \max_<\{e_\sigma : \alpha_\sigma \neq 0\}$ is called the *initial monomial of* f. Also, for every homogeneous ideal $J \subset E$, The *initial ideal of* J is the monomial ideal generated by $\{\mathrm{in}_<(f) : f \in J\}$. A monomial ideal $J \subset E$ is called *strongly stable* if $e_\tau \in J$ and $\tau \prec_p \sigma$ means $e_\sigma \in J$.

Let $GL_n(K)$ denote the general linear group with coefficients in K. Any $\varphi = (a_{ij}) \in GL_n(K)$ induce an automorphism of graded K-algebra E as follows:

$$\varphi(f(e_1, \dots, e_n)) = f(\sum_{i=1}^n a_{i1} e_i, \dots, \sum_{i=1}^n a_{in} e_i) \text{ for all } f \in E.$$

If $J \subset E$ is a homogeneous ideal, then each $\varphi \in GL_n(K)$ gives another homogeneous ideal $\varphi(J) = \{\varphi(f) : f \in J\}$. Now, we recall the fundamental theorem of generic initial ideals.

Lemma 1.1 ([**1**, Theorem 1.6]). *Let K be an infinite field. Fix a term order $<$ with $e_1 < \cdots < e_n$. Then, for each homogeneous ideal $J \subset E$, there exists a nonempty Zariski open subset $U \subset GL_n(K)$ such that* $\mathrm{in}_<(\varphi(J)) = \mathrm{in}_<(\varphi'(J))$ *for all $\varphi, \varphi' \in U$ and this* $\mathrm{in}_<(\varphi(J))$ *is strongly stable.*

This monomial ideal $\mathrm{in}_<(\varphi(J))$ is called the *generic initial ideal of $J \subset E$ with respect to the term order* $<$ and will be denoted $\mathrm{Gin}_<(J)$. In particular, we write $\mathrm{Gin}(J) = \mathrm{Gin}_{<_{\mathrm{rev}}}(J)$, where $<_{\mathrm{rev}}$ is the degree reverse lexicographic order with $e_1 < e_2 < \cdots < e_n$. In other words, for $\sigma \subset [n]$ and $\tau \subset [n]$ with $\sigma \neq \tau$, define $e_\sigma <_{rev} e_\tau$ if (i) $|\sigma| < |\tau|$ or (ii) $|\sigma| = |\tau|$ and the minimal integer in symmetric difference $(\sigma \setminus \tau) \cup (\tau \setminus \sigma)$ belongs to σ. Also, we define $\sigma <_{rev} \tau$ by the same way.

A *shifting operation* on $[n]$ is an operator which associates with each simplicial complex Γ on $[n]$ a simplicial complex $\Delta(\Gamma)$ on $[n]$ and which satisfies the following conditions:

(S_1) $\Delta(\Gamma)$ is shifted;
(S_2) $\Delta(\Gamma) = \Gamma$ if Γ is shifted;
(S_3) $f(\Gamma) = f(\Delta(\Gamma))$;
(S_4) $\Delta(\Gamma') \subset \Delta(\Gamma)$ if $\Gamma' \subset \Gamma$.

(Exterior algebraic shifting) Let Γ be a simplicial complex on $[n]$. The *exterior face ideal of* Γ is a monomial ideal of E generated by all monomials $e_\sigma \in E$ with $\sigma \notin \Gamma$. The *exterior algebraic shifted complex* of Γ is the simplicial complex $\Delta^e(\Gamma)$ defined by

$$J_{\Delta^e(\Gamma)} = \mathrm{Gin}(J_\Gamma).$$

The shifting operation $\Gamma \mapsto \Delta^e(\Gamma)$ which is in fact a shifting operation ([**6**, Proposition 8.8]), is called *exterior algebraic shifting.*

(Combinatorial shifting) Erdös, Ko and Rado [**3**] introduced combinatorial shifting. Let Γ be a collection of r-subsets of $[n]$, where $r \leq n$. For $1 \leq i < j \leq n$, write $\mathrm{Shift}_{ij}(\Gamma)$ for the collection of r-subsets of $[n]$ whose elements are $C_{ij}(\sigma) \subset [n]$, where $\sigma \in \Gamma$ and where

$$C_{ij}(\sigma) = \begin{cases} (\sigma \setminus \{j\}) \bigcup \{i\}, & \text{if } j \in \sigma,\ i \notin \sigma \text{ and } (\sigma \setminus \{j\}) \bigcup \{i\} \notin \Gamma, \\ \sigma, & \text{otherwise.} \end{cases}$$

We can define $\mathrm{Shift}_{ij}(\Gamma)$ for a simplicial complex Γ by the same way. It follows from, e.g., [**6**, Corollary 8.6] that there exists a finite sequence of pairs of integers $(i_1, j_1), (i_2, j_2), \dots, (i_q, j_q)$ with each $1 \leq i_k < j_k \leq n$ such that

$$\mathrm{Shift}_{i_q j_q}(\mathrm{Shift}_{i_{q-1} j_{q-1}}(\cdots(\mathrm{Shift}_{i_1 j_1}(\Gamma))\cdots))$$

is shifted. Such a shifted complex is called a *combinatorial shifted complex* of Γ and will be denoted by $\Delta^c(\Gamma)$. A combinatorial shifted complex $\Delta^c(\Gamma)$ of Γ is, however, not necessarily unique. The shifting operation $\Gamma \mapsto \Delta^c(\Gamma)$, which is in fact a shifting operation ([**6**, Lemma 8.4]), is called *combinatorial shifting.*

Algebraic shifting behaves nicely. For example, algebraic shifting preserves the Cohen-Macaulay property and preserves the dimension of reduced homology groups. On the other hand, combinatorial shifting does not behave nicely. However, the advantage of combinatorial shifting is that we can easily compute them by purely combinatorial methods. Hence the following problem naturally occurs.

Problem (Kalai [**5**, Problem 24]). What are the relations between combinatorial shifting and algebraic shifting?

We will remark a relation between combinatorial shifting and exterior algebraic shifting. For every $\sigma \subset [n]$ and for every shifted simplicial complex Γ on $[n]$, define

$$m_{\leq\sigma}(\Gamma) = |\{\tau \in \Gamma : \tau \leq_{rev} \sigma \text{ and } |\tau| = |\sigma|\}|.$$

Then we have the following relation between $\Delta^c(\Gamma)$ and $\Delta^e(\Gamma)$.

Lemma 1.2. *Let Γ be a simplicial complex on $[n]$. Then, for any combinatorial shifted complex $\Delta^c(\Gamma)$ and for any subset $\sigma \subset [n]$, one has*

$$m_{\leq\sigma}(\Delta^e(\Gamma)) \geq m_{\leq\sigma}(\Delta^c(\Gamma)).$$

Proof. It is not hard to see that (see [**6**, Lemma 8.3]), for all integers $1 \leq i < j \leq n$, there is $\varphi_{ij} \in GL_n(K)$ such that $\mathrm{in}_{<_{rev}}(\varphi_{ij}(I_\Gamma)) = I_{\mathrm{Shift}_{ij}(\Gamma)}$. For each $\varphi \in GL_n(K)$, define a simplicial complex $\Delta_\varphi(\Gamma)$ by

$$I_{\Delta_\varphi(\Gamma)} = \mathrm{in}_{<_{rev}}(\varphi(I_\Gamma)).$$

Then, it follows from [**8**, Theorem 3.1] that, for every $\sigma \subset [n]$, one has

$$m_{\leq\sigma}(\Delta^e(\Gamma)) \geq m_{\leq\sigma}(\Delta^e(\Delta_\varphi(\Gamma))). \tag{1.1}$$

By the definition of combinatorial shifting, there exists a finite sequence of pairs of integers $(i_1, j_1), (i_2, j_2), \ldots, (i_q, j_q)$ such that $\Delta^c(\Gamma) = \Delta_{\varphi_{i_q j_q}}(\Delta_{\varphi_{i_{q-1} j_{q-1}}}(\cdots(\Delta_{\varphi_{i_1 j_1}}(\Gamma))\cdots))$. Also, since $\Delta^c(\Gamma)$ is shifted, the conditions of shifting operation say $\Delta^e(\Delta^c(\Gamma)) = \Delta^c(\Gamma)$. Then, by (1.1), we have

$$\begin{aligned} m_{\leq\sigma}(\Delta^e(\Gamma)) &\geq m_{\leq\sigma}(\Delta^e(\Delta_{\varphi_{i_q j_q}}(\Delta_{\varphi_{i_{q-1} j_{q-1}}}(\cdots(\Delta_{\varphi_{i_1 j_1}}(\Gamma))\cdots)))) \\ &= m_{\leq\sigma}(\Delta^e(\Delta^c(\Gamma))) \\ &= m_{\leq\sigma}(\Delta^c(\Gamma)), \end{aligned}$$

for every $\sigma \subset [n]$, as desired. $\square$

Note that Lemma 1.2 induces some other relations between combinatorial shifting and exterior algebraic shifting. For example, it was used in [**9**] to compare the graded Betti numbers of the Stanley-Reisner ideal of $\Delta^e(\Gamma)$ and $\Delta^c(\Gamma)$.

1.2. The shifting-theoretic upper bound relation. The shifting-theoretic upper bound relation was considered from the viewpoint of symmetric algebraic shifting. Thus, first, we recall symmetric algebraic shifting which was introduced in [**4**]. We refer the reader to [**10**] for the definition of Cohen-Macaulay complexes and Gorenstein* complexes.

(Symmetric algebraic shifting) Let K be a field of characteristic 0 and $R = K[x_1, \ldots, x_n]$ the polynomial ring. Let Γ be a simplicial complex on $[n]$. The *Stanley-Reisner ideal* I_Γ *of* Γ is a monomial ideal generated by all squarefree monomials $x_{i_1} x_{i_2} \cdots x_{i_r}$ with $\{i_1, i_2, \ldots, i_r\} \notin \Gamma$ and $\{i_1, i_2, \ldots, i_r\} \subset [n]$. The ring $R(\Gamma) = R/I_\Gamma$ is called *the face ring of* Γ.

Let $y_1, y_2, \ldots, y_n$ be generic linear forms in $x_1, x_2, \ldots, x_n$ and M the set of monomials in $y_1, y_2, \ldots, y_n$. For every monomial m in M, denote its image in $R(\Gamma)$ by $\tilde{m}$. Define

$$\mathrm{GIN}(\Gamma) = \{m \in M : \tilde{m} \notin \mathrm{span}\{\tilde{l} : \deg(l) = \deg(m),\ l <_{rev} m\}\}.$$

For every monomial $m \in \mathrm{GIN}(\Gamma)$ with $\deg(m) = r \leq n$ which does not involve $y_1, y_2, \dots, y_{r-1}$, write $m = y_{i_1} y_{i_2} \cdots y_{i_r}$ with $i_1 \leq i_2 \leq \cdots \leq i_r$, and define

$$S(m) = \{i_1 - r + 1, i_2 - r + 2, \dots, i_{r-1} - 1, i_r\}.$$

The *symmetric algebraic shifted complex* $\Delta^s(\Gamma)$ of Γ is defined by

$$\Delta^s(\Gamma) = \{S(m) : m \in \mathrm{GIN}(\Gamma),\ \deg(m) = r \leq n \text{ and } y_i \text{ does not divides } m \text{ for } i \leq r-1\}.$$

The shifting operation $\Gamma \to \Delta^s(\Gamma)$ which is in fact a shifting operation ([**6**, §8]), is called *symmetric algebraic shifting.*

Second, we recall h-vectors. Let Γ be a $(d-1)$-dimensional simplicial complex and $(f_0, f_1, \dots, f_{d-1})$ f-vectors of Γ. The *h-vector* of Γ is defined by the relation

$$\sum_{i=0}^{d} h_i(\Gamma) x^{d-i} = \sum_{i=0}^{d} f_{i-1}(x-1)^{d-i},$$

where we let $f_{-1} = 1$. This is equivalent to

$$h_i(\Gamma) = \sum_{j=0}^{i} (-1)^{i-j} \binom{d-j}{d-i} f_{j-1} \text{ and } f_{i-1} = \sum_{j=0}^{i} \binom{d-j}{d-i} h_i(\Gamma).$$

(The Lefschetz property) Let Γ be a $(d-1)$-dimensional Cohen-Macaulay simplicial complex and $\vartheta_1, \vartheta_2, \dots, \vartheta_d$ generic linear forms. Then $\vartheta_1, \vartheta_2, \dots, \vartheta_d$ is a system of parameters of $R(\Gamma)$. Let

$$\bigoplus_{i=0}^{d} H_i(\Gamma) = R(\Gamma)/ < \vartheta_1, \vartheta_2, \dots, \vartheta_d >,$$

where $H_i(\Gamma)$ is the i-th homogeneous component of $R(\Gamma)/ < \vartheta_1, \vartheta_2, \dots, \vartheta_d >$. It is well known [**10**, pp. 53–58] that

$$h_i(\Gamma) = \dim_K H_i(\Gamma).$$

Let ϑ_{d+1} be an additional general linear form and $s = \max\{k : h_k(\Gamma) \neq 0\}$. A $(d-1)$-dimensional Cohen-Macaulay simplicial complex Γ is called *(strongly) Lefschetz* if, for $0 \leq i \leq \lfloor \frac{s}{2} \rfloor$, the multiplication

$$\vartheta_{d+1}^{s-2i} : H_i(\Gamma) \to H_{s-i}(\Gamma)$$

is an isomorphism. Note that the boundary complex of every simplicial polytope is Lefschetz. The important aspect of Lefschetz property is that proving the Lefschetz property for all Gorenstein* complexes implies the g-conjecture. See [**10**, pp75–78] for the detail.

Next, we recall some basic property of algebraic shifting.

LEMMA 1.3 ([**6**, Lemma 8.]). *Let Γ be a simplicial complex. The followings are equivalent:*

(i) Γ *is Cohen-Macaulay;*
(ii) $\Delta^e(\Gamma)$ *is Cohen-Macaulay;*
(iii) $\Delta^e(\Gamma)$ *is pure.*

Lemma 1.3 is also true for symmetric algebraic shifting Δ^s. Also, if Γ is Cohen-Macaulay, then h-vectors of Γ appears in $\Delta^e(\Gamma)$ and $\Delta^s(\Gamma)$ by the following way.

LEMMA 1.4 (Kalai [**4**, Lemma 7.1]). *Let Γ be a pure shifted $(d-1)$-dimensional simplicial complex. Let $W_i(\Gamma) = \{\sigma \in \Gamma : |\sigma| = d,\ [d-i] \subset \sigma$ and $d-i+1 \notin \sigma\}$. Then $h_i(\Gamma) = |W_i(\Gamma)|$.*

PROOF. For every monomial $m \in K[y_1, \dots, y_n]$ with $\deg(u) = i$, denote its image in $R(\Gamma)/ < y_1, y_2, \dots, y_d >$ by $[m]$. Let

$$L_i(\Gamma) = \{m \in \mathrm{GIN}(\Gamma) : \deg(m) = i \text{ and } m \in K[y_{d+1}, \dots, y_n]\}.$$

First, we will show that $\dim_K H_i(\Gamma) = |L_i(\Gamma)|$. If m is a monomial in $K[y_{d+1}, \dots, y_n]$ with $deg(m) = i$ and l is a monomial in $< y_1, \dots, y_d >$ with $deg(l) = i$, then $l <_{rev} m$. Since $\mathrm{GIN}(\Gamma) = \{m \in M : \tilde{m} \notin \mathrm{span}\{\tilde{l} : l <_{rev} m\}\}$, it follows that the set of monomials $\tilde{m}$ with $m \in \mathrm{GIN}(\Gamma)_i \cap < y_1, \dots, y_d > = \{\mathrm{GIN}(\Gamma)_i \setminus L_i(\Gamma)\}$ is a K-basis of $\{R(\Gamma) \cap < y_1, \dots, y_d >\}_i$. Thus $\{[m] : m \in L_i(\Gamma)\}$ is a K-basis of $\{R(\Gamma)/ < y_1, \dots, y_d >\}_i$.

On the other hand, since $y_1, y_2, \ldots, y_n$ are generic linear forms, it follows that $y_1, \ldots, y_d$ are generic system of parameters. Thus $H_i(\Gamma) = \{R(\Gamma)/ < y_1, \ldots, y_d >\}_i$ and, therefore, $\dim_K H_i(\Gamma) = |L_i(\Gamma)|$.

Second, we will show that if $\Delta^s(\Gamma)$ is pure and shifted, then, for all $0 \leq i \leq d$, we have

$$W_i(\Delta^s(\Gamma)) = \{[d-i] \cup S(m) : m \in L_i(\Gamma)\}. \tag{1.2}$$

For any $m \in L_i(\Gamma)$, we have $\min(S(m)) \geq d-i+2$ and $|S(m)| = i$. Since $\Delta^s(\Gamma)$ is pure and shifted, we have $[d-i] \cup S(m) \in W_i(\Delta^s(\Gamma))$. Conversely, if $[d-i] \cup \sigma \in W_i(\Delta^s(\Gamma))$, then $\sigma \in \Delta^s(\Gamma)$ and $\min(\sigma) \geq d-i+2$. Hence there is $m \in L_i(\Gamma)$ with $S(m) = \sigma$.

Since Γ is shifted, we have $\Delta^s(\Gamma) = \Gamma$. Then the relation (1.2) says $|W_i(\Gamma)| = |L_i(\Gamma)| = \dim_K H_i(\Gamma) = h_i(\Gamma)$. □

Lemma 1.5 (Kalai). *Let Γ be a $(d-1)$-dimensional Cohen-Macaulay simplicial complex on $[n]$ with $h_d(\Gamma) \neq 0$. The followings are equivalent:*

(i) *Γ is Lefschetz;*
(ii) *$\Delta^s(\Gamma) \subset \Delta^s(C(n,d))$ and $h_i(\Gamma) = h_{d-i}(\Gamma)$ for all $0 \leq i \leq d$.*

Proof. ((i) $\Rightarrow$ (ii)) The relation $h_i(\Gamma) = h_{d-i}(\Gamma)$ immediately follows from the definition of Lefschetz property. Note that $\Delta^s(\Gamma)_{d-1} = \bigcup_{j=0}^{d} W_i(\Delta^s(\Gamma))$. We will show $W_i(\Delta^s(\Gamma)) \subset W_i(\Delta^s(C(n,d)))$ for all $0 \leq i \leq d$. For $0 \leq i \leq \frac{d}{2}$, the inclusion $W_i(\Sigma) \subset W_i(\Delta^s(C(n,d)))$ is true for an arbitrary simplicial complex Σ. Since $y_1, y_2, \ldots, y_n$ are generic linear forms, it follows that $y_1, \ldots, y_d$ are generic system of parameters and y_{d+1} is an additional generic linear form. Then, by assumption, the multiplication $y_{d+1}^{d-2i} : L_i(\Gamma) \to L_{d-i}(\Gamma)$ is a bijection. Then, for $0 \leq i \leq \frac{d}{2}$, $L_{d-i}(\Gamma)$ is of the form $L_{d-i}(\Gamma) = \{y_{d+1}^{d-2i} m : m \in L_i(\Gamma)\}$. Also, for every $m \in L_i(\Gamma)$ with $0 \leq i \leq \frac{d}{2}$, we have

$$S(y_{d+1}^{d-2i} m) = \{i+2, \ldots, d-i+1\} \cup S(m) \tag{1.3}$$

Thus, for $0 \leq i \leq \frac{d}{2}$, relation (1.2) says that $W_{d-i}(\Delta^s(\Gamma))$ is of the form

$$W_{d-i}(\Delta^s(\Gamma)) = \{[i] \cup \{i+2, \ldots, d-i+1\} \cup S(m) : m \in L_i(\Gamma)\} \subset W_{d-i}(\Delta^s(C(n,d))).$$

((ii) $\Rightarrow$ (i)) If $\Delta^s(\Gamma) \subset \Delta^s(C(n,d))$, then, for $0 \leq i \leq \frac{d}{2}$, each $W_{d-i}(\Delta^s(\Gamma))$ is of the form

$$W_{d-i}(\Gamma) = \{[i] \cup \{i+2, \ldots, d-i+1\} \cup \sigma \in \Gamma : |\sigma| = i\}.$$

Since $\Delta^s(\Gamma)$ is shifted, there is a natural injection form $W_{d-i}(\Gamma)$ to $W_i(\Gamma)$ as follows:

$$[i] \cup \{i+2, \ldots, d-i+1\} \cup \sigma \mapsto [d-i] \cup \sigma. \tag{1.4}$$

Since $h_i(\Gamma) = h_{d-i}(\Gamma)$, Lemma 1.4 says this injection is a bijection. Then (1.2) and (1.3) implies that the multiplication $y_{d+1}^{d-2i} : L_i(\Gamma) \to L_{d-i}(\Gamma)$ is a bijection. □

Let Γ be a $(d-1)$-dimensional Gorenstein* complex on $[n]$. Since $\Delta^s(\Delta^e(\Gamma)) = \Delta^e(\Gamma)$, Lemma 1.5 says that $\Delta^e(\Gamma)$ is Lefschetz if and only if $\Delta^e(\Gamma) \subset \Delta^s(C(n,d))$ and $h_i(\Gamma) = h_{d-i}(\Gamma)$ for $i = 0, 1, \ldots, \lfloor \frac{d}{2} \rfloor$. Since $h_i(\Gamma) = h_{d-i}(\Gamma)$, where $0 \leq i \leq \lfloor \frac{d}{2} \rfloor$, are true for arbitrary Gorenstein* complex, if we can prove the relation $\Delta^e(\Gamma) \subset \Delta^s(C(n,d))$ for arbitrary $(d-1)$-dimensional Gorenstein* complex Γ on $[n]$, then we can prove the g-conjecture. However, the relation $\Delta^e(\Gamma) \subset \Delta^s(C(n,d))$ is unknown even for the boundary complex of simplicial polytopes. We say that a $(d-1)$-dimensional simplicial complex Γ on $[n]$ satisfies the shifting-theoretic upper bound relation if Γ satisfies $\Delta^e(\Gamma) \subset \Delta^e(C(n,d))$.

We will show that if $\Delta^c(\Gamma)$ is Lefschetz, then $\Delta^e(\Gamma)$ is also Lefschetz by using Lemma 1.2.

Theorem 1.6. *Let Γ be a $(d-1)$-dimensional Cohen-Macaulay complex on $[n]$ with $h_d(\Gamma) \neq 0$ and with $h_i(\Gamma) = h_{d-i}(\Gamma)$ for $i = 0, 1, \ldots, \lfloor \frac{d}{2} \rfloor$.*

(i) *If $\Delta^c(\Gamma)_{d-1} \subset \Delta^s(C(n,d))_{d-1}$, then this $\Delta^c(\Gamma)$ is pure.*
(ii) *If there is a combinatorial shifted complex $\Delta^c(\Gamma)$ of Γ with $\Delta^c(\Gamma)_{d-1} \subset \Delta^s(C(n,d))_{d-1}$, then one has $\Delta^e(\Gamma) \subset \Delta^s(C(n,d))$.*

Proof. (i) Fix a combinatorial shifted complex $\Delta^c(\Gamma)$ which satisfies the assumption $\Delta^c(\Gamma)_{d-1} \subset \Delta^s(C(n,d))_{d-1}$. We will show $|W_i(\Delta^c(\Gamma))| = h_i(\Gamma)$ for all $0 \leq i \leq d$.

Let $\sigma(i,n) = [d-i] \cup \{n-i+1,\dots,n\}$. Then, for every $\sigma \subset [n]$ with $|\sigma| = d$, we have $\sigma \leq_{rev} \sigma(i,n)$ if and only if $[d-i] \subset \sigma$. This implies that, for every $(d-1)$-dimensional pure shifted simplicial complex Σ, we have

$$m_{\leq \sigma(i,n)}(\Sigma) = \sum_{j=0}^{i} |W_i(\Sigma)|.$$

Then Lemma 1.2 says that , for $0 \leq i \leq \frac{d}{2}$, we have

$$\sum_{j=0}^{i} |W_j(\Delta^e(\Gamma))| \geq \sum_{j=0}^{i} |W_j(\Delta^c(\Gamma))| \text{ and } \sum_{j=0}^{i} |W_{d-j}(\Delta^e(\Gamma))| \leq \sum_{j=0}^{i} |W_{d-j}(\Delta^c(\Gamma))|. \tag{1.5}$$

On the other hand, since $\Delta^c(\Gamma)_{d-1} \subset \Delta^s(C(n,d))_{d-1}$, the injection (1.4) says $|W_i(\Delta^c(\Gamma))| \geq |W_{d-i}(\Delta^c(\Gamma))|$ for $0 \leq i \leq \frac{d}{2}$. In particular, we have

$$\sum_{j=0}^{i} |W_j(\Delta^c(\Gamma))| \geq \sum_{j=0}^{i} |W_{d-j}(\Delta^c(\Gamma))|. \tag{1.6}$$

Since Γ is Cohen-Macaulay and $h_i(\Gamma) = h_{d-i}(\Gamma)$, Lemmas 1.3 and 1.4 say $|W_i(\Delta^e(\Gamma))| = |W_{d-i}(\Delta^e(\Gamma))| = h_i(\Gamma)$. Thus these inequalities (1.5) and (1.6) are all equal. Inductively, we have $|W_i(\Delta^c(\Gamma))| = |W_i(\Delta^e(\Gamma))| = h_i(\Gamma)$ for all $0 \leq i \leq d$.

Let L be the pure simplicial complex generated by $\Delta^c(\Gamma)_{d-1}$. Then Lemma 1.4 says L and Γ have the same h-vector, that is, they have the same f-vector. Since $\Delta^c(\Gamma) \supset L$, we have $\Delta^c(\Gamma) = L$. Thus this $\Delta^c(\Gamma)$ is pure.

(ii) We will show $W_i(\Delta^e(\Gamma)) \subset W_i(\Delta^s(C(n,d)))$ for all $0 \leq i \leq d$. Let $\sigma_0(i) = \max_{<_{rev}}\{W_i(\Delta^s(C(n,d)))\}$, $\sigma_c(i) = \max_{<_{rev}}\{W_i(\Delta^c(\Gamma))\}$ and $\sigma_e(i) = \max_{<_{rev}}\{W_i(\Delta^e(\Gamma))\}$.

Since $\Delta^c(\Gamma) \subset \Delta^s(C(n,d))$, we have $\sigma_0(i) \geq_{rev} \sigma_c(i)$ for all i. On the other hand, since $|W_i(\Delta^c(\Gamma))| = h_i$, we have

$$m_{\leq \sigma_c(i)}(\Delta^c(\Gamma)) = \sum_{k=0}^{i} |W_k(\Delta^c(\Gamma))| = \sum_{k=0}^{i} h_k(\Gamma)$$

and

$$m_{\leq \sigma_e(i)}(\Delta^e(\Gamma)) = \sum_{k=0}^{i} |W_k(\Delta^e(\Gamma))| = \sum_{k=0}^{i} h_k(\Gamma).$$

Then Lemma 1.2 says $\sigma_c(i) \geq_{rev} \sigma_e(i)$. Thus we have $\sigma_0(i) \geq_{rev} \sigma_e(i)$ for all i.

On the other hand, $W_i(\Delta^s(C(n,d)))$ is the set of smallest $h_i(\Gamma)$ elements w.r.t. $<_{rev}$ which contain $\{1,\dots,d-i\}$ and which do not contain $\{d-i+1\}$, that is,

$$W_i(\Delta^s(C(n,d))) = \{\sigma \subset [n] : [d-i] \subset \sigma,\ d-i+1 \notin \sigma \text{ and } \sigma \leq_{rev} \sigma_0(i)\}.$$

Thus we have $\Delta^e(\Gamma) \subset \Delta^s(C(n,d))$. □

2. Exterior algebraic shifting of Cyclic polytopes and stacked polytopes

2.1. Cyclic polytopes. We recall the definition of cyclic polytopes. We refer the reader to [**2**] for the basic theory of convex polytopes.

Let $\mathbb{R}$ denote the set of real numbers. For any subset M of the d-dimensional Euclidean space $\mathbb{R}^d$, there is a smallest convex set containing M. This convex set is called *convex hull of* M and will be denoted by $\mathrm{conv}(M)$. For $d \geq 2$, the *moment curve* in $\mathbb{R}^d$ is the curve parameterized by

$$t \to x(t) = (t, t^2, \dots, t^d) \in \mathbb{R}^d.$$

The *cyclic d-polytope* with n vertices is the convex hull P of the form

$$P = \mathrm{conv}(\{x(t_1), x(t_2), \dots, x(t_n)\}),$$

where $t_1, t_2, \dots, t_n$ are distinct real numbers.

The main result of this section is the following.

Theorem 2.1. *Let $C(n,d)$ be the boundary complex of the cyclic d-polytope with n vertices. Then there is a combinatorial shifted complex $\Delta^c(C(n,d))$ such that $\Delta^c(C(n,d)) = \Delta^s(C(n,d))$. Thus, in particular, one has $\Delta^e(C(n,d)) = \Delta^s(C(n,d))$.*

Proof. (sketch) By virtue of Theorem 1.6, what we have to do is finding a combinatorial shifted complex $\Delta^c(C(n,d))$ which satisfies $\Delta^c(C(n,d)) = \Delta^s(C(n,d))$.

Also, by Gale's evenness condition ([**2**, Theorem 13.6]), we know that $C(n,d)_{d-1}$ is the collection of d-subsets σ of $[n]$ which satisfies, for every $i<j$ with $i,j \notin \sigma$, the number $|\{i,i+1,\dots,j\}\cap\sigma|$ is even.

Define

$$\mathrm{Shift}_{n\downarrow i}(\Gamma) = \mathrm{Shift}_{ii+1}(\cdots(\mathrm{Shift}_{in-1}(\mathrm{Shift}_{in}(\Gamma)))\cdots)$$

and

$$\mathrm{Shift}_{n\uparrow i}(\Gamma) = \mathrm{Shift}_{in}(\cdots(\mathrm{Shift}_{ii+2}(\mathrm{Shift}_{ii+1}(\Gamma)))\cdots).$$

(i) In case of d is even, then

$$\mathrm{Shift}_{n-1\downarrow n}(\mathrm{Shift}_{n-2\downarrow n}(\cdots(\mathrm{Shift}_{1\downarrow n}(C(n,d))\cdots)) = \Delta^s(C(n,d)).$$

(ii) In case of d is odd, then

$$\mathrm{Shift}_{n-1\uparrow n}(\mathrm{Shift}_{n-2\uparrow n}(\cdots(\mathrm{Shift}_{1\uparrow n}(C(n,d))\cdots)) = \Delta^s(C(n,d)).$$

Since computations of (i) and (ii) are complicated, we omit the proof. □

2.2. Stacked polytopes. We recall the construction of stacked polytopes. Starting with a d-simplex, one can add new vertices by building a shallow pyramids over facets to obtain a simplicial convex d-polytope with n vertices. Such convex polytopes are called *stacked d-polytopes*. Let $P(n,d)$ be the boundary complex of a stacked d-polytope with n vertices. Note that the combinatorial type of $P(n,d)$ is not unique. Then we have the following result for algebraic shifting of stacked polytopes.

Theorem 2.2. *Let $L(n,d)$ be the pure $(d-1)$-dimensional simplicial complex generated by*

$$\{\{2,\dots,d+1\}\}\cup\{(\{1,\dots,d\}\setminus\{i\})\cup\{j\} : 1<i\le d,\ j>d \text{ or } j=i\}.$$

Let $P(n,d)$ be the boundary complex of a stacked d-polytope with n vertices. Then

(i) *One has $\Delta^e(P(n,d)) = \Delta^s(P(n,d)) = L(n,d)$.*

(ii) *If Γ is the boundary complex of a simplicial d-polytope with n vertices, then one has*

$$\Delta^s(P(n,d)) \subset \Delta^s(\Gamma).$$

Proof. (sketch) The equality $\Delta^s(P(n,d)) = L(n,d)$ and (ii) easily follows from the Lefschetz property of the boundary complex of simplicial polytopes.

We will show $\Delta^e(P(n,d)) = \Delta^s(P(n,d))$. The case $d=2$ is easy. In case of $d\ge 3$, by using Lemma 1.4, it is not hard to show that if $\Delta^e(P(n,d)) \ne L(n,d)$ then $\{d+1,d+2\}\in\Delta^e(P(n,d))$. Note that $\{d+1,d+2\}\notin\Delta^s(P(n,d))$. On the other hand, it is known that 1-skeleton of $P(n,d)$ is a chordal graph if $d\ge 3$. It follows from [**7**, Theorem 4.8] that if G is a chordal graph then $\Delta^e(G) = \Delta^s(G)$. This says that $\{d+1,d+2\}\notin\Delta^e(P(n,d))$ and $\Delta^e(P(n,d)) = \Delta^s(P(n,d)) = L(n,d)$. □

References

[1] A. Aramova, J. Herzog and T. Hibi, Gotzmann theorems for exterior algebras and combinatorics, *J. Algebra*, **191** (1997), 174–211.

[2] A Brøndsted, *An Introduction to Convex Polytopes*, Springer-Verlag, New York, 1982.

[3] P. Erdös, C. Ko and R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser.* (2) **12** (1960), 313 – 320.

[4] G. Kalai, The diameter of graphs of convex polytopes and f-vector theory, *in* "Applied Geometry and Discrete Mathematics" (P. Gritzmann and B. Sturmfels, Ed.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, **4** (1991), pp. 387–411.

[5] G. Kalai, Algebraic shifting, *in* "Computational Commutative Algebra and Combinatorics" (T. Hibi, Ed.), Advanced Studies in Pure Math., Volume 33, (2002), pp. 121 – 163.

[6] J. Herzog, Generic initial ideals and graded Betti numbers, *in* "Computational Commutative Algebra and Combinatorics" (T. Hibi, Ed.), Advanced Studies in Pure Math., Volume 33, (2002), pp. 75 – 120.

[7] S. Murai, Algebraic shifting of finite graphs, 2005, arXiv:math.CO/0505010, (preprint).

[8] S. Murai, Generic initial ideals and exterior algebraic shifting of join of simplicial complexes, arXiv:math.CO/0506298, 2005, (preprint).

[9] S. Murai and T. Hibi, The behavior of graded Betti numbers via algebraic shifting and combinatorial shifting, 2005, arXiv:math.AC/0503685 (preprint).
[10] R. P. Stanley, *Combinatorics and Commutative Algebra, Second Edition*, Birkhäuser, Boston, Basel, Berlin, 1996.

Department of Pure and Applied Mathematics, Graduate School of Information Science and Technology, Osaka University, Toyonaka, Osaka, 560-0043, Japan,
E-mail address: s-murai@ist.osaka-u.ac.jp

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A general bijection for a class of walks on the slit plane

Philippe Nadeau

Abstract. We study walks in the plane $\mathbb{Z}^2$, with steps in a given finite set $\mathfrak{S}$, which start from the origin but otherwise never hit the half-line $\mathcal{H} = \{(k,0), k \leqslant 0\}$. These ***walks on the slit plane*** have received some attention these last few years, since in particular their enumeration leads to simple closed formulas; but only one bijection has been found so far, in the case of the square lattice, that explains such formulas.

Let $p = p(\mathfrak{S})$ be the smallest possible abscissa x such that there is a walk on the slit plane ending at $(x, 0)$. Suppose that $|j| \leqslant 1$ for each step $(i, j) \in \mathfrak{S}$. The main result of this paper is the construction of a length preserving bijection between $\mathfrak{S}$-walks on the slit plane with a marked step ending at $(p, 0)$, and a certain class of walks on the plane whose enumeration is much simpler. This allows us to interpret combinatorially previously known enumerations, and to give many new ones.

Résumé. Nous étudions des chemins dans le plan $\mathbb{Z}^2$, dont les pas appartiennent à un ensemble $\mathfrak{S}$ donné, qui partent de l'origine mais qui sinon évitent la demi-droite $\mathcal{H} = \{(k,0), k \leqslant 0\}$. Ces ***chemins sur le plan incisé*** ont éveillé un certain intérêt ces dernières années, notamment en raison d'énumérations menant à des formules closes simples ; cependant une seule bijection a jusqu'ici été trouvée, dans le cas du réseau carré, pour expliquer de telles formules.

Soit $p = p(\mathfrak{S})$ la plus petite abscisse x telle qu'il existe un chemin sur le plan incisé terminant en $(p, 0)$. On suppose que $|j| \leqslant 1$ pour tout pas $(i, j) \in \mathfrak{S}$. Le résultat principal de cet article est la construction d'une bijection, préservant la longueur, entre les $\mathfrak{S}$-chemins dans le plan incisé avec un pas marqué terminant en $(p, 0)$, et une certaine classe de chemins du plan dont l'énumération est plus aisée. Cela nous permet de donner des interprétations combinatoires de résultats d'énumération déjà connus, et d'en donner de nombreux autres.

1. Introduction

Walks on the slit plane were introduced in **[2]**. Given a finite set of steps $\mathfrak{S} \subset \mathbb{Z}^2$, they are defined as walks on the plane $\mathbb{Z}^2$ with steps in $\mathfrak{S}$ that start at the origin O, and otherwise avoid the half-line $\mathcal{H} = \{(k,0), k \leqslant 0\}$. In the paper **[2]** and the following paper **[3]**, the goal was to give closed forms for various generating functions related to these walks, and study when such generating functions were algebraic.

One of the main results of **[3]** is that, in the case where all elements (i, j) of $\mathfrak{S}$ verify $|j| \leqslant 1$ ($\mathfrak{S}$ is said to have the *small height variation property*), then the generating functions $S(x, y, t), S_j(x, t), S_{i,j}(t)$ are algebraic, these series enumerating respectively walks according to length and endpoint, walks ending at height j according to length and final abscissa, and walks ending at (i, j) according to length.

For specific steps $\mathfrak{S}$, the closed form of the generating functions allows in fact to obtain expressions for the number of walks of length n ending at certain points (i, j). For instance, for $\mathfrak{S} = \{(\pm1, 0), (0, \pm1)\}$ (the square lattice) or $\mathfrak{S} = \{(\pm1, \pm1)\}$ (the diagonal lattice), closed formulas are proved in **[2]** and **[6]** for various endpoints. There is a specific endpoint for which closed formulas are often obtained : it is $(p, 0)$, where $p = p(\mathfrak{S})$ is the smallest possible abscissa x such that there is a walk on the slit plane ending at $(x, 0)$; for instance, p equals 1 for the square lattice and 2 for the diagonal lattice.

In this paper, we construct bijections that will explain these closed formulas for the endpoint $(p, 0)$. First we will define a bijection which is valid for sets $\mathfrak{S}$ with small variations and that are symmetric with respect

2000 *Mathematics Subject Classification.* 05A15.
Key words and phrases. Enumerative combinatorics, bijective combinatorics, cyclic lemma, slit plane, lattice walks.

to the x-axis; the main idea for the construction comes from the paper **[1]**, where it was used to enumerate these walks on the square lattice. Then we will modify this bijection in the case of sets $\mathfrak{S}$ that are not necessarily symmetric; we will thus be able to bijectively prove, for instance, that if $\mathfrak{S} = \{(\pm 1, 1), (0, -1)\}$ then there are $4^{2n+1}\binom{2n+1}{n}/(4n+2)$ walks of length $4n+2$ on the slit plane that end at $(1, 0)$.

2. Preliminaries

In this section we will define standard notions concerning walks on the lattice $\mathbb{Z}^2$, and give specific definitions for the case of walks on the slit plane.

For the rest of this section, we let $\mathfrak{S}$ be a finite subset of $\mathbb{Z}^2$.

2.1. Walks. A *walk* with steps in $\mathfrak{S}$ is a finite sequence $w = (w_0, w_1, \ldots, w_n)$ of points of $\mathbb{Z}^2$ such that $w_0 = (0, 0)$ and $w_i - w_{i-1} \in \mathfrak{S}$ for $1 \leqslant i \leqslant n$. We shall also say that w is a $\mathfrak{S}$-walk. Note that our walks are always assumed to start at the origin. The number of steps n is the *length* of w. The *endpoint* of w is w_n, and it is denoted $end(w)$. We also denote the (final) height and abscissa of w by $y(w)$ and $x(w)$, that is $end(w) = (x(w), y(w))$. A walk with a *marked step* is the data of a walk $w = (w_0, w_1, \ldots, w_n)$ together with an integer $i \in [\![0, n-1]\!]$, so a step $w_{i+1} - w_i$ is distinguished (or more precisely, an occurrence of this step in w). For the figures of this paper, marked steps will be distinguished by a thicker line.

2.2. Walks and words. A walk is characterized by a finite sequence of steps of $\mathfrak{S}$. Hence, it will be convenient to consider walks as words on the alphabet $\mathfrak{S}$. The set of words $\mathfrak{S}^*$ is equipped with the usual concatenation product; as usual, ε denotes the empty word.

Any word w of $\mathfrak{S}^*$ will thus be thought of as a walk starting from $(0, 0)$, and we will in fact make no distinction between the walk and the word: for instance, if w_1 and w_2 are two walks, then $w_1 w_2$ is the walk w_1 followed by the walk w_2 (which is attached at the endpoint of w_1). Note that *end* is then a morphism from the monoid $\mathfrak{S}^*$ to $\mathbb{Z}^2$, where elements of $\mathbb{Z}^2$ are added componentwise. That is, when w_1 and w_2 are two walks, then we have $end(w_1 w_2) = end(w_1) + end(w_2)$.

2.3. Walks on the Slit Plane. We will use the terminology introduced in **[2, 3]**. We say that the walk w avoids the half-line $\mathcal{H} = \{(k, 0), k \leqslant 0\}$ if none of the vertices $w_1, \ldots, w_n$ belong to $\mathcal{H}$. We then call w a ***walk on the slit plane***. For $(i, j) \neq (0, 0)$, we denote by $\mathcal{S}_{i,j}(n)$ the set of walks w on the slit plane of length n and such that $end(w) = (i, j)$. We also denote the cardinality of this set by $S_{i,j}(n)$. In this paper, we will consider $\mathcal{S}_{p,0}(n)$, where $p = p(\mathfrak{S})$ is the minimum positive integer x such that there is a $\mathfrak{S}$-walk ending at $(x, 0)$; we shall always assume that we deal with sets $\mathfrak{S}$ such that p is well defined.

If $s = (x, y) \in \mathbb{Z}^2$, we note $\tilde{s} = (x, -y)$ its symmetric with respect to the x-axis. We extend this definition to walks: if $w = s_1 s_2 \cdots s_n \in \mathfrak{S}^*$, then $\widetilde{w} = \tilde{s_1}\tilde{s_2} \cdots \tilde{s_n}$. Geometrically, $\widetilde{w}$ is the walk symmetric of w with respect to the x-axis.

We now define two properties of sets of steps, illustrated on Figure 1 :

DEFINITION 2.1. Let $\mathfrak{S}$ be a set of steps.

- $\mathfrak{S}$ is *symmetric* (with respect to the x-axis) if for all $s \in \mathfrak{S}$, then $\tilde{s} \in \mathfrak{S}$.
- The set $\mathfrak{S}$ is said to have *small height variations* if, for all $(i, j) \in \mathfrak{S}$, $|j| \leqslant 1$.

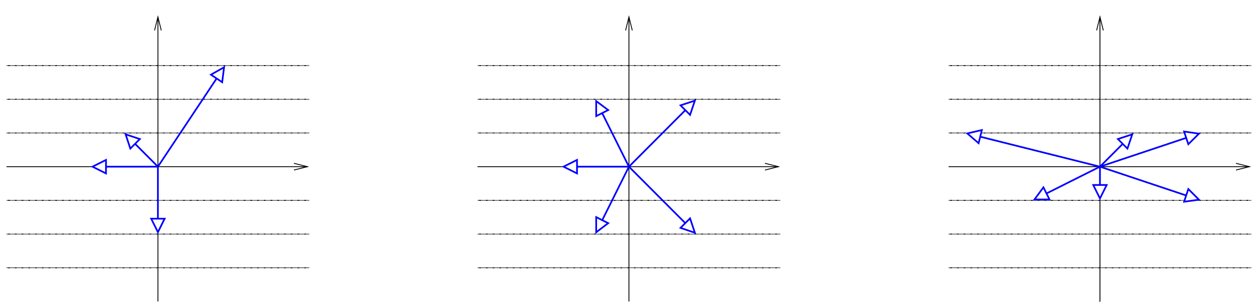


FIGURE 1. From left to right: a set $\mathfrak{S}$ which is neither symmetric nor with small variations, a symmetric set, and a set with small variations.

3. The main theorems

We consider in this section (and in the rest of the paper) a finite set $\mathfrak{S} \subset \mathbb{Z}^2$ of *steps* which has the *small height variations* property. Let $w = (w_0, \ldots, w_n)$ be an element of $\mathcal{S}_{p,0}(n)$, i.e. a walk on the slit plane of length n that ends at $(p, 0)$. Let also $i \in [\![0, n-1]\!]$.

First suppose that, in addition, $\mathfrak{S}$ is *symmetric*. Then define $\Psi(w,i) = (W_0, W_1, \ldots, W_n)$ as the walk

- whose beginning $(W_0, W_1, \ldots, W_{n-i})$ is obtained by reflecting $(w_i, \ldots, w_n)$ off the x-axis and translating it so that $W_0 = O$;
- and whose ending is obtained by appending $w = (w_0, \ldots, w_i)$ to $(W_0, W_1, \ldots, W_{n-i})$ through the translation of vector $W_{n-i} - w_0$

See Figure 2 for an illustration of this construction. If w is considered as a word in $\mathfrak{S}^*$ (see Section 2), the choice of i corresponds to a factorization $w = uv$ in $\mathfrak{S}^*$ with $v \neq \varepsilon$. Then one has simply $\Psi(w,i) = \tilde{v}u$.

We can now state our first theorem:

THEOREM 3.1. *The construction Ψ is a bijection between the following two sets:*

(1) $\mathfrak{S}$-walks on the slit plane of length n ending at $(p, 0)$ with a marked step.
(2) $\mathfrak{S}$-walks on the plane of length n ending at $(p, 2k)$ for a certain $k \in \mathbb{Z}$.

As an immediate corollary, we have:

COROLLARY 3.1. *Let $S_{p,0}(n)$ be the number of walks on the slit plane that end at $(p, 0)$. Let also $W_{p,even}(n)$ be the number of walks on the plane that end at $(p, 2k)$ for a certain integer k.*

Then we have the identity

$$n \cdot S_{p,0}(n) = W_{p,even}(n)$$

We now generalize this bijection in the case where $\mathfrak{S}$ is not assumed to be symmetric (but still has the small height variation property). Note that the construction of the previous theorem cannot function as such, because, after reflection, some of the steps may not be elements of $\mathfrak{S}$.

Let us write $\mathfrak{S}_{sym}$ for the set of steps whose elements are the symmetric of those of $\mathfrak{S}$ off the x-axis; in other words $\mathfrak{S}_{sym}$ equals $\widetilde{\mathfrak{S}}$. Then we define $\overline{\mathfrak{S}} = \mathfrak{S} \cup \mathfrak{S}_{sym}$. We also need to define $\mathfrak{S}^\delta = \{s \in \mathfrak{S} \mid y(s) = \delta\}$ for $\delta \in \{-1, 0, 1\}$, and similarly $\overline{\mathfrak{S}^1}$ and $\overline{\mathfrak{S}^{-1}}$. We can now state our second theorem:

THEOREM 3.2. *Let $\mathfrak{S}$ be a set of steps with small variations, and n be a positive integer. Assume that there is a $\mathfrak{S}$-walk ending on the positive x-axis, and let $p = p(\mathfrak{S})$ be the smallest positive abscissa that can be reached. Then we have a bijection between the following sets:*

(1) Walks on the slit plane of length n with steps in $\mathfrak{S}$ that end at $(p, 0)$.
(2) Walks of length n with steps in $\overline{\mathfrak{S}}$, ending at abscissa p, with an even number $2m$ of steps in $\overline{\mathfrak{S}^1} \cup \overline{\mathfrak{S}^{-1}}$, such that, among these steps, the first m ones are in $\overline{\mathfrak{S}^1}$ and the last m ones are in $\overline{\mathfrak{S}^{-1}}$

Notice that if $\mathfrak{S}$ is symmetric then the walks in 3.1(2) and 3.2(2) coincide; in fact the bijections will be identical in this case.

4. Proof of Theorem 3.1

We consider here a set $\mathfrak{S}$ of steps that is symmetric and has the small height variation property. We keep the notations of Section 3 concerning Theorem 3.1; in particular (u, v) is the factorization of w afforded by the marked step. Note first that the steps of $\Psi(w,i)$ are in $\mathfrak{S}$, since we assumed that $\mathfrak{S}$ is symmetric. An example of $\Psi(w,i)$ is shown on Figure 2.

The proof will proceed as follows : first we show that Ψ is well defined. Then we construct a function Γ from $\mathcal{W}_{p,even}(n)$ to $\mathcal{S}_{p,0}(n) \times [\![0, n-1]\!]$. Finally, we prove that Ψ and Γ are actually inverse to one another.

Ψ *is well-defined.* We have to show that $\Psi(w,i)$ is a walk with endpoint at abscissa p and even ordinate. This is obvious from the geometric construction, and follows from a simple computation. Define (h, k) by $end(u) = (h, k)$. Since $end(w) = (p, 0)$, it follows that $end(v) = end(w) - end(u) = (p - h, -k)$, and consequently $end(\tilde{v}) = (p - h, k)$. Finally, we have $\Psi(w,i) = \tilde{v}u$, so $end(\Psi(w,i)) = end(\tilde{v}) + end(u) = (p, 2k)$, which shows that indeed $\Psi(w,i)$ is an element of $\mathcal{W}_{p,even}(n)$.

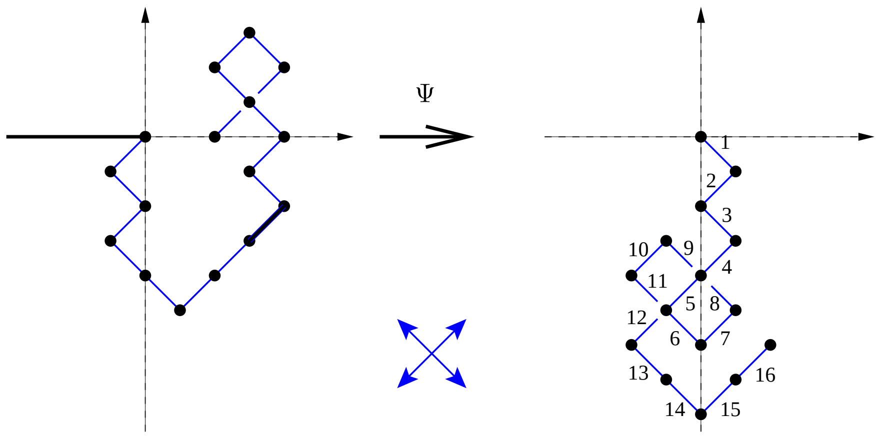


FIGURE 2. Example of the bijection Ψ in the case of the diagonal lattice. Here $n = 16$, $i = 7$, and steps are numbered in the second walk for easier understanding.

Definition of the inverse. Let us now define a function Γ from $\mathcal{W}_{p,even}(n)$ to $\mathcal{S}_{p,0}(n) \times [\![0, n-1]\!]$. Let $W = s_1 s_2 \cdots s_n$ be a $\mathfrak{S}$-walk that ends at $(p, 2l)$ with $l \in \mathbb{Z}$. Since $\mathfrak{S}$ has the small height variation property, there are points in W with ordinate l. Among such points, let (m, l) be the one with minimal abscissa. Let finally $i \in [\![1, n]\!]$ be maximal such that s_i is a step starting from (m, l). Note that such a step always exists because (m, l) cannot be the endpoint of W. Indeed, this is clear if $l \neq 0$; and in the case $l = 0$, we have $m \leqslant 0$ because W starts from O, whereas the walk W ends in $(p, 0)$ with $p > 0$.

Let us define $W_1 = s_1 \cdots s_{i-1}$ and $W_2 = s_i \cdots s_n$; we then set

$$\begin{aligned} \Gamma(W) &= (W_2 \widetilde{W_1}, n - i + 1) \text{ if } i > 1 \\ &= (W_2, 0) \text{ if } i = 1 \end{aligned}$$

See Figure 3 for an example.

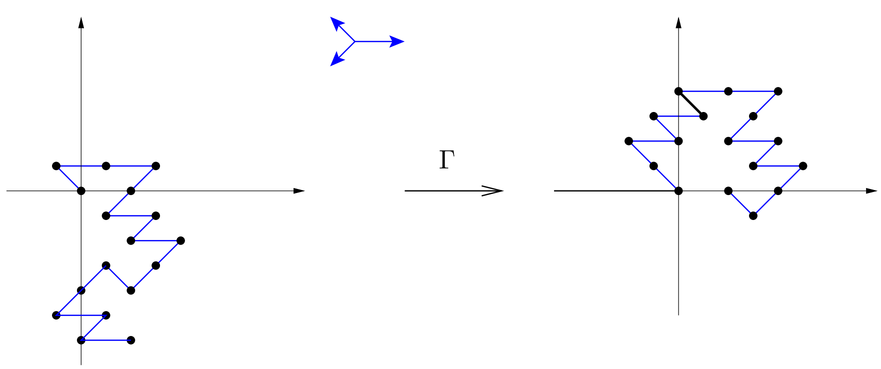


FIGURE 3. Example of the inverse bijection Γ when $\mathfrak{S} = \{(2,0), (-1,1), (-1,-1)\}$. Here we have $n = 16$, $(m, l) = (1, -3)$ and $i = 12$.

Let us rephrase the properties of W_1 in the language of words; we will state this as a lemma for easier reference:

LEMMA 4.1. *Let U be a prefix of W such that $end(U) = (k, l)$. Then we have $k \geqslant m$, and the inequality is strict if U is a prefix strictly longer than W_1.*

We claim that the walk thus obtained is an element of $\mathcal{S}_{p,0}(n)$; by abuse of notation, we will write $\Gamma(W)$ for $W_2 \widetilde{W_1}$. Firstly, we have $end(W_1) = (m, l)$, so that $end(\widetilde{W_1}) = (m, -l)$, and $end(W_2) = (p - m, l)$. So we have $end(\Gamma(W)) = (p, 0)$ which shows that $\Gamma(W)$ has the good endpoint.

Now we have to prove that for every nonempty prefix w_1 of $\Gamma(W)$ such that $end(w_1) = (x, 0)$, then x is a positive integer. Let w_1 be such a prefix: there are two cases to consider, depending on whether w_1 is a shorter prefix than W_2 or not. If $w_1 \cdot u = W_2$, then $W_1 w_1$ is a prefix of W strictly longer than W_1 and its endpoint is $(m + x, l)$. By Lemma 4.1, we have indeed $x > 0$. If $W_2 \cdot u = w_1$, then u is a prefix of $\widetilde{W_1}$, so that $\tilde{u}$ is a prefix of W_1. But $end(\widetilde{u}) = end(\widetilde{w_1}) - end(\widetilde{W_2}) = (m - p + x, l)$. By Lemma 4.1, we have $x - p \geqslant 0$, which implies again $x > 0$ because $p > 0$.

This completes the proof that Γ is a well-defined function from $\mathcal{W}_{p,even}(n)$ to $\mathcal{S}_{p,0}(n)$.

End of the proof of Theorem 3.1. We finally have to show that Γ is the inverse of Ψ. It is clear that $\Psi(\Gamma(W)) = W$, so that we need to prove that $\Gamma(\Psi(w, i)) = (w, i)$. This is clearly equivalent to showing that $\widetilde{v}$ is equal to the prefix W_1 of $\Psi(w, i)$ defined in the construction of Γ. If $end(u) = (h, k)$, we have already computed that $end(\widetilde{v}) = (p - h, k)$ and $end(\Psi(w, i)) = end(\widetilde{v}u) = (p, 2k)$. So what we have to show is that (1) $p - h = m$, where m is defined as in the construction of Γ, and that (2) if U is a prefix of $\Psi(w, i) = \widetilde{v}u$, longer than $\tilde{v}$, and whose endpoint equals (x, k) for a certain x, then $x > m$.

Suppose that $p - h > m$. Assume first that $\widetilde{v}$ is a strict prefix of W_1, so that there exists $u_0 \neq \varepsilon$ such that $W_1 = \widetilde{v}u_0$, and $end(u_0) = (x, 0)$ where $x = m - (p - h) < 0$. Since $W_1 W_2 = \widetilde{v}u$, this implies $u = u_0 W_2$, which is absurd since u_0 is a prefix of w hitting the forbidden half-line. We have a contradiction, so $\widetilde{v}$ is a suffix of W_1, or, equivalently, $\widetilde{W_1}$ is a prefix of v: there follows that $u\widetilde{W_1}$ is a prefix of w with endpoint $(m + h, 0)$. But we supposed that $m + h < p$, so this contradicts the definition of p. (Note that this is the only place where the definition of p is used in the proof). Finally $p - h \leqslant m$, and the definition of m forces $p - h \geqslant m$, so (1) is proved.

Now let U be a prefix of $\widetilde{v}u$, with $end(U) = (x, k)$, such that $U = \widetilde{v}u_0$ with $u_0 \neq \varepsilon$. Then u_0 is a nonempty prefix of u with $end(u) = (x + h - p, 0) = (x - m, 0)$. Since $w = uv$ is a walk on the slit plane, we must have $x - m > 0$, which proves (2), and completes the proof of Theorem 3.1.

REMARK 4.2. This construction is a generalization of the one in **[1]**. Indeed, in the special case of the square lattice, what they did is marking a specific step in each walk, namely the last one with origin at the smallest possible abscissa. If this is the marked step in our bijection, then it reduces to theirs.

Marking a unique step in this fashion is not always feasible, but it is possible for a certain category of sets $\mathfrak{S}$. We state this a corollary, and it is a direct generalization of the construction of **[1]** :

COROLLARY 4.1. *Let n be a positive integer, $\mathfrak{S}$ be a symmetric set of steps with small variations, which contains only one step with positive abscissa, namely $(1, 0)$. Then there is a bijection between $\mathcal{S}_{1,0}(n)$ and walks of length $n - 1$ that end at $(0, 2k)$ for a certain $k \in \mathbb{Z}$ and stay in the right half-plane $x \geqslant 0$.*

REMARK 4.3. The construction Ψ is easily seen to remain injective when $\mathfrak{S}$ is not symmetric (and even when $\mathfrak{S}$ has "large variations"). Let us describe the image of Ψ in this case, as we will use it in the next Section:

PROPOSITION 4.1. *Let $\mathfrak{S}$ have the small heigth variation property, and n be a positive integer. Then Ψ is a bijection between:*

(1) $\mathfrak{S}$-walks on the slit plane of length n ending at $(p, 0)$ with a marked step.
(2) $\overline{\mathfrak{S}}$-walks on the plane of length n ending at $(p, 2k)$ for a certain $k \in \mathbb{Z}$, such that if W_1 and W_2 are such as defined in the construction of Γ, then W_1 has its steps in $\mathfrak{S}^{sym}$ and W_2 has its steps in $\mathfrak{S}$.

Note that this proposition is not true in general if $\mathfrak{S}$ does not have small variations, because in this case there may be walks with steps in $\overline{\mathfrak{S}}$ that end at $(p, 2k)$ that do not have any point at height k, so that W_1 and W_2 are not well-defined.

5. Proof of Theorem 3.2

We will deal in all this section with a set of steps $\mathfrak{S}$ with the small height variation property. We also still assume that $p(\mathfrak{S})$ is well defined, that is, there exists a $\mathfrak{S}$-walk that ends on the positive axis.

The bijection announced in Theorem 3.2 will be defined by first introducing some intermediate objects. The reader is advised to have a look at Figure 4 for an illustration of the different constructions of this section.

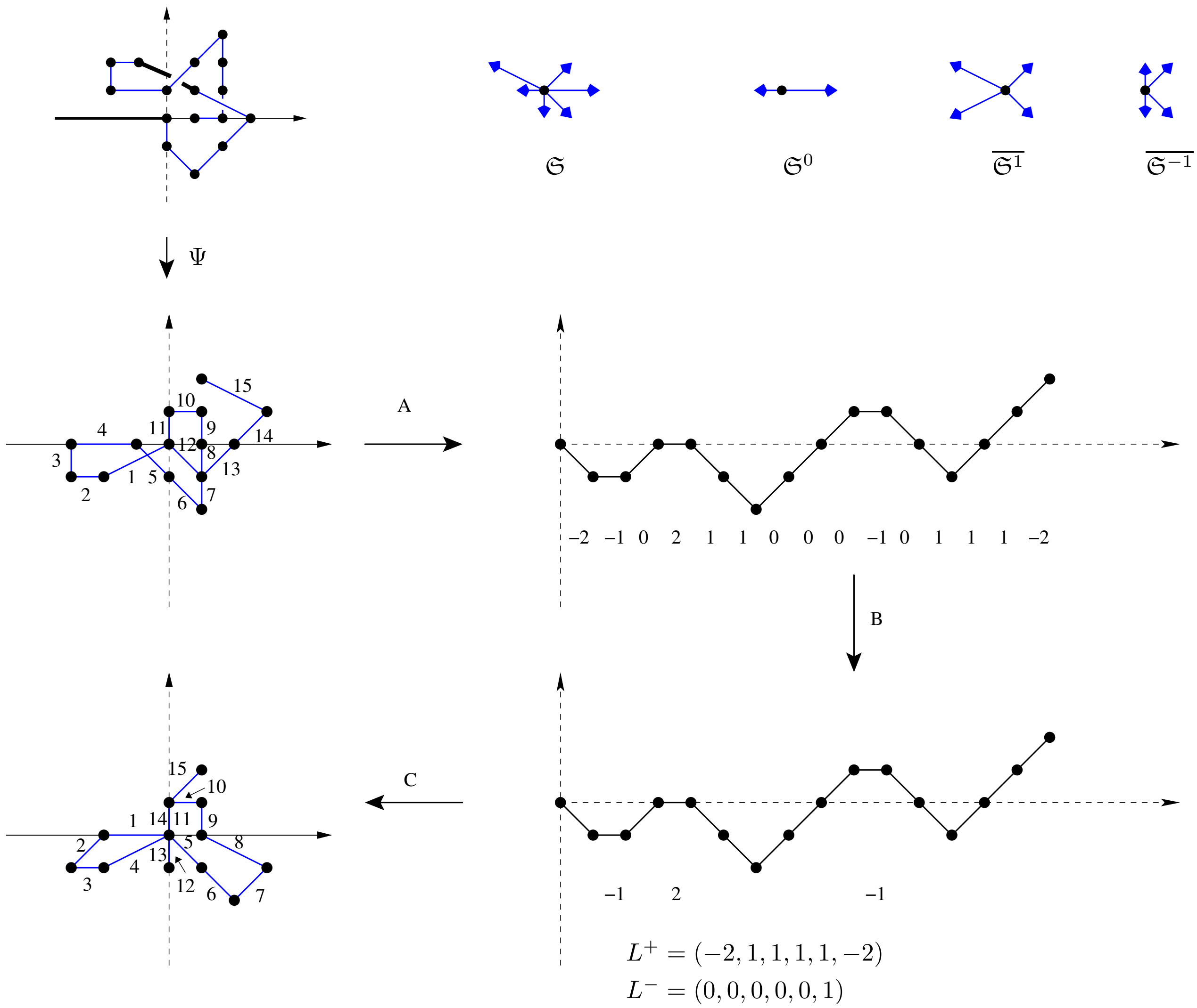


Figure 4. Illustration of the different steps of the bijection of Theorem 3.2. Construction A is involved in Lemma 5.3, construction B is the bijection of Theorem 5.4, and construction C is explained at the end of Section 5.

Let $\mathfrak{M}$ be the set of steps $\{(1,1),(1,0),(1,-1)\}$; the letter $\mathfrak{M}$ stands for Motzkin, since $\mathfrak{M}$-walks that end on the x-axis and remain in the upper half plane are the famous *Motzkin paths*, see for instance [**5**] for more information.

Definition 5.1. Let n be a positive integer, and $M = m_1 \cdots m_n$ be a $\mathfrak{M}$-walk of length n. A *labeling* of M is a function l from $[\![1,n]\!]$ to $\mathbb{Z}$. We also associate to a labeling l the function $\hat{l}$ defined by $\hat{l}(i) = l(1) + \cdots + l(i)$ for $i \in [\![0,n]\!]$.

One should think of l as being associated to the occurrences of the steps of M, and $\hat{l}$ as associated to the point of abscissa i in M. Now for $\delta \in \{-1,0,1\}$, we define $S^\delta \subset \mathbb{Z}$ as the abscissas of the walks in $\mathfrak{S}^\delta$.

Definition 5.2. Let $M = m_1 \cdots m_n$ be an $\mathfrak{M}$-walk of length n ending at height $2k$ ($k \in \mathbb{Z}$), together with a labeling l. Consider all integers i such that (i,k) is a point of M, and consider among them those with $\hat{l}(i)$ minimal; denote $i_{max} = i_{max}(M,l)$ the maximal integer with this property.

Then M is *well* $\mathfrak{S}$*-labeled* by l (or l is a *good* $\mathfrak{S}$*-labeling* of M) if the following conditions are verified:
– for all $i \leqslant i_{max}$ and $\delta \in \{-1,0,1\}$, $l(i) \in S^\delta$ iff $m_i = (1,-\delta)$;
– for all $i > i_{max}$ and $\delta \in \{-1,0,1\}$, $l(i) \in S^\delta$ iff $m_i = (1,\delta)$;

This seemingly artificial definition is explained by the following lemma:

LEMMA 5.3. *There is a bijection between*

(1) $\mathfrak{S}$*-walks on the slit plane of length* n *ending at* $(p, 0)$
(2) couples (M, l) *where* M *is an* $\mathfrak{M}$*-walk of length* n *ending at even height, well* $\mathfrak{S}$*-labeled by* l*, and verifying* $\hat{l}(n) = p$*.*

PROOF. First apply Proposition 4.1. Then transform a walk $w = s_1 \cdots s_n$ thus obtained in the following way: for every $i \in [\![1, n]\!]$, define $m_i = (1, y(s_i))$ and $l(i) = x(s_i)$. The resulting path $M = m_1 \cdots m_n$ with the function l give then the desired bijection, as is easily seen: notice that the length of W_1 in the intermediate walk is equal to the integer i_{max} in the definition of a good labeling. □

We will now state the main step of the bijection:

THEOREM 5.4. *Let* $M = m_1 \cdots m_n$ *be an* $\mathfrak{M}$*-path of length* n *ending at even height. There is a bijection between:*

(1) good $\mathfrak{S}$*-labelings* l *of* M *with* $\hat{l}(n) = p$*.*
(2) 3-uples (L^+, L^-, l_0) *such that* l_0 *is a function from* $\{i \in [\![1, n]\!] \mid m_i = (1, 0)\}$ *to* S^0*, and* L^+ *(respectively* L^-*) is a sequence of* m *elements of* S^+ *(resp.* S^-*) where* $2m$ *is defined as the number of* i *such that* $m_i \in \{(1, 1), (1, -1)\}$*.*

SKETCH OF THE PROOF. Let us define this bijection. First, $l_0(i)$ is simply defined as $l(i)$: the labels for horizontal steps remain unchanged. To define $L^+(i)$ and $L^-(i)$, recall the definition of i_{max} given in 5.2. For i increasing, L^+ consists first of all labels $l(i)$ of steps $m_i = (1, -1)$ for $i \leqslant i_{max}$, followed by the labels of steps $m_i = (1, 1)$ for $i > i_{max}$. Similarly, L^- consists of all labels $l(i)$ of steps $m_i = (1, 1)$ for $i \leqslant i_{max}$ followed by the labels of steps $m_i = (1, -1)$ for $i > i_{max}$. One checks easily that this is well defined.

The main problem is to inverse this construction, and for this we must determine the abscissa i_{max}, so that we can be assured that we obtain a *good labeling*. Indeed, let us try to define the function l given by L^+, L^- and l_0. Of course we have to set $l(i) = l_0(i)$ for all i such that m_i is horizontal. There is clearly only one way to define labels for the first abscissas i, up until we hit an abscissa i_0 whose corresponding point m_{i_0} is at height k (where k is defined by $y(M) = 2k$): we have to use (in their original order) the elements of L^+ to label steps of the form $(1, -1)$, and elements of L^- to label those of the form $(1, 1)$. Now we have to know whether i_{max} is equal to i_0 or not, in order to know if we have to switch the roles of L^+ and L^-. Clearly, if there is only one way to define i_{max}, then we have found the only possible inverse construction.

Here is the way to do it: if there is no other point of height k, then clearly $i_{max} = i_0$. If not, let $i_1, \ldots, i_t$ be all other abscissas whose corresponding points m_i are at height k. The key point is that, for every j, $\hat{l}(i_j)$ does not depend on whether we have already decided (haphazardly) to switch the roles of L^+ and L^- after a certain i_l or not. Indeed, one verifies that we have to use the same elements of both lists L^+ and L^- whatever our choice: this is a direct consequence of the fact that between two points m_{i_l}, there are as many steps up and down since all these points are at the same height k. So, now that we know the values of $\hat{l}$ for the abscissas i_l, there is only one way to define i_{max}, and we can construct the labeling l.

It is clear that these constructions are inverse to one another; then it remains to check that these constructions are well defined to complete the proof. This is easy and will be omitted in this abstract. □

Finally, let us show that this theorem implies Theorem 3.2. Let M and (L^+, L^-, l_0) be as in the above theorem, and we will bijectively associate to such data a $\overline{\mathfrak{S}}$-walk such as described in Theorem 3.2(2). The construction is simple: let $i_1 < \ldots < i_m < j_1 < \ldots < j_{2m}$ be the indices i such that $m_i \in \{(1, 1), (1, -1)\}$. Then define a $\overline{\mathfrak{S}}$-walk $w = s_1 \cdots s_n$ by $s_i = (l_0(i), 0)$ when $m_i = (1, 0)$, $s_{i_t} = (L^+(t), y(m_{i_t}))$ and $s_{j_t} = (L^-(t), y(m_{j_t}))$ for $t \in [\![1, m]\!]$ (see Figure 4). It is straightforward to show that this construction is well defined and bijective. By Lemma 5.3, this completes the proof of Theorem 3.2.

REMARK 5.5. We will quickly explain explain where the idea for the walks described in Theorem 3.2(2) comes from. Let $A_\delta(x) = \sum_{i \in S^\delta} x^i$ for $\delta \in \{-1, 0, 1\}$. Mireille Bousquet-Mélou [**3**] proves the following:

THEOREM 5.6 ([3]). *Let $\Delta(x;t)$ be the following polynomial in x, x^{-1} and t:*

$$\Delta(x;t) = (1 - tA_0(x))^2 - 4t^2 A_1(x)A_{-1}(x)$$

Then the generating function for walks on the slit plane ending at $(p,0)$ is

$$\sum_{n=1}^{\infty} S_{p,0}(n)t^n = [x^p] \log\left(\frac{1}{\sqrt{\Delta(x;t)}}\right)$$

For $i \in \mathbb{Z}$, define $a_i = |\{(j,k) \ / \ (j,1),(k,-1) \in \mathfrak{S} \text{ and } i = j+k\}|$. That is, a_i is equal to the number of couples $(s^+, s^-) \in \mathfrak{S}^+ \times \mathfrak{S}^-$ such that $end(s^+s^-) = (i,0)$. Then we have easily $A_1(x)A_{-1}(x) = \sum_i a_i x^i$, and

$$(A_1A_{-1})^m = (\sum_i a_i x^i)^m = \sum_{\substack{m_1,\dots,m_k \\ \sum_i m_i = m}} \binom{m}{m_1,\dots,m_k} \left(\prod_i a_i^{m_i}\right) x^{\sum_i i m_i}$$

Then some standard calculations using Theorem 5.6 lead to an expression of $S_{p,0}(n) = [x^p t^n] \log\left(\frac{1}{\sqrt{\Delta(x;t)}}\right)$, that can naturally be interpreted as in Theorem 3.2(2).

6. Applications

6.1. Examples. We will apply both theorems to particular sets of steps for which closed formulas exist.

First, let us deal with the diagonal lattice (for which $p = 2$). Let $C_n = \binom{2n}{n}/(n+1)$ be the nth Catalan number.

PROPOSITION 6.1 ([3]). *Let n be a positive integer. There are $4^n C_n/2$ walks on the slit plane with steps in $\{(\pm 1, \pm 1)\}$ of length $2n$ that end in $(2,0)$.*

PROOF. Let D_n be this number. By Theorem 3.1, we have to enumerate walks of length $2n$ that end at $(2,2l)$ where $l \in \mathbb{Z}$. In fact the condition that walks end at an even ordinate is superfluous because the walks are of even length, so we have to enumerate walks that end at abscissa 2.

Let us first choose the occurrences of steps with positive abscissa $(1,1)$ and $(1,-1)$; since the walks end at abscissa $(2,0)$, there are $n+1$ occurrences, so that there are $\binom{2n}{n+1}$ choices. To define a walk completely, it remains to choose if the steps go up or down, and there are clearly $2^{2n} = 4^n$ ways to do that.

Finally, by Theorem 3.1, we have

$$2nD_n = 4^n \binom{2n}{n+1},$$

which is equivalent to the desired formula. □

PROPOSITION 6.2 ([4]). *Let n be a positive integer. There are $4^n \binom{3n}{n}/(n+1)$ walks on the slit plane with steps in $\{(2,0),(-1,1),(-1,-1)\}$ of length $3n+1$ that end in $(2,0)$.*

PROOF. Let K_n be this number. We have to enumerate walks of length $3n+1$ with endpoint at abscissa 2 and at an even ordinate. Let a be the number of steps $(2,0)$ in such a walk. By focusing at abscissas we have $2a - (3n+1-a) = 2$, so that $a = n+1$. Choosing the occurrences of these steps is then counted by $\binom{3n+1}{n+1}$, and then it remains to choose for the remaining $2n$ steps between $(-1,1)$ and $(-1,-1)$. Note that here again the condition that the endpoint has an even ordinate is superfluous. We finally obtain

$$(3n+1)K_n = 4^n \binom{3n+1}{n+1},$$

which gives us the desired enumeration. □

Let us give an example of application of Theorem 3.2.

PROPOSITION 6.3. *Let n be a positive integer. There are $4^{2n+1}\binom{2n+1}{n}/(4n+2)$ walks on the slit plane with steps in $\{(0,-1),(-1,1),(1,1)\}$ of length $4n+2$ that end in $(1,0)$.*

PROOF. Let M_n be this number. By Theorem 3.2, after having marked a step, we have to enumerate walks of length $4n+2$ whose $2n+1$ first steps are elements of $\{(\pm 1, \pm 1)\}$ and $2n+1$ last steps are elements of $\{(0, \pm 1\}$, and ending at abscissa 1. To ensure the condition on the abscissa, there has to be $n+1$ steps in $\{(1, \pm 1)\}$. One then easily obtains the following identity

$$(4n+2)M_n = 2^{2n+1}\binom{2n+1}{n} \cdot 2^{2n+1},$$

which concludes the proof. □

In fact, *for all possible sets $\mathfrak{S}$ of 3 steps, not all of them horizontal*, we can find and prove bijectively closed formulas for $S_{p,0}(n)$. Actually, for such sets the number of occurrences of each step for $w \in \mathcal{S}_{p,0}(n)$ is determined by the length n, and the enumeration of the corresponding walks through the bijection becomes easy. One also always obtains closed formulas for sets of cardinality 4 of the form $\{(a, \pm 1), (-b, \pm 1)\}$ where a and b are nonnegative integers.

6.2. Mean number of returns to a given ordinate. There is an obvious refinement of the bijections of Theorems 3.1 and 3.2: the walks on the slit plane marked at height k are sent to walks that end at height $2k$. We note that the case $k=0$ is a consequence of the "cyclic lemma" stated in [**2**].

This can be applied to answer the following question: given $\mathfrak{S}$ and n, assume uniform distribution on the walks of $\mathcal{S}_{p,0}(n)$, how many times on average do these walks hit the height j? Let H_j^n be the random variable defined on $\mathcal{S}_{p,0}(n)$ (with uniform distribution) by: $H_j^n(w) = |\{i > 0 \mid y(w_i) = j\}$. The following proposition, whose proof is immediate, gives a precise answer:

PROPOSITION 6.4. *Let $\mathfrak{S}$ be symmetric with small height variations, n be a positive integer, and j an integer. Then the expectation $\mathbb{E}(H_j^n)$ is equal to n times the quotient of the number of walks of length n ending at ordinate $2j$ by the number of walks of length n ending at even ordinate.*

A generalization can be stated for non symmetric $\mathfrak{S}$. From this proposition, one can obtain closed formulas for a great number of sets $\mathfrak{S}$. For instance, on the square lattice one can gets :

$$\mathbb{E}(H_j^{2n+1}) = \frac{(2n+2)\binom{2n+1}{n+j+1}\binom{2n+1}{n-j+1}}{\binom{4n+2}{2n+1}}$$

By Stirling's formula, this is asymptotically equivalent to $\frac{2\sqrt{2}}{\sqrt{n}}$ for fixed j and n going to infinity.

7. Concluding remarks

7.1. Cyclic lemma for walks in the upper plane. There is a variation on walks on the slit plane considered already in [**3**]. It deals with walks on the slit plane that in addition stay in the upper half plane $\{y \geqslant 0\}$. Let us consider, for a given set of steps $\mathfrak{S}$, the walks that end at $(p, 0)$ where $p = p(\mathfrak{S})$ is defined as before. Then we have the following theorem :

THEOREM 7.1. *Let n be a positive integer. Walks of length n in the region $(y \geqslant 0) \cap \mathcal{H}$ with a marked step ending at $(p, 0)$ are in bijection with walks in the plane of length n ending at $(p, 0)$.*

SKETCH OF THE PROOF. The bijection is defined in a manner similar Theorem 3.1, and is actually simpler because there is no reflection involved: if $w = uv$ is the factorization given by the marked step, then we associate to it the walk $w' = uv$.

For the inverse bijection, let W be a walk on the plane ending at $(p, 0)$. Then define W_1 as the longest prefix of W such that $y(W_1)$ is minimal, and $x(W_1)$ is minimal among all the prefixes U with $y(U)$ minimal. Then, if $W = W_1W_2$, the inverse construction is defined by W_2W_1 where the first step of W_1 is marked (or the first step if $W_1 = \varepsilon$).

The proof that these are well defined functions which are inverse to one another follows then the same lines as the proof of Theorem, and is actually simpler. 3.1. □

7.2. Other endpoints. In this paper we have dealt with walks on the slit plane ending at a specific endpoint, which is the most natural in that it respects the symmetry of the half line and is as close as possible to it. But, at least for certain sets $\mathfrak{S}$, there are other endpoints that lead to closed formulas for enumeration. Though most of them can be proved using generating functions, only for those ending at $(p, 0)$ exists a bijective proof which was the object of this paper

So an obvious problem is to find bijective proofs for such formulas; our belief is that an approach similar to this paper is feasible (though the construction described here does not work as such), and that a construction for a given set of steps can be most certainly generalized to a whole class of steps.

Another question is to relate some of the problems just mentioned between them; for instance, it was noticed in **[2]** that among the walks of length $2n+1$ that go from $(0,0)$ to $(1,0)$ on the square square lattice, exactly as many avoid the horizontal half-line $\{(k,0) \mid k \leqslant 0\}$ as the diagonal half-line $\{(k,k) \mid k \leqslant 0\}$. There are many examples with various endpoints or steps of this phenomenon, which clearly needs a direct combinatorial explanation.

References

[1] E. Barcucci, E. Pergola, R. Pinzani, S. Rinaldi, A bijection for some paths on the slit plane, Adv. in Appl. Math. 26 (2001), No. 2, 89–96.
[2] M. Bousquet-Mélou and G. Schaeffer, Walks on the slit plane, Probab. Theory Related Fields 124 (2002), No. 3, 305–344.
[3] M. Bousquet-Mélou, Walks on the slit plane: other approaches, Adv. in Appl. Math. 27 (2001), No. 2-3, 243–288.
[4] M. Bousquet-Mélou, Algebraic Generating Functions in Enumerative Combinatorics and Context-Free Languages, STACS 2005, Springer, Lecture Notes in Computer Science 3404, 18–35.
[5] Richard P. Stanley, *Enumerative combinatorics*, vol. 2, Cambridge University Press, 1999.
[6] G. Xin, Proof of a conjecture on the slit plane problem Discrete Math. 282 (2004), No. 1-3, 281–287.

LABORATOIRE DE RECHERCHE EN INFORMATIQUE, UNIVERSITÉ PARIS 11, ORSAY, FRANCE
E-mail address: nadeau@lri.fr

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

On the complexity of computing Kostka numbers and Littlewood-Richardson coefficients

Hariharan Narayanan

Abstract. Kostka numbers and Littlewood-Richardson coefficients appear in combinatorics and representation theory. Interest in their computation stems from the fact that they are present in quantum mechanical computations since Wigner ([**17**]). In recent times, there have been a number of algorithms proposed to perform this task ([**1**], [**13**], [**14**], [**2**], [**3**]). The issue of their computational complexity was explicitly asked by E. Rassart ([**13**]). We prove that the computation of either quantity is $\#P$-complete. This, implies that unless $P = NP$, which is widely disbelieved, there do not exist efficient algorithms that compute these numbers.

Résumé. Les nombres de Kostka et les coefficients de Littlewood-Richardson apparaissent en combinatoire et en théorie de la représentation. Il est intéressant de les calculer car ils apparaissent dans certains calculs en mécanique quantique depuis Wigner ([**17**]). Récemment, plusieurs algorithmes ont été proposés pour les calculer ([**1**], [**13**], [**14**], [**2**], [**3**]). Le problème de la complexité de ce calcul a été posé par E. Rassart ([**13**]). Nous démontrons que le calcul des nombres de Kotska et des coefficients de Littlewood-Richardson est #P-complet. Cela implique que, à moins que P=NP, il n'existe pas d'algorithme efficace pour calculer ces nombres.

1. Introduction

Let $\mathbb{N} = \{1, 2, \ldots\}$ be the set of positive integers and $\mathbb{Z}_{\geq 0} = \mathbb{N} \cup \{0\}$. Let $\lambda = (\lambda_1, \ldots, \lambda_s) \in \mathbb{N}^s$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s \geq 1$, $\mu = (\mu_1, \ldots, \mu_t) \in \mathbb{Z}_{\geq 0}^t$, $\nu = (\nu_1, \ldots, \nu_u) \in \mathbb{Z}_{\geq 0}^u$ and $\alpha = (\alpha_1, \ldots, \alpha_v) \in \mathbb{N}^v$, $\alpha_1 \geq \cdots \geq \alpha_v \geq 1$. The Kostka number $K_{\lambda\mu}$ and the Littlewood-Richardson coefficient $c_{\lambda\alpha}^{\nu}$ play an essential role in the representation theory of the symmetric groups and the special linear groups. Their combinatorial definitions can be found in Section 2. These have been present in quantum mechanical computations since the time of Wigner ([**17**]). Recently, in [**13**], E. Rassart asked whether there exist fast (polynomial time) algorithms to compute Kostka numbers and Littlewood Richardson coefficients (Question 1, page 99). We prove that the two quantities are $\#P$-complete (see Theorems 1, 2). It is known that if a $\#P$-complete quantity were computable in polynomial time, $P = NP$. An explanation of this fact is sketched in Section 2. Thus, under the widely believed hypothesis that $P \neq NP$, there do not exist efficient (polynomial time) algorithms to compute Kostka numbers and Littlewood-Richardson coefficients.

In [**1**], Barvinok and Fomin show how the set of all non-zero $K_{\lambda\mu}$ for a given μ can be produced in time that is polynomial in the total size of the input and output. They also give a probabilistic algorithm running in time, polynomial in the total size of input and output, that computes the set of all non-zero Littlewood-Richardson coefficients $c_{\lambda\mu}^{\nu}$ given λ and μ. In [**3**], methods for the explicit computation of the Kostka numbers and Littlewood-Richardson coefficients using vector partition functions are discussed. Practical implementations of Littlewood-Richardson calculators have been developed by Anders Buch and J. Stembridge.

$K_{\lambda\mu}$ is the multiplicity of the weight μ in the representation V_λ of the lie algebra $sl_{r+1}(\mathbb{C})$ of the special linear group having highest weight λ and $c_{\lambda\alpha}^{\nu}$ is the multiplicity of V_ν in the tensor product $V_\lambda \otimes_{\mathbb{C}} V_\alpha$. They

2000 *Mathematics Subject Classification.* Primary 05E10.

Key words and phrases. Complexity, Kostka numbers, LR coefficients, polynomial time.

also appear in the representation theory of the symmetric groups (see chapter 7, [**6**]). Schur polynomials form a linear basis of the ring of symmetric functions, and the Littlewood-Richardson coefficients appear in the multiplication rule,

$$s_\lambda \cdot s_\mu = \sum_\nu c^\nu_{\lambda\mu} s_\nu.$$

They also appear in the linear expansion of a skew Schur functionin terms of the Schur function basis,

$$s_{\nu/\lambda} = \sum_\mu c^\nu_{\lambda\mu} s_\mu.$$

While there are formulas for $K_{\lambda\mu}$ and $c^\nu_{\lambda\alpha}$ due to Kostant and Steinberg respectively ([**3**], [**2**]), the number of terms is, in general, exponential in the bit-length of the input. These numbers have interesting propoerties such as, for fixed $\lambda, \mu, \alpha, \nu, K_{N\lambda N\mu}$ and $c^{N\nu}_{N\lambda N\mu}$ are polynomials in N. These facts were established by Kirillov [**7**] and Derksen-Weyman [**4**] respectively.

Whether $K_{\lambda\mu} > 0$ can be answered in polynomial time (see Proposition 1), and so can the question of whether $c^\nu_{\lambda\alpha} > 0$, though the latter is a non-trivial fact established by K. Mulmuley and M. Sohoni [**10**], and uses the proof of the Saturation Conjecture by Knutson and Tao [**9**]. This fact plays an important role in the approach to the P vs NP question [**11**] due to K. Mulmuley and M. Sohoni.

We reduce the $\#P$-complete problem of finding the number $|\mathbb{I}(\mathbf{a}, \mathbf{b})|$ of $2 \times k$ contingency tables to that of finding some Kostka number $K_{\lambda\mu}$. Kostka numbers are known to be also Littlewood-Richardson (LR) coefficients. Thus, their computation reduces to computing some LR coefficient $c^\nu_{\lambda\alpha}$, where λ, μ, α and ν can be computed in time polynomial in the size of $(\mathbf{a}, \mathbf{b})$. The main tool used in the reduction to finding Kostka numbers is the R-S-K correspondence ([**6**] pages 40-41) between the set $\mathbb{I}(\mathbf{a}, \mathbf{b})$ of contingency tables and pairs of tableaux having contents $\mathbf{a}$ and $\mathbf{b}$ respectively.

2. Preliminaries and Notation

NP is the class of decision problems, $e : \cup_{n\in\mathbb{N}}\{0,1\}^n \to \{0,1\}$, for which there exists a polynomial time Turing machine M and a polynomial p such that $(\forall n \in \mathbb{N})$, $(\forall x \in \{0,1\}^n), e(x) = 1$ if and only if $\exists y, y \in \{0,1\}^{p(n)}$ such that M accepts $(x, y)\}$.

The class $\#P$ is the class of functions $f : \cup_{n\in\mathbb{N}}\{0,1\}^n \to \mathbb{Z}_{\geq 0}$, for which there exists a polynomial time Turing machine M and a polynomial p such that $(\forall n \in \mathbb{N}), (\forall x \in \{0,1\}^n), f(x) = |\{y \in \{0,1\}^{p(n)}$ such that M accepts $(x,y)\}|$. Valiant defined the counting class $\#P$ in his seminal paper [**15**]. Many counting problems are naturally in $\#P$. For example, counting the number of integer points in a polytope, membership queries to which can be answered in polynomial time is a problem in $\#P$.

A problem $W \in NP$ is NP-complete, if given a black box that solves instances of W in polynomial time, any problem in NP can be solved in polynomial time. Similarly, a counting problem $X \in \#P$ is $\#P$-complete if given a black box that provides solutions to instances of X in polynomial time, any problem in the class $\#P$ can be solved in polynomial time. Note that by definition, counting the number of solutions to any problem in NP is in $\#P$. Thus if a $\#P$-complete counting problem could be solved in polynomial time, we could find the number of solutions to any problem in NP efficiently (in polynomial time.) and thereby solve it, by checking if the number of solutions is ≥ 1.

The following problem of computing the number of $2\times k$ contingency tables is known to be $\#P$-complete. Let $\mathbf{a} = (a_1, a_2) \in \mathbb{Z}^2_{\geq 0}, a_1 \geq a_2$ and $\mathbf{b} = (b_1, \dots b_k) \in \mathbb{Z}^k_{\geq 0}$. We denote by $\mathbb{I}(a, b)$ the set of $2 \times k$ arrays of nonnegative integers whose row sums are a_1 and a_2 respectively and whose column sums are $b_1, \dots, b_k$. Geometrically, $\mathbb{I}(a, b)$ can be viewed as the set of integer points in the intersection of the multidimensional rectangular block defined by the column sums, and the diagonal hyperplane given by the first row sum. Counting $\mathbb{I}(a, b)$ was proved to be $\#P$-complete by R. Kannan, M. Dyer and J. Mount in [**5**].

A **Young diagram** ([**6**], page 1) is a collection of boxes, arranged in left justified rows, such that from top to bottom, the number of boxes in a row is monotonically (weakly) decreasing. The first two shapes in Figure 1 are Young diagrams. A **filling** is a numbering of the boxes of a Young diagram with positive integers, that are not necessarily distinct. A **Young tableau** or simply **tableau** is a filling such that the entries are

(1) weakly increasing from left to right across each row, and
(2) strictly increasing from top to bottom, down each column.

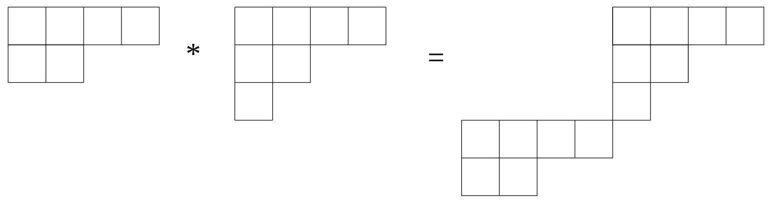


Figure 1. Left to right, the shapes λ, α and the skew shape $\lambda * \alpha$.

P and Q, in Figure 2, are Young tableaux. A **skew diagram** is the diagram obtained from removing a smaller Young diagram out of a larger one. The third shape in Figure 1 is a skew shape. A **skew tableau** is a filling of the boxes of a skew diagram with positive integers, non-decreasing in rows, and strictly increasing in columns (see Figure 5). Let $\lambda := (\lambda_1, \ldots, \lambda_s)$. If the number of boxes in the i^{th} row of a tableau, for $1 \leq i \leq s$ is λ_i, the tableau is said to have **shape** λ. If the tableau houses μ_j copies of j for $j \leq t$ and $\mu := (\mu_1, \ldots, \mu_t)$, it is said to have **content** μ. Thus, in Figure 2, P and Q have the same shape $(5, 2)$, but contents $(3, 2, 2)$ and $(4, 3)$ respectively.

Given two shapes λ and α, $\lambda * \alpha$ is defined to be the skew-shape obtained by attaching the lower left corner of α to the upper right corner of λ as in Figure 1 (see [**6**], page 60). $\mathbf{size}(\lambda, \mu)$ denotes the number of bits used in the description of this tuple of vectors. For $\lambda := (\lambda_1, \ldots, \lambda_s)$, let $|\lambda| = \sum_{i=1}^{s} \lambda_i$. For vectors λ, μ, we say that $\lambda \trianglerighteq \mu$ if $|\lambda| = |\mu|$ and $\forall i, \sum_{j \leq i} \lambda_j \geq \sum_{j \leq i} \mu_j$. In addition, if $\lambda \neq \mu$, we say $\lambda \triangleright \mu$. This ordering is called the dominance ordering.

We call a tableau **Littlewood-Richardson** or **LR**, if, when its entries are read right to left, top to bottom, at any moment, the number of copies of i encountered is greater than or equal to the number of copies of $i+1$ encountered ([**6**], page 63). We denote the set of all (possibly skew) tableaux of shape λ and content μ by $\mathbb{T}(\lambda, \mu)$, and its subset consisting of all LR (possibly skew) tableaux by $\mathrm{LR}\mathbb{T}(\lambda, \mu)$. The **Kostka number** $\mathbf{K}_{\lambda\mu}$ is the number of tableaux of shape λ and content μ, i.e $|\mathbb{T}(\lambda, \mu)|$ ([**6**], page 25). The **Littlewood-Richardson coefficient** $\mathbf{c}_{\lambda\alpha}^{\nu}$ is the number of LR skew tableaux of shape $\lambda * \alpha$ of content ν, i.e $|\mathrm{LR}\mathbb{T}(\lambda * \alpha, \nu)|$ (this follows from Corollary 2, (v), page 62 and Lemma 1, page 65 of [**6**]).

3. The problems are in $\#P$

The particular representation of partitions used above seems to be the most reasonable in the context of computing Kostka numbers and Littlewood-Richardson coefficients. The answer to whether or not a problem is in $\#P$ depends on the format in which the input is specified. If for example, we store partitions by their transposes, then these problems are no longer in the class $\#P$. This can be seen by considering the Kostka number equal to the number of standard tableaux on a $n \times 2$ rectangular array. By the hook length formula, the number of such tableaux is the Catalan number $\binom{2n}{n}/(n+1)$ which is exponential in n. However if the shape and content were represented as the transposes of the corresponding partitions, they occupy only $O(\log n)$ space. And so the Kostka number is doubly exponential in the size of the input. It is not hard to see that this is impossible for counting problems in the class $\#P$. On the other hand, if the partitions were represented in unary, it is not clear what the complexity of computing Kostka numbers and LR coefficients is. In unary, the partition $(3, 2, 1)$ would be represented as $(111, 11, 1)$. Thus unlike in the binary case, one cannot represent partitions with very large parts efficiently. It is clear that the problems are in $\#P$ for the unary case, but it is not clear whether they are $\#P$-complete.

The tableau shapes λ, α and contents μ, ν are described by vectors with integer coefficients. The Littlewood-Richardson coefficient number $c_{\lambda\alpha}^{\nu}$ counts the number of integer points of a polytope of dimension $O(\mathbf{size}(\lambda, \mu)^2)$, given by the intersection of $O(\mathbf{size}(\lambda, \mu)^2)$ halfspaces. The defining coefficients of these halfspaces have size $O(\mathbf{size}(\lambda, \mu))$. This follows from the encoding of relevant skew tableaux in the form of Littlewood-Richardson triangles (see [**12**].) Therefore the computation of Littlewood-Richardson coefficients is in $\#P$. The Kostka number $K_{\lambda\mu}$ is known to correspond to Littlewood-Richardson coefficients in parameters whose sizes are polynomial in $\mathbf{size}(\lambda, \mu)$. For the sake of completeness, an explicit correspondence has been established in Lemma 2. It follows that the problem of computing the Kostka number $K_{\lambda\mu}$ is in $\#P$.

Proposition 1. *Given λ and μ, whether or not $K_{\lambda\mu} > 0$ can be answered in polynomial time.*

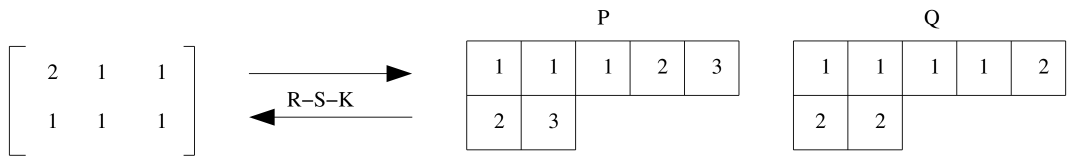


FIGURE 2. An instance of the correspondence between $\mathbb{I}(\mathbf{a},\mathbf{b})$ and $\cup_{\check\lambda}\mathbb{T}(\check\lambda,\mathbf{a})\times\mathbb{T}(\check\lambda,\mathbf{b})$ for $\mathbf{a}=(4,3)$, $\mathbf{b}=(3,2,2)$.

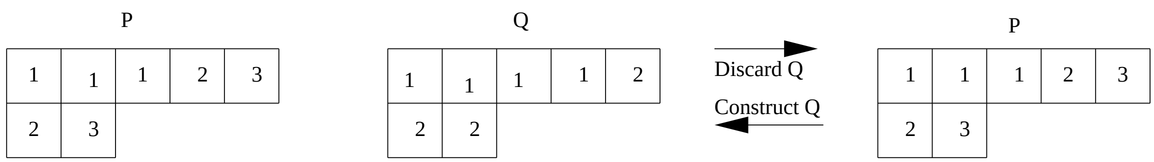


FIGURE 3. An instance of the correspondence between $\cup_{\check\lambda}\mathbb{T}(\check\lambda,\mathbf{a})\times\mathbb{T}(\check\lambda,\mathbf{b})$ and $\cup_{\check\lambda\trianglerighteq\mathbf{a}}\mathbb{T}(\check\lambda,\mathbf{b})$ for $\mathbf{a}=(4,3)$ and $\mathbf{b}=(3,2,2)$.

Proof:
Let λ,μ be defined as in section 1. For any permutation σ of the set $\{1,\ldots,t\}$, let $\sigma(\mu)$ be the vector $(\mu_{\sigma(1)},\ldots,\mu_{\sigma(t)})$. It is a known fact that $K_{\lambda\mu}=K_{\lambda\sigma(\mu)}$ (see [**6**], page 26). Let σ be a permutation such that $\forall i\le t-1$, $\mu_{\sigma(i)}\ge\mu_{\sigma(i+1)}$. For any $\breve\mu$, whose components are arranged in non-increasing order, it is known that $K_{\lambda\breve\mu}>0$ if and only if $\lambda\trianglerighteq\breve\mu$ (see [**6**], page 26). Whether $\lambda\trianglerighteq\sigma(\mu)$ can be checked in time that is $O(\mathbf{size}(\lambda,\mu))$. Thus, whether or not $K_{\lambda\mu}>0$ can be answered in time $O(\mathbf{size}(\lambda,\mu)\ln(\mathbf{size}(\lambda,\mu))$, which is the time it takes to find a permutation σ that arranges the components of μ in non-increasing order.
□

4. Hardness Results

LEMMA 1. *Given* $\mathbf{a}=(a_1,a_2)\in\mathbb{Z}^2_{\ge 0}$, $a_1\ge a_2$, *and* $\mathbf{b}=(b_1,\ldots,b_k)\in\mathbb{Z}^k_{\ge 0}$, *let* $\lambda=(a_1+a_2,a_2)$ *and* $\mu=(b_1,\ldots,b_k,a_2)$. *Then,* $|\mathbb{I}(\mathbf{a},\mathbf{b})|=K_{\lambda\mu}$.

Proof:

The R-S-K (Robinson-Schensted-Knuth) correspondence ([**6**] pages 40-41) gives a bijection between $\mathbb{I}(\mathbf{a},\mathbf{b})$, the set of $2\times k$ contingency tables with row sums $\mathbf{a}$ and column sums $\mathbf{b}$, and pairs of tableaux (T_1,T_2) having a common shape but contents $\mathbf{a}$ and $\mathbf{b}$ respectively. In other words, we have a bijection between $\mathbb{I}(\mathbf{a},\mathbf{b})$ and $\cup_{\check\lambda}\mathbb{T}(\check\lambda,\mathbf{a})\times\mathbb{T}(\check\lambda,\mathbf{b})$. A sample correspondence is shown in Figure 2.

CLAIM 1. *For every shape* $\check\lambda=(\check\lambda_1,\check\lambda_2)$, *such that that* $\check\lambda\trianglerighteq\mathbf{a}$, *there is exactly one tableau having shape* $\check\lambda$ *and content* $\mathbf{a}$. *For any other shape* $\check\lambda$ *there is no tableau having shape* $\check\lambda$ *and content* $\mathbf{a}$.

It follows from the proof of Proposition 1 that the existence of a tableau with shape $\check\lambda$ and content $\mathbf{a}$ is equivalent to the condition $\check\lambda\trianglerighteq\mathbf{a}$. Any tableau with content $\mathbf{a}=(a_1,a_2)$ can have at most two rows, since the entries in a single column are all distinct. The filling in which the first a_1 boxes of the top row contain 1 and all others contain 2 is a tableau (see Q in Figure 3). Since all the copies of 1 must be in the first row and must be in a contiguous stretch including the leftmost box, this is the only tableau in $\mathbb{T}(\lambda,\mathbf{a})$. Hence the claim is proved.

Thus there is a bijection between $\cup_{\check\lambda}\mathbb{T}(\check\lambda,\mathbf{a})\times\mathbb{T}(\check\lambda,\mathbf{b})$ and the set of tableaux of content $\mathbf{b}$ having some shape $\check\lambda\trianglerighteq\mathbf{a}$. i.e, there is a bijection between $\cup_{\check\lambda}\mathbb{T}(\check\lambda,\mathbf{a})\times\mathbb{T}(\check\lambda,\mathbf{b})$ and $\cup_{\check\lambda\trianglerighteq\mathbf{a}}\mathbb{T}(\check\lambda,\mathbf{b})$. An example of this is provided in Figure 3. Let us now consider the set $\cup_{\check\lambda\trianglerighteq\mathbf{a}}\mathbb{T}(\check\lambda,\mathbf{b})$.

CLAIM 2. *Any tableau in* $\cup_{\check\lambda\trianglerighteq\mathbf{a}}\mathbb{T}(\check\lambda,\mathbf{b})$ *can be extended to a tableau of the shape* $\lambda=(a_1+a_2,a_2)$ *by filling the boxes that are in* λ *but not* $\check\lambda$, *with* $k+1$. *This extension is a bijection between* $\cup_{\check\lambda\trianglerighteq\mathbf{a}}\mathbb{T}(\check\lambda,\mathbf{b})$ *and* $\mathbb{T}(\lambda,\mu)$.

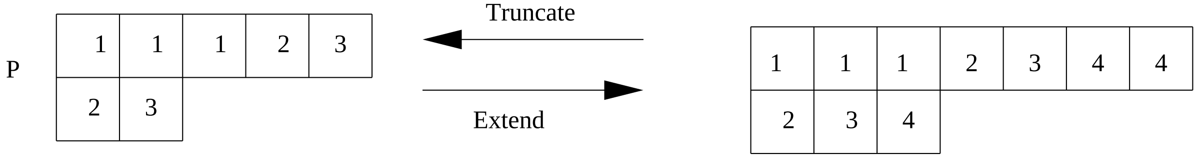


FIGURE 4. An instance of the correspondence between $\cup_{\check{\lambda} \trianglerighteq \mathbf{a}} \mathbb{T}(\check{\lambda}, \mathbf{b})$ and $\mathbb{T}(\lambda, \mu)$, where $a = (4, 3), b = (3, 2, 2), \lambda = (7, 3)$ and $\mu = (3, 2, 2, 3)$.

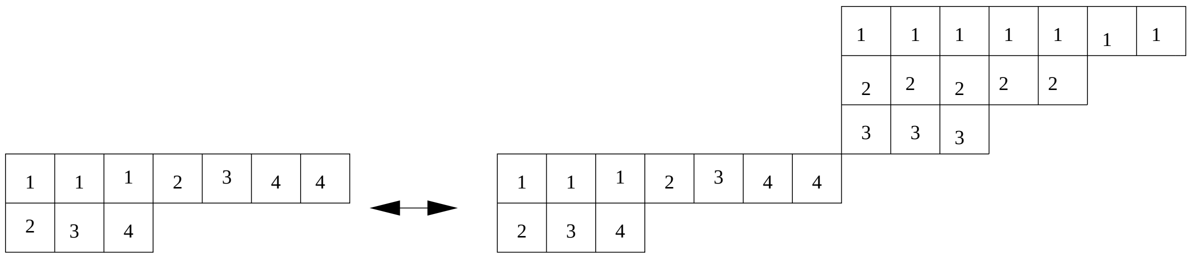


FIGURE 5. An instance of the correspondence between $\mathbb{T}(\lambda, \mu)$ and $\mathrm{LR}\mathbb{T}(\lambda * \alpha, \nu)$ for $\lambda = (7, 3)$ and $\mu = (3, 2, 2, 3)$, $\alpha = (7, 5, 3)$ and $\nu = (10, 7, 5, 3)$.

If there is a tableau of shape $\check{\lambda}$ and content $\mathbf{a}$, $\check{\lambda}_1 \leq a_1 + a_2$, and $\check{\lambda}_2 \leq a_2$. $\check{\lambda} \trianglerighteq \mathbf{a} \implies \check{\lambda}_1 \geq a_2 = \lambda_2$. Therefore no two of the boxes in λ which are not in $\check{\lambda}$ belong to the same column. Those of these boxes, that are present in a given row, occupy a contiguous stretch that includes the rightmost box. Therefore by filling them with $k+1$ we get a tableau in $\mathbb{T}(\lambda, \mu)$. Conversely, given a tableau T in $\mathbb{T}(\lambda, \mu)$, deleting all boxes of T filled with $k+1$ gives a tableau in $\cup_{\check{\lambda} \trianglerighteq \mathbf{a}} T(\check{\lambda}, \mathbf{b})$. These two maps are inverses of each other and hence provide a bijection between $\cup_{\check{\lambda} \trianglerighteq \mathbf{a}} T(\check{\lambda}, \mathbf{b})$ and $\mathbb{T}(\lambda, \mu)$. Hence the claim is proved.

An example of this correspondence has been illustrated in Figure 4. Therefore, $|\mathbb{I}(\mathbf{a}, \mathbf{b})| = |\cup_{\check{\lambda}} \mathbb{T}(\check{\lambda}, \mathbf{a}) \times \mathbb{T}(\check{\lambda}, \mathbf{b})| = |\cup_{\check{\lambda} \trianglerighteq \mathbf{a}} \mathbb{T}(\check{\lambda}, \mathbf{b})| = |\mathbb{T}(\lambda, \mu)| = K_{\lambda\mu}$.
□

THEOREM 1. *The problem of computing $K_{\lambda\mu}$, even when λ has only 2 rows, is #P-complete.*

Proof:
Computing $K_{\lambda\mu}$ is in #P as shown in Section 3. Now the result follows from Lemma 1 because the computation of $|\mathbb{I}(\mathbf{a}, \mathbf{b})|$ is known to be #P-complete ([**5**]).
□

LEMMA 2. *Given $\lambda = (\lambda_1, \lambda_2) \in \mathbb{Z}^2_{\geq 0}$, $\lambda_1 \geq \lambda_2$, and $\mu = (\mu_1, \dots, \mu_\ell) \in \mathbb{Z}^\ell_{\geq 0}$, let $\alpha = (\alpha_1, \dots, \alpha_{\ell-1})$ where $(\forall i)\alpha_i = \sum_{j>i} \mu_i$, and $\nu = (\nu_1, \dots, \nu_\ell)$, where $\forall i \leq \ell - 1, \nu_i = \alpha_i + \mu_i$, and $\nu_\ell = \mu_\ell$. Then $K_{\lambda\mu} = c^\nu_{\lambda\alpha}$.*

Proof:
$c^\nu_{\lambda\alpha}$ is, by definition, $|\mathrm{LR}\mathbb{T}(\lambda * \alpha, \nu)|$, which is the number of LR tableaux on the skew shape $\lambda * \alpha$ that have content ν. The skew shape $\lambda * \alpha$ consists of a copy of λ and a copy of α, as in Figures 1 and 5. For any skew tableau S of shape $\lambda * \alpha$, we shall denote by $S|_\alpha$, the restriction of S to the copy of α and by $S|_\lambda$, the restriction of S to the copy of λ. Thus, $S|_\alpha$ is a tableau of shape α and $S|_\lambda$ is a tableau of shape λ.

Let $S \in \mathrm{LR}\mathbb{T}(\lambda * \alpha, \nu)$. For $i \leq \ell - 1$, it follows from the LR and tableau constraints that the i^{th} row of $S|_\alpha$ must consist entirely of copies of i.

Consequently, $S|_\lambda$ must have content $\nu - \alpha = \mu$. In other words, $S|_\lambda \in \mathbb{T}(\lambda, \mu)$. Conversely, given any tableau $T \in \mathbb{T}(\lambda, \mu)$, let $S(T)$ be the skew tableau of shape $\lambda * \alpha$ in which $S(T)|_\lambda = T$ and the i^{th} row of $S(T)|_\alpha$ consists entirely of copies of i. It is not difficult to see that $S(T) \in \mathrm{LR}\mathbb{T}(\lambda * \alpha, \nu)$. $S(T)|_\lambda = T$, thus we have a bijection between $\mathrm{LR}\mathbb{T}(\lambda * \alpha, \nu)$, the set of LR skew tableaux of shape $\lambda * \alpha$ having content ν and $\mathbb{T}(\lambda, \mu)$, the set of tableaux of shape λ having content μ. Hence $K_{\lambda\mu} = |\mathbb{T}(\lambda, \mu)| = |\mathrm{LR}\mathbb{T}(\lambda * \alpha, \nu)| = c^\nu_{\lambda\alpha}$ as claimed.
□

THEOREM 2. *The problem of computing $c^\nu_{\lambda\alpha}$, even when λ has only 2 rows is #P-complete.*

Proof:
By the explanation in Section 3, computing $c^{\nu}_{\lambda\alpha}$ is in $\#P$. We have already proved in Theorem 1, that the computation of $K_{\lambda\mu}$ is $\#P$-complete. The result now follows from Lemma 2.
□

5. Conclusion

We proved that the computation of Kostka numbers and Littlewood-Richardson coefficients is $\#P$-complete. The reduction to computing Kostka numbers was from the $\#P$-complete problem [**5**] of computing the number of contingency tables having given row and column sums. The problem of computing Kostka numbers was then reduced to that of computing Littlewood-Richardson coefficients. FPRAS (Fully Polynomial Randomized Approximation Schemes) are known to exist for contingency tables with two rows. Thus we obtain FPRAS for a restricted class of Kostka numbers from the correspondence in Lemma 1. It would be of interest to know if such schemes exist for Kostka numbers and Littlewood-Richardson coefficients with general parameters.

6. Acknowledgements

I wish express my gratitude to Ketan Mulmuley for suggesting the topic of this paper and for many valuable discussions. Many thanks are due to Ravi Kannan for informing me about [**5**]. I also sincerely thank Etienne Rassart, Rahul Santhanam, Lászlό Babai and the anonymous referee for helpful comments.

References

[1] A. Barvinok and S.V. Fomin, *Sparse interpolation of symmetric polynomials*, Advances in Applied Mathematics , 18 (1997), 271-285, MR 98i:05164.
[2] S. Billey, V. Guillemin, E. Rassart, *A vector partition function for the multiplicities of* $sl_k(\mathbb{C})$, Journal of Algebra, 278 (2004) no. 1, 251-293.
[3] C. Cochet, *Kostka Numbers and Littlewood-Richardson Coefficients*, preprint (2003).
[4] H. Derksen and J. Weyman. On the Littlewood-Richardson polynomials, *J. Algebra*, 255(2):247-257, 2002.
[5] M. Dyer, R. Kannan, J. Mount, *Sampling Contingency tables*, Random Structures and Algorithms, (1997) 10 487-506.
[6] W. Fulton, *Young Tableaux*, London Mathematical Society Student Texts 35 (1997).
[7] Anatol N. Kitillov. Ubiquity of Kostka polynomials. In *Physics and Combinatorics 1999*, pages 85-200 World Scientific, Singapore, 2001. Proceedings of the Nagoya 1999 International Workshop on Physics and Combinatorics, Nagoya University, August 23-27, 1999. arXiv:math.QA/9912094.
[8] D. E. Knuth, *Permutations, matrices, and generalized Young tableaux*, Pacific Journal of Mathematics, vol. 34, (1970), pp. 709–727.
[9] A. Knutson, T. Tao, *The honeycomb model of tensor products I: Proof of the saturation conjecture*, J. Amer. Math. Soc. 12 (1999), 1055-1090.
[10] K. Mulmuley, M, Sohoni, *Geometric Complexity III: on deciding positivity of Littlewood-Richardson coefficients*, http://arxiv.org/abs/cs.CC/0501076
[11] K. Mulmuley, M. Sohoni, *Geometric complexity theory, P vs. NP, and explicit obstructions.*, Proceedings, International Conference on Algebra and Geometry, Hyderabad, (2001).
[12] I. Pak, E. Vallejo, *Combinatorics and geometry of Littlewood-Richardson cones*, Europ. J. Combinatorics, vol. 26 (2005), 995-1008.
[13] E. Rassart, *Geometric approaches to computing Kostka numbers and Littlewood-Richardson coefficients*, Thesis for the Ph.D. degree in Mathematics, Massachusetts Institute of Technology (MIT), 2004.
[14] E. Rassart, *A polynomiality property for Littlewood-Richardson coefficients*, Journal of Combinatorial Theory, Series A 107 (2004), no. 2, 161–179.
[15] L.G. Valiant, *The complexity of computing the permanent*, Theoret. Comp. Sci., (1979), 8:189-201.
[16] M. A. A. van Leeuwen, *The Littlewood-Richardson rule, and related combinatorics*, Math. Soc. of Japan Memoirs 11, Interaction of Combinatorics and Representation Theory; arXiv:math.CO/9908099.
[17] E. Wigner. *On the consequences of the symmetry of the nuclear Hamiltonian on the spectroscopy of the nuclei*, Physical Review, (1937), 51:106-119.

Department of Computer Science, University of Chicago
E-mail address: `hari@cs.uchicago.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Dual Graded Graphs and Fomin's r-correspondences associated to the Hopf Algebras of Planar Binary Trees, Quasi-symmetric Functions and Noncommutative Symmetric Functions (EXTENDED ABSTRACT)

Janvier Nzeutchap

Abstract. Fomin (1994) introduced a notion of duality between two graded graphs on the same set of vertices. By a construction similar to the plactic monoid, Hivert, Novelli and Thibon (2001) introduced a monoid structure on the set of binary search trees, the Robinson-Schensted insertion algorithm being replaced by the binary search tree insertion algorithm. Using this monoid they gave a new construction of the algebra of Planar Binary Trees of Loday-Ronco. In this construction, one can build pairs of graded graphs of which we study the duality as in Fomin's setting. We observe that the sylvester congruence defining this algebra is in fact an r-correspondence as defined by Fomin. We also observe graph duality in the algebras of noncommutative symmetric functions, and quasi symmetric functions, and we identify an r-correspondence of two graded graphs built in these algebras, with the hypoplactic congruence introduced by Krob and Thibon (1997). We also present a combinatorial description of the Schensted-Fomin algorithm for dual graded graphs and we use this description to give a proof of a bijection between pairs of paths in any pair of dual graded graphs and permutations of the symmetric group. We conclude with the statement of a possible connection between graded graphs duality and the construction of dual Hopf algebras.

Résumé. Fomin (1994) a introduit une notion de dualité entre graphes gradués. D'autre part, par une construction analogue à celle définissant le monoïde plaxique, Hivert, Novelli et Thibon (2001) introduisent une structure de monoïde sur l'ensemble des arbres binaires de recherche, la congruence plaxique étant remplacée dans cette construction par l'algorithme d'insertion dans un arbre binaire de recherche, encore appelé congruence sylvestre par ces auteurs. Cette construction donne lieu à une nouvelle réalisation de l'algèbre de Hopf des arbres binaires de Loday-Ronco. Dans cette algèbre, il est possible de construire des paires de graphes gradués dont nous étudions la dualité au moyen d'un isomorphisme avec des graphes définis par Fomin. Nous identifions par la suite la congruence sylvestre à une r-correspondance que nous définissons dans ces graphes. La congruence hypoplaxique introduite par Krob et Thibon (1997) est quant à elle identifiée à une r-correspondance définie sur des graphes en dualité dans les algèbres des fonctions symétriques non commutatives et des fonctions quasi-symétriques. Nous donnons aussi une description combinatoire de l'algorithme de Schensted-Fomin pour les graphes gradués en dualité, et nous l'utilisons pour faire une preuve d'une bijection entre paires de chemins dans de tels graphes, et les permutations du groupe symétrique. Compte tenu du mode de construction des graphes étudiés, nous concluons par une possible relation entre graphes gradués en dualité et construction d'algèbres de Hopf duales.

1. Introduction and preliminary definitions

The Young lattice is defined on the set of partitions of integers, with covering relations given by the natural inclusion order. This lattice is associated to the operation of multiplication of Schur functions [**6**] s_λ by s_1, where there is an edge connecting λ and μ if s_μ appears with a nonzero coefficient in the expansion of $s_1 s_\lambda$. The distributive lattice nature of this graph was generalized by S. Fomin (1994) with the introduction of graph duality [**7**]. With this extension he introduced a generalization of the classical Robinson-Schensted algorithm, giving a general scheme for establishing bijective correspondences between pairs of paths in dual

2000 *Mathematics Subject Classification.* Primary 05-06; Secondary 05E99.
Key words and phrases. graded graph, duality, Schensted-Fomin algorithm, Hopf algebra.

graded graphs both starting at a vertex of rank 0 and having a common end point of rank n, on the one hand, and permutations of the symmetric group $\mathfrak{S}_n$ on the other hand.

Later, Krob and Thibon (1997), Hivert, Novelli and Thibon (2001) showed that using two congruence relations on words, namely the hypoplactic congruence [**2**] and the binary search tree insertion algorithm [**1, 3**], one can realize as polynomials (commutative or not), two pairs of dual Hopf algebras. The first pair of algebras is the dual pair formed by the algebra of quasi-symmetric functions ($QSym$) and the algebra of noncommutative symmetric functions (**Sym**). The second is made of the algebra of planar binary trees (**PBT**) of Loday-Ronco [**5**] and its dual (**PBT***) which is isomorphic to itself. In those algebras one builds pairs of graded graphs analogue to the Young lattice, and associated to the operations of multiplication. One aim of this paper is first to prove the duality of these pairs of graphs, next to find natural r-correspondences [**7**] associated to those graphs, and last to show that these correspondences convert the Schensted-Fomin algorithm for dual graded graphs into a parallel version of the hypoplactic insertion algorithm and sylvester insertion algorithm respectively.

The paper is organized as follows: we first recall definitions related to graph duality, and then we examine the case of **PBT** in the second section, **Sym** and $QSym$ are treated in the next section. In the fourth section we present a combinatorial description of the Schensted-Fomin algorithm for dual graded graphs. We also use this algorithmic description to give a proof of the bijection between pairs of paths in dual graded graphs and permutations of the symmetric group. Last we apply this algorithm to the graphs of the previous sections. Now let us introduce graph duality.

Definition 1.1. A graded graph [**7**] is a triple $G = (P, \rho, E)$ where P is a discrete set of vertices, $\rho : P \to \mathbb{Z}$ is a rank function and E is a multi-set of edges (x, y) satisfying $\rho\,(y) = \rho\,(x) + 1$.

Let $G_1 = (P, \rho, E_1)$ and $G_2 = (P, \rho, E_2)$ be a pair of graded graphs with a common set of vertices and a common rank function.

Definition 1.2. An oriented graded graph [**7**] $G = (P, \rho, E_1, E_2)$ is defined by directing the G_1-edges, E_1 up (in the direction of increasing rank) and the G_2-edges, E_2 down (in the direction of decreasing rank).

Let $G = (P, \rho, E_1, E_2)$ be an oriented graded graph and $\mathbb{K}$ a field of characteristic zero, define $\mathbb{K}P$ as the vector space formed by linear combinations of vertices of P. One can now define two linear operators U (Up) and D (Down) acting on $\mathbb{K}P$ as follows:

$$U\,x \;=\; \sum_{(x,y)\,\in\,E_1} m_1(x,y)\;y \qquad ; \qquad D\,y \;=\; \sum_{(x,y)\,\in\,E_2} m_2(x,y)\;x \tag{1.1}$$

where $m_i(x, y)$ is the multiplicity or the weight of the edge (x, y) in E_i.

Definition 1.3. G_1 and G_2 are said to be dual [**7**] if U and D satisfy the commutation relation:

$$D_{n+1}\,U_n = U_{n-1}\,D_n + I_n \tag{1.2}$$

where U_n (resp. D_n) denote the restriction of the operator U (resp. D) to the n^{th} level of the graph, and I_n the identical operator at the same level.

Generalizations of this definition are also found in [**7**], notably the case of an r-duality with $r > 1$ where the commutation relation generalizes to:

$$D_{n+1}\,U_n = U_{n-1}\,D_n + r\,I_n \quad \text{and} \quad D_{n+1}\,U_n = U_{n-1}\,D_n + r_n\,I_n \tag{1.3}$$

A well-known example of a graded graph is the Young lattice of partitions of integers, which describes the multiplication of Schur functions s_λ by s_1 (Fig. 1). This is a first and natural example of graph duality in relation with the operation of multiplication in two dual Hopf algebras. In fact, the Young lattice is a self-dual graded graph or distributive lattice. Its duality expresses the fact that for any partition λ, there is one more partition obtained by adding a single part to λ than by deleting a single part from λ, and for two partitions λ and μ there are as many partitions simultaneously contained by λ and μ than those simultaneously containing λ and μ. On the other hand, the collection of Schur functions span a self-dual Hopf algebra, that is the algebra of symmetric functions [**6**]. So the self-dual Hopf algebra of symmetric functions is described by the Young lattice which is a self-dual graded graph.

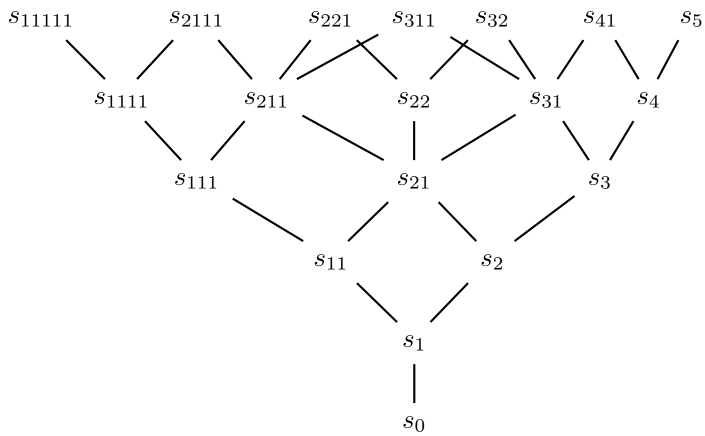


FIGURE 1. The Young graph: multiplication of Schur functions s_λ by s_1

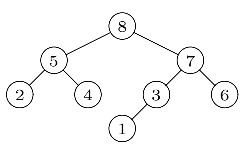
2. Dual graded graphs in PBT

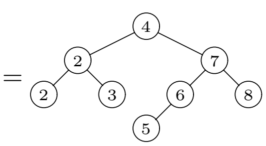
This section is devoted to **PBT**, the Hopf algebra of planar binary trees, for which Loday and Ronco [**5**] gave an explicit embedding as a subalgebra of the convolution algebra of permutations, via the construction of the decreasing tree of a permutation. After recalling the construction of this algebra, we will describe a second example of graph duality in relation with the operation of multiplication in two dual Hopf algebras. In all that follows, we will be considering only words on a totally ordered alphabet, for instance $A = \{1, 2, 3, \cdots\}$.

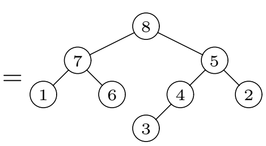
2.1. Definitions.

DEFINITION 2.1. A decreasing tree $\mathcal{T}$ is a labeled binary tree such that the label of each internal node is greater than the labels of all the nodes in its subtrees.

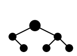
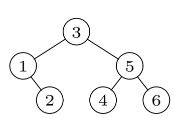
Let w be a word with no repetition of letters. Its *decreasing tree* $\mathcal{T}(w)$ is obtained as follows: its root is labeled with the greatest letter n of w, and if $w = u\,n\,v$, where u and v are words with no repetition of letters, then the left subtree of $\mathcal{T}(w)$ is $\mathcal{T}(u)$ and its right subtree is $\mathcal{T}(v)$. Another tree associated to a word w is its right strict binary search tree, this is a labeled binary tree labeled with w's letters such that for each internal node, its label is greater or equal to the labels of the nodes in its left subtree and strictly smaller than the labels of the nodes in it's right subtree. The binary search tree associated to a word w will be denoted $\mathcal{P}(w)$. It is obtained by applying the well-known binary search tree insertion algorithm [**1**] to w, but *reading w from right to left.* During this insertion process one can use a second tree denoted $\mathcal{Q}(w)$ to record the positions in w of the letters inserted at each step. $\mathcal{Q}(w)$ coincides with $\mathcal{T}(std(w)^{-1})$ where $std(w)$ is the standardized word of w. The user not familiar with the standardization process may consult [**3**] for definitions. For example let us consider the two words $w_1 = 25481376$ and $w_2 = 28567324$, then we have:

$\mathcal{T}(w_1) =$ 

; $\mathcal{P}(w_2) =$ 

; $\mathcal{Q}(w_2) =$ 

The map $w \mapsto (\mathcal{P}(w), \mathcal{Q}(w))$ is known as the *sylvester correspondence* and is associated to a congruence, the sylvester congruence, defined on words on the alphabet A by: $u \equiv_{syv} v \Leftrightarrow \mathcal{P}(u) = \mathcal{P}(v)$. See [**3**] for a plactic-like characterization on words. The *sylvester canonical permutation* associated to an unlabeled binary tree T is the right-to-left postfix reading of the only binary search tree that is the left-to-right infix labeling of T. The sylvester canonical permutation of a permutation σ is the right-to-left postfix reading of $\mathcal{P}(\sigma)$. For example let us consider:

$T =$ 
; the labeling is 

and the sylvester canonical permutation is $\sigma_T = 645213$.

645213 is also the sylvester canonical permutation of the permutation 465213. Now let us recall the definition of the algebra of free quasi-symmetric functions. This definition is needed to introduce **PBT**.

DEFINITION 2.2. Let σ be a permutation. The *Free Quasi-Ribbon* $\mathbb{F}_\sigma$ is the noncommutative polynomial

$$\mathbb{F}_\sigma \quad = \sum_{w \,:\, std\,(w) \,=\, \sigma^{\,-1}} w$$

where $std(w)$ denotes the standardized word of w, and w runs over the words on the alphabet A. The free quasi-ribbons span a subalgebra of the free associative algebra. This subalgebra is the algebra of free quasi-symmetric functions (**FQSym**), and its multiplication rule is the following:

$$\mathbb{F}_\alpha \, \mathbb{F}_\beta = \sum_{\sigma \,\in\, (\,\alpha \,\sqcup\!\sqcup\, \beta\,[|\alpha|]\,)} \mathbb{F}_\sigma$$

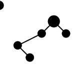
where $\alpha \sqcup\!\sqcup \beta\,[|\alpha|]$ is the shifted shuffle of the two permutations α and β. The user not familiar with the standardization process and shuffles may consult [**3**] for definitions. The dual basis of the $\mathbb{F}_\sigma$ are the $\mathbb{G}_\sigma$ defined by:

$$\mathbb{G}_\sigma \quad = \quad \mathbb{F}_{\sigma^{-1}} \quad = \sum_{w \,:\, std\,(w) \,=\, \sigma} w$$

An embedding of **PBT** in **FQSym** is given as the linear span of the $(\mathbf{P}_T)$ defined [**3**] by:

$$\mathbf{P}_T \quad = \sum_{w \,:\, shape\,(\mathcal{T}(std(w))) \,=\, T} w \quad = \sum_{\sigma \,:\, shape\,(\mathcal{P}(\sigma)) \,=\, T} \mathbb{F}_\sigma \tag{2.1}$$

where T is an unlabeled binary tree, σ a permutation, the shape of a labeled tree being the corresponding unlabeled tree. For example:

$\mathbf{P}$  $= \mathbf{P}_{52134} = \mathbb{F}_{21354} + \mathbb{F}_{21534} + \mathbb{F}_{25134} + \mathbb{F}_{52134}$

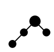
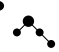
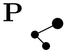
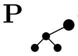
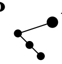
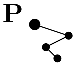
The multiplication rule in **PBT** is given by:

$$\mathbf{P}_{T_1}\,\mathbf{P}_{T_2} \quad = \sum_{T \,\in\, \text{shuffle}\,(T_1,T_2)} \mathbf{P}_T \tag{2.2}$$

where shuffle(T_1, T_2) is the set of unlabeled binary trees whose canonical sylvester permutations appear in $\sigma_1 \sqcup\!\sqcup \sigma_2\,[|\sigma_1|]$, σ_i being the canonical sylvester permutations associated to T_i. For example:

$$12 \sqcup\!\sqcup 21[2] \;=\; 12 \sqcup\!\sqcup 43 \;=\; (1243 + 1423 + 4123) \quad + \quad (1432 + 4132 + 4312)$$

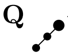
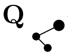
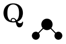
so we will have:

$\mathbf{P}$  $\mathbf{P}$  $=$ $\mathbf{P}$  $+$ $\mathbf{P}$  ; and one can also check that $\mathbf{P}_\bullet$ $\mathbf{P}$  $=$ $\mathbf{P}$  $+$ $\mathbf{P}$  $+$ $\mathbf{P}$ 

The dual basis of the $(\mathbf{P}_T)$ are the $(\mathbf{Q}_T)$ defined by $\mathbf{Q}_T = \pi\,(\mathbf{G}_{\sigma_T})$ where $\pi \;:\; \mathbb{C}\langle A\rangle \longrightarrow \mathbb{C}\langle A\rangle / \equiv_{sylv}$ is the canonical projection sending a sum of permutations to the sum of the corresponding sylvester canonical permutations. The multiplication rule in $\mathbf{PBT}^*$ is given by:

$$\mathbf{Q}_{T_1}\,\mathbf{Q}_{T_2} \quad = \sum_{T \,\in\, \text{Conv}\,(T_1,T_2)} \mathbf{Q}_T$$

where Conv$\,(T_1, T_2)$ is defined as follows: let σ_i be the canonical sylvester permutation associated to T_i, then Conv$\,(T_1, T_2)$ is the set of unlabeled binary trees whose canonical sylvester permutations appear in the convolution product $\mathbf{G}_{\sigma_1}\mathbf{G}_{\sigma_2}$. For example,

$\mathbb{G}_1\mathbb{G}_{12} = \mathbb{G}_{123} + \mathbb{G}_{213} + \mathbb{G}_{312}$, so one will have $\mathbf{Q}_\bullet\mathbf{Q}$  $=$ $\mathbf{Q}$  $+$ $\mathbf{Q}$  $+$ $\mathbf{Q}$ 

Using the multiplication rules in **PBT** and $\mathbf{PBT}^*$, it is possible to build a pair of graded graphs (Fig. 2 and Fig. 3) whose set of vertices of degree n are the binary trees of size n. In those graphs, there is an edge between T and T' if T' appears in the product $\mathbf{P}_\bullet\,\mathbf{P}_T$ (resp. $\mathbf{Q}_\bullet\,\mathbf{Q}_T$), where $\bullet$ is the tree of size 1. All edges are weighted 1 since there are no multiplicities in the products in **PBT** and $\mathbf{PBT}^*$. A second pair of graphs describes the right multiplication by $\mathbf{P}_\bullet$ and $\mathbf{Q}_\bullet$ respectively. See [**3**] for related figures.

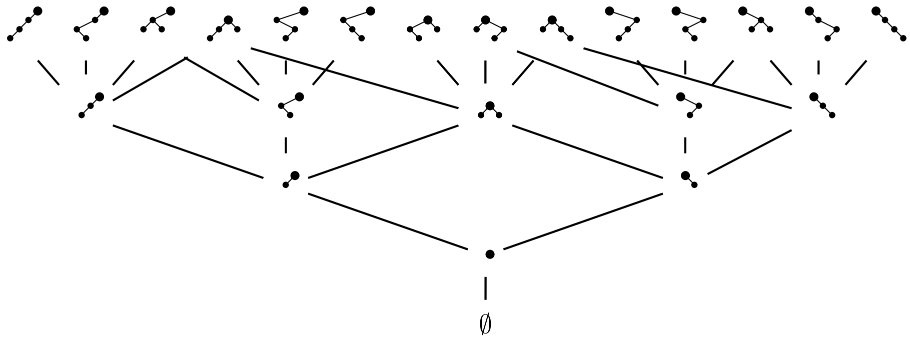


FIGURE 2. $\Gamma^{left}_{Q_\bullet}$: left multiplication by $\mathbf{Q}_\bullet$ in **PBT**

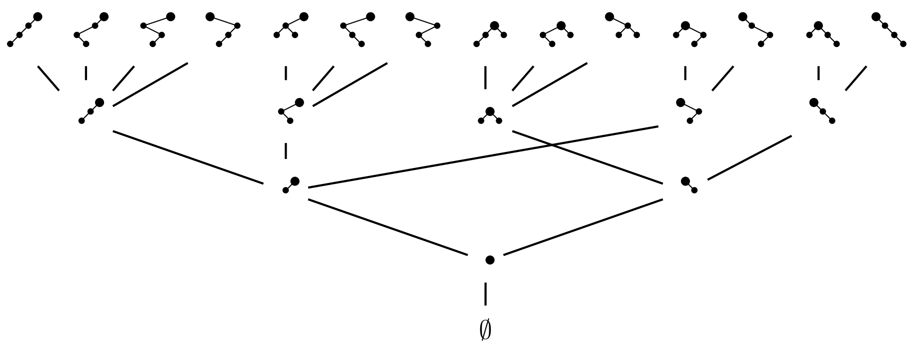


FIGURE 3. $\Gamma^{left}_{P_\bullet}$: left multiplication by $\mathbf{P}_\bullet$ in **PBT**

2.2. Graph's duality.

It was already stated in [**3**], but without proof, that the graphs $\Gamma^{left}_{Q_\bullet}$ and $\Gamma^{left}_{P_\bullet}$ above could be in duality. We prove this using the fact that they are isomorphic to two dual graded graphs studied by Fomin [**7**] and known as the lattice of binary trees and the bracket tree (Fig. 4 and Fig. 5). The lattice of binary trees is defined as follows: it's vertices of rank n are the syntactically correct formulae defining different versions of calculation of a non-associative product of $n+1$ entries. So any vertex of rank n is a valid sequence of $n-1$ opening and $n-1$ closing brackets inserted into $x_1 . x_2 \cdots x_n$. In the bracket tree, two vertices are linked if one results from the other by deleting the first entry, and then removing subsequent unnecessary brackets, and renumbering the new expression.

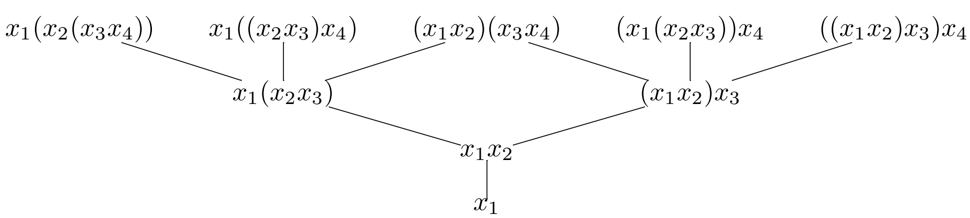


FIGURE 4. The lattice of binary trees

REMARK 2.3. There is a one-to-one correspondence between unlabeled binary trees and bracketed expressions. In this correspondence, an unlabeled binary tree is identified with the expression obtained by completing the tree, adding one leaf to any node having a single child-node, and two leaves to any childless node. Then label the leaves of the resulting *complete* unlabeled binary tree according to the left-to-right infix order. If the final tree is empty then the expression is x_1, or else the expression is obtained by recursively reading its left and right subtrees in that order.

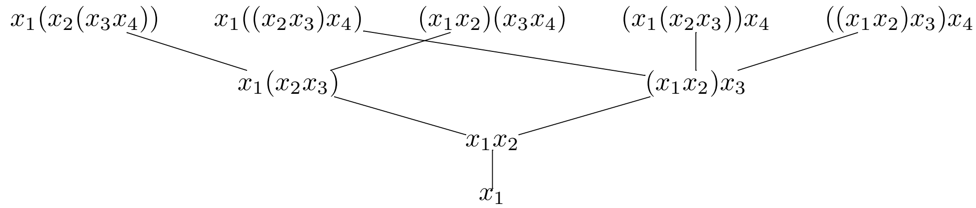
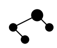
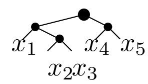


FIGURE 5. The bracket tree, dual of the lattice of binary trees

An example to illustrate the correspondence described in remark 2.3 is the following:

$T =$  ; the labeling is 

and the expression: $(x_1(x_2x_3))(x_4x_5)$

PROPOSITION 2.1. *$\Gamma_{Q_\bullet}^{left}$ and $\Gamma_{P_\bullet}^{left}$ are respectively isomorphic to the lattice of binary trees and it's dual, the bracket tree.*

The proof of **Proposition 2.1** is made using a combinatorial description of the covering relations in the lattice of binary trees and in the bracket tree, identifying each bracketed expression with an unlabeled binary tree as described in **Remark 2.3**. These relations are:

(1) In the lattice of binary trees, a tree T is covered by the set of trees obtained from it by addition of a single node, in all possible ways.
(2) In the bracket tree, a tree T' covers a single tree T obtained from T' by deleting it's left-most node if any, or its root otherwise, and replacing the deleted node by its own right subtree if any.

It can then be shown that performing $\mathbf{Q}_\bullet\mathbf{Q}_T$ and $\mathbf{P}_\bullet\mathbf{P}_T$ respectively corresponds exactly to applying the above operations to T. Hence from Fomin's statement that the lattice of binary trees is dual to the bracket tree, we have:

COROLLARY 2.1. *$\Gamma_{Q_\bullet}^{left}$ and $\Gamma_{P_\bullet}^{left}$ are dual as defined by Fomin.*

3. Dual graded graphs in Sym and *QSym*

In the same way as in **PBT**, one can build two graded graphs in the algebras of noncommutative symmetric functions (**Sym**), and of quasi symmetric functions (*QSym*). They are associated to the operation of multiplication of ribbon Schur functions (resp., of quasi-ribbon functions) by R_1 (resp., F_1), see [**2**] for definitions. This gives us a third example of graph duality arising from multiplications in two dual Hopf algebras. The graphs are illustrated below (Fig. 6 and Fig. 7).

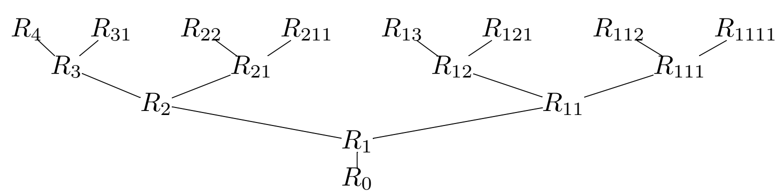


FIGURE 6. $\Gamma_{R_1}^{right}$: right multiplication by R_1 in **Sym**

Investigating the duality of the two graphs defined above, we found that they are isomorphic to two dual graded graphs studied by Fomin [**7**] and known as the lifted binary tree and Binword (Fig. 8 and Fig. 9). Their vertices are words on the alphabet $\{0,1\}$. In the first graph, a word w is covered by the two words $w.0$ and $w.1$ (where . denotes the usual concatenation of words), except 0 which is only covered by 1. In Binword, there exists an edge from u to v if u is obtained by deleting a single letter (but not the first) from v, and in addition, there is an edge from 0 to 1.

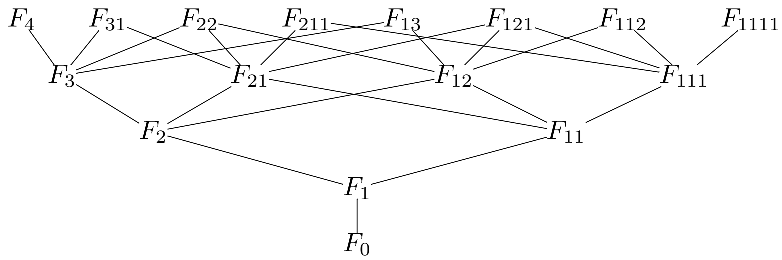


FIGURE 7. Γ_{F_1} : multiplication by F_1 in $QSym$

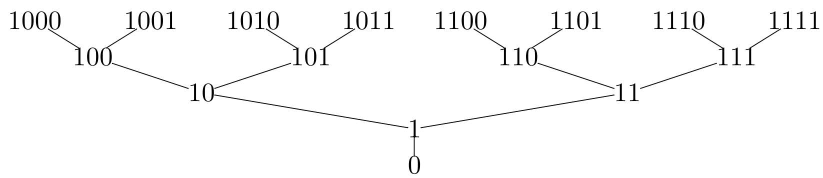


FIGURE 8. The lifted binary tree

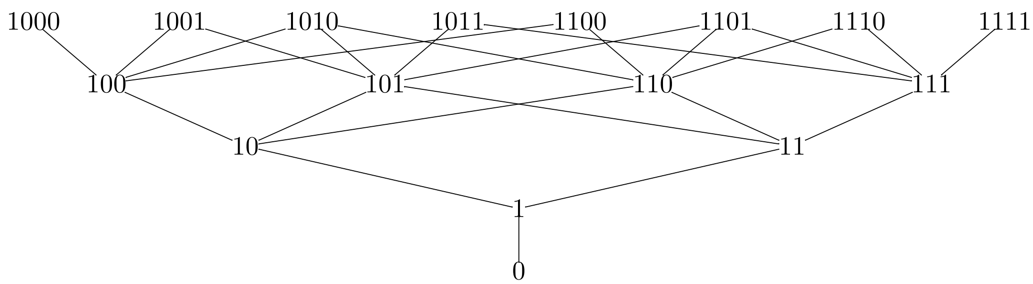
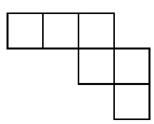
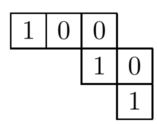


FIGURE 9. Binword, a dual of the lifted binary tree

REMARK 3.1. There is a one-to-one correspondence between compositions of an integer n and the vertices of rank n in the lifted binary tree. A composition I is identified with the word w_I obtained by filling its ribbon diagram from left to right and from top to bottom, with 1 in the first box and in any box following a descent, 0 elsewhere.

$I = (3, 2, 1) \quad =$  $\; ; \quad w_I \quad =$  $\quad = \quad 100101$

PROPOSITION 3.1. *$\Gamma^{right}_{R_1}$ and Γ_{F_1} are isomorphic to the lifted binary tree and it's dual binword, respectively.*

COROLLARY 3.1. *$\Gamma^{right}_{R_1}$ and Γ_{F_1} are dual as defined by Fomin.*

4. Schensted-Fomin algorithm for dual graded graphs

In all that follows and unless otherwise stated, G_1 and G_2 will denote two graded graphs in r-duality, with a zero denoted $\widehat{0}$.

4.1. r-correspondences.

As introduced in [**8**], r-correspondences are bijective realizations of equation (1.2) and its generalizations (1.3). Let ϕ be a bijective map associating pairs (b_1, b_2) to triples (a_1, a_2, α) where a_1 and b_1 are edges in

G_1, a_2 and b_2 are edges in G_2 such that *start* $(a_1) =$ *start* (a_2), *end* $(b_1) =$ *end* (b_2), and $\alpha \in \{0, 1, ..., r\}$. The map ϕ is said to be an r-correspondence if the following conditions are satisfied:

(i) if $\phi(b_1, b_2) = (a_1, a_2, \alpha)$ then $end(a_1) = start(b_2)$ and $end(a_2) = start(b_1)$
(ii) if b_1 and b_2 are degenerated ($b_1 = b_2 = (x_0, x_0)$) then $\phi(b_1, b_2) = (b_1, b_2, 0)$

An important lemma is the following:

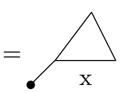
Lemma 4.1. *There exists an r-correspondence between two graded graphs G_1 and G_2 if and only if G_1 and G_2 are in r-duality* [**7**].

One goal of Fomin's construction is to use r-correspondences to establish bijective maps between pairs of paths in G_1 and G_2 both starting at $\widehat{0}$ and having a common end point of rank n, and permutations of the symmetric group $\mathfrak{S}_n$. In this section we define two natural r-correspondences associated to the pairs of dual graded graphs of section 2 and 3. Using the Schensted-Fomin algorithm for dual graded graphs, we will later see that these r-correspondences are parallel versions of the hypoplactic and sylvester insertion algorithms. A combinatorial description of this algorithm is given in section 4.2.

4.1.1. *A natural r-correspondence in* ***PBT***'*s graphs.*

The following is an algorithm to find $a_1 = (t, x)$, $a_2 = (t, y)$ and α, when $b_1 = (y, z)$ and $b_2 = (x, z)$ are given, $r = 1$ in this case.

Function *getAr*:

Inputs: x, y, z;
Outputs: t, α;
Begin
  If $x = z$ **then** $t = y$ **and** $\alpha = 0$
  Else if $y = z$ **then** $t = x$ **and** $\alpha = 0$
  Else if $x \neq y$ **then**
    $t =$ (y without its left-most node, replaced by its own right subtree if any) **and** ($\alpha = 0$) ; (1)
  Else if $z = x' :=$ 

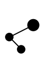
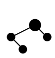
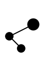
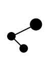
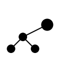
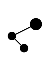
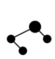
 (that is x + one node added to the left of its left-most node) **then** (2)
    ($t = x$) **and** ($\alpha = 1$) ; (3)
  Else
    $t =$ (y without its left-most node) **and** ($\alpha = 0$) ; (4)
  End_if
End.

(1): for this correspondence to be well defined, one should prove that t is covered by x in $\Gamma^{left}_{Q_\bullet}$. Indeed, (b_1, b_2) is in this case a DU-path from y to x and it is the only DU-path from y to x since $\Gamma^{left}_{P_\bullet}$ is a tree. And since $\Gamma^{left}_{Q_\bullet}$ and $\Gamma^{left}_{P_\bullet}$ are dual graphs (**corollary 2.1**), there is a single UD-path from y to x, necessary having t as middle point.
(2): this serves to define x'.
(4): (b_1, b_2) is the 2^{nd} DU-loop of the form (x, x); it will match the unique UD-loop (x, x). The first is processed in (3).

Of course this algorithm is invertible.

Proposition 4.1. *The previous algorithm defines an r-correspondence in $\Gamma^{left}_{Q_\bullet}$ and $\Gamma^{left}_{P_\bullet}$.*

Now a few examples to illustrate this correspondence.

	y	x	z	t	α	**comments**
1.					0	($x \neq y$) $\Rightarrow$ ($t = y$ without its left-most node)
2.					1	($x = y$ and $z = x'$) $\Rightarrow$ ($t = x$)
3.					0	($x = y$ and $z \neq x'$) $\Rightarrow$ ($t = y$ without its left-most node)

We can do the same for the two graphs $\Gamma^{right}_{R_1}$ and Γ_{F_1} that we have defined in **Sym** and *QSym*. In this case, we observe that the natural choice of an r-correspondence in $\Gamma^{right}_{R_1}$ and Γ_{F_1} converts the Schensted-Fomin algorithm for dual graded graphs (applied to those graphs) into a parallel version of the hypoplactic insertion algorithm. The construction is the following.

4.1.2. *A natural r-correspondence in* ***Sym*** *'s and QSym 's graphs.*

The following algorithm is an adaptation of the one described in [**8**] for the lattice of binary trees and the bracket tree, using our **Proposition 3.1**. Of course this algorithm is invertible.

Function *getAr*:

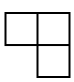
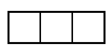
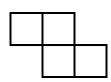
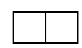
Inputs: x, y, z ;
Outputs: t, α ;
Begin
  If $x = z$ **then** $t = y$ **and** $\alpha = 0$
  Else if $y = z$ **then** $t = x$ **and** $\alpha = 0$
  Else if $(x \neq y)$ or (the last box of z does not follow a descent) **then** (1)
    $t =$ (x without its last box) **and** $(\alpha = 0)$;
  Else /* $x = y$ *and the last box of* z *follows a descent* */ (2)
    $(t = x)$ **and** $(\alpha = 1)$;
  End_if
End.

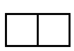
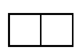
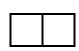
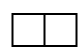
(1): that is w_z ends with 0.
(2): that is w_z ends with 1. w_z is defined in **Proposition 3.1**.

PROPOSITION 4.2. *The previous algorithm defines an r-correspondence in* $\Gamma^{right}_{R_1}$ *and* Γ_{F_1}.

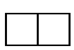
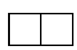
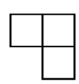
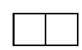
Now a few examples to illustrate this correspondence.

(i): Two cases where $(x \neq y)$ or (the last box of z does not follow a descent)

$y =$  $= 21$; $x =$  $= 3$; $z =$  $= 22$; $t =$  $= 2$; $\alpha = 0$

$y =$  $= 2$; $x =$  $= 2$; $z =$  $= 2$; $t =$  $= 2$; $\alpha = 0$

(ii): A case where $(x = y)$ and (the last box of z follows a descent)

$y =$  $= 2$; $x =$  $= 2$; $z =$  $= 21$; $t =$  $= 2$; $\alpha = 1$

4.2. Schensted-Fomin algorithm for dual graded graphs.

This algorithm was introduced in [**8**]. Given an r-correspondence ϕ, the algorithm establishes a bijective correspondence between pairs of paths in G_1 and G_2 starting at $\widehat{0}$ and having a common end point of rank n, and permutations of $\mathfrak{S}_n$. *We describe a combinatorial version of this algorithm and apply it to the two pairs of graded graphs whose duality has been studied in the previous sections. We also use this description to give a simpler proof of a bijection between permutations and pairs of paths in dual graded graphs, for the case* $r = 1$. The two paths (inputs) are given as two sequences of vertices $v = (v_0, v_1, \cdots, v_n)$ and $w = (w_0, w_1, \cdots, w_n = v_n)$. We'll be using a double entry matrix $M_{\phi,\sigma}$ initialized with v on its last column and w on its last line. In this matrix, lines and columns of odd indices will form an $(n+1)\times(n+1)$ matrix M_ϕ representing a correspondence table for ϕ, while those of even indices will form an $n \times n$ matrix M_σ representing the permutation σ generated from the two paths v and w.

Given (k, l), evaluating $\sigma(k, l)$ requires the values of $M_\phi(k, l-1)$ and $M_\phi(k-1, l)$, and will inform us of the value of $M_\phi(k-1, l-1)$. To do this, we set $z = M_\phi(k, l)$, $x = M_\phi(k, l-1)$, $y = M_\phi(k-1, l)$ and define t to be equal to $M_\phi(k-1, l-1)$. So x, y, and z are known values, and t is being searched for. Setting $b_1 = (y, z)$ and $b_2 = (x, z)$, one can evaluate $\phi(b_1, b_2)$ which can be expressed as $[a_1 := (t, x),\ a_2 := (t, y),\ \alpha]$. Now t is known and it only remains to set $\sigma(k, l) = \alpha$ and we are done. This is illustrated below:

$$M_{\phi,\,\sigma} \quad = \quad \begin{pmatrix} & & & & & & . & \vdots \\ & & & \mathbf{t = ?} & & \mathbf{y} & \cdots & v_k \\ & & & & \sigma_{k,l} \mathbf{= ?} & & . & \\ & & & \mathbf{x} & & \mathbf{z} & \cdots & v_{k+1} \\ & & & \vdots & & \vdots & . & \vdots \\ w_0 & w_1 & \cdots & w_l & & w_{l+1} & \cdots & v_n \end{pmatrix}$$

M_ϕ (M in short) and σ will be filled from their $(n,n)^{th}$ component to the $(0,0)^{th}$ for M_ϕ, or $(1,1)^{th}$ for σ, following diagonals:

Step 1: use $x = w_{n-1}$, $y = v_{n-1}$ and $z = v_n = w_n$ to compute $M_{n-1,n-1} = t$ and $\sigma_{n,n} = \alpha$, satisfying $(a_1, a_2, \alpha) = \phi(b_1, b_2)$ with $b_1 = (y,z)$, $b_2 = (x,z)$, $a_1 = (t,x)$ and $a_2 = (t,y)$.

$$M_{\phi,\,\sigma} \quad = \quad \begin{pmatrix} & & & & & & v_0 \\ & & & & & & v_1 \\ & & & & & & v_2 \\ & & & & & & \vdots \\ & & & & & & v_{n-2} \\ & & & & M_{n-1,\,n-1} & & v_{n-1} \\ & & & & & \sigma_{n,n} & \\ w_0 & w_1 & \cdots & w_{n-2} & w_{n-1} & & v_n \end{pmatrix}$$

Step 2: $M_{n-1,\,n-2}$, $\sigma_{n,n-1}$, $M_{n-2,\,n-1}$ and $\sigma_{n-1,n}$ are computed.

Step 2n-1: M_ϕ and σ are completely filled. The number $2n-1$ of steps is determined by the total number of diagonals in M_ϕ which is an $(n+1)\times(n+1)$ matrix.

Below is a computer implementable description of the algorithm:

Function *permutation_from_paths*:

Inputs: v, w, ϕ
Outputs: σ
Temporary variables: $M_\phi, b_1, b_2, a_1, a_2, \alpha, L, L_0, k, l$;
Begin
$n = \mathrm{length}(v) - 1$; /* *length(v): number of consecutive points defining v, that is 1 more than the number of edges in the path represented by v* */
For all k **from** 0 **to** n **do**
$M_\phi(k,n) = v(k)$ **and** $M_\phi(n,k) = w(k)$;
End_loop.
$KL = \{\, (n,n) \,\}$; /* *first pair of indices to process* */
While $L \neq \{\,\}$ **do**
$L_0 = \{\,\}$; /* *the empty set* */
For all (k,l) in L **do**
$x = M_\phi(k, l-1)$ **and** $y = M_\phi(k-1, l)$ **and** $z = M_\phi(k,l)$;
$b_1 = (y,z)$ **and** $b_2 = (x,z)$;
$(a_1, a_2, \alpha) = \phi(b_1, b_2)$;
$M_\phi(k-1, l-1) = t$; /* *common origin to the edges* a_1 *and* a_2 */
$\sigma(k,l) = \alpha$;
If $k > 1$ **then** $L_0 = L_0 \cup \{\, (k-1, l) \,\}$;
If $l > 1$ **then** $L_0 = L_0 \cup \{\, (k, l-1) \,\}$;
End_loop.
$L = L_0$; /* *next pairs of indices to process* */
End_loop.
End.

It seems not obvious from this description that M_σ is indeed a permutation matrix. Below is a simple proof in the case of a simple duality ($r = 1$), together with the proof of the bijection between permutations and pairs of paths. The reasoning remains valid only for $r = 1$.

Proof. Let M'_ϕ stands for M_ϕ where each vertex is replaced by its rank in the graph. One first observes that M'_ϕ is an $(n+1)\times(n+1)$ matrix satisfying the following conditions: the first line and the first column are initialized with 0 while the last ones are initialized with integer entries increasing from 0 to n ; entries increase at most by 1 on lines and columns ; and for any 2×2 sub-matrix the difference of the sums on the first and second diagonals is 0 or 1. This is formally equivalent to:

$$\begin{cases} m(i,n) = i \ ; \ m(n,j) = j & i,j = 0,\cdots,n \\ m(i,j+1) - m(i,j) \ \in \ \{0,1\} & \\ m(i+1,j) - m(i,j) \ \in \ \{0,1\} & \\ m(i+1,j+1) + m(i,j) - m(i+1,j) - m(i,j+1) \ \in \ \{0,1\} & \end{cases} \tag{4.1}$$

Next one uses the definition and properties of ϕ to see that:

$$M_\sigma(i,j) = M'_\phi(i+1,j+1) + M'_\phi(i,j) - M'_\phi(i+1,j) - M'_\phi(i,j+1) \tag{4.2}$$

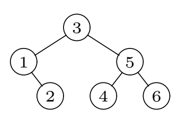
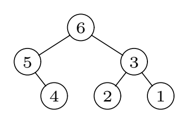
Finally, one establishes a bijective map between matrices satisfying (4.1) on the one hand, and permutations of the symmetric group $\mathfrak{S}_n$ on the other hand, using (4.2) to determine the permutation matrix associated to any matrix satisfying (4.1). So the described algorithm (*permutation_from_paths*) sends a pair of paths to a permutation.

As for the proof of the bijection between permutations and pairs of paths, first notice that the above algorithm is naturally invertible. Given a permutation σ, initialize the first line and the first column of $M_{\phi,\sigma}$ with the common zero of the two graphs. Then fill the permutation matrix M_σ with 1's and 0's, and starting from the upper left corner, fill M_ϕ using ϕ^{-1}. Hence establishing a bijective correspondence between pairs of paths in G_1 and G_2 starting at $\widehat{0}$ and having a common end point of rank n, on the one hand, and permutations of $\mathfrak{S}_n$, on the other hand. ☐

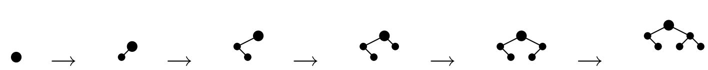
Now let us apply this algorithm to the graphs and r-correspondences we studied in the previous sections.

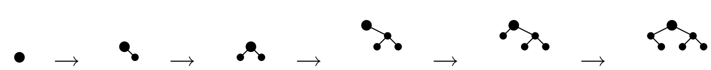
4.2.1. *Schensted-Fomin algorithm applied to* ***PBT****'s graphs.*

We identify any permutation $\alpha \in \mathfrak{S}_n$ with two *natural* paths v_α and w_α, the first in $\Gamma^{left}_{Q_\bullet}$ and the second in $\Gamma^{left}_{P_\bullet}$. These two paths are both paths from the empty tree $\emptyset$ to $T_\alpha = \text{shape}(\mathcal{P}(\alpha))$. The path in $\Gamma^{left}_{Q_\bullet}$ is the sequence of shapes of partial binary search trees corresponding to inserting the last k letters of α, for $k = 0\,..\,n$. As for the path in $\Gamma^{left}_{P_\bullet}$, the shapes defining it correspond to selecting in α only letters greater than k, for $k = n\,..\,0$. For example, let us consider $\alpha = 645213$, applying the sylvester correspondence to α leads to:

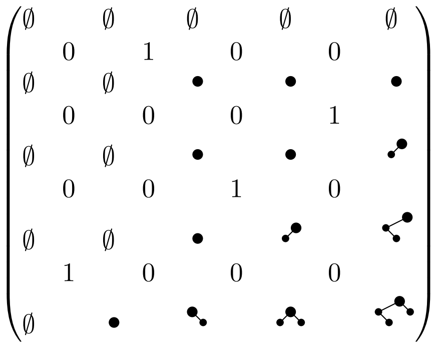
$\mathcal{P}(\alpha) \quad = \quad$  $\quad ; \quad \mathcal{Q}(\alpha) \quad = \quad \mathcal{T}(\alpha^{-1}) \quad = \quad$ 

The two natural paths are then:

$v_\alpha \quad = \quad \emptyset \quad \to$ 

$w_\alpha \quad = \quad \emptyset \quad \to$ 

Now let us choose a smaller example to which we will apply the Schensted-Fomin algorithm for dual graded graphs, say $\gamma = 4213$. Below are $M_{\phi,\sigma}$; γ and σ:



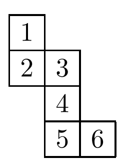
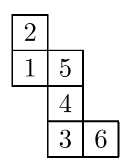
$$; \begin{pmatrix} 0 & 0 & 1 & \mathbf{0} \\ 0 & 1 & \mathbf{0} & 0 \\ 0 & \mathbf{0} & 0 & 1 \\ \mathbf{1} & 0 & 0 & 0 \end{pmatrix} ; \begin{pmatrix} 0 & 1 & 0 & \mathbf{0} \\ 0 & 0 & \mathbf{0} & 1 \\ 0 & \mathbf{0} & 1 & 0 \\ \mathbf{1} & 0 & 0 & 0 \end{pmatrix} $$

$\gamma = 4213$ and $\sigma = 4132$, so the permutation produced by the Schensted-Fomin algorithm in **PBT** using our natural r-correspondence (see **4.1.1**), differs from the initial permutation. But one observation can be made on the relation between the two permutations: *one is obtained from the other by reflexion on the second diagonal*. It is known [**8**] that a natural choice of an r-correspondence in the Young lattice converts the Schensted-Fomin algorithm for dual graded graphs into a parallel version of the Robinson-Schensted algorithm.

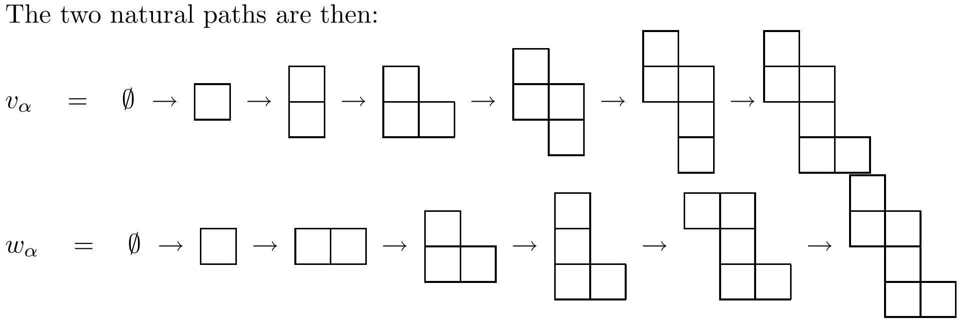
PROPOSITION 4.3. *Our natural choice of r-correspondence in $\Gamma^{left}_{Q_\bullet}$ and $\Gamma^{left}_{P_\bullet}$ converts the Schensted-Fomin algorithm for dual graded graphs into a parallel version of the sylvester insertion algorithm.*

4.2.2. *Schensted-Fomin algorithm applied to **Sym**'s and QSym's graphs.*

We identify any permutation $\alpha \in \mathfrak{S}_n$ with two *natural* paths v_α and w_α, the first in $\Gamma^{right}_{R_1}$ and the second in Γ_{F_1}. These two paths are both paths from the empty composition $\emptyset$ to the recoil composition of α. The paths in $\Gamma^{right}_{R_1}$ is the sequence of recoil compositions of restrictions of α to $[1..k]$, for $k = 0..n$. As for the path in Γ_{F_1}, it is made of descent compositions of restrictions of α^{-1} to $[k..n]$, for $k = (n+1)..1$. For example, consider $\alpha = 215436$, applying the hypoplactic insertion algorithm [**4**] leads to the following quasi-ribbon and ribbon diagrams:

$Q_r(\alpha) \quad = $  ; $\quad R(\alpha) \quad = \quad Q_r(\alpha^{-1}) \quad =$ 

The two natural paths are then:



Now let us choose a smaller example to which we will apply the Schensted-Fomin algorithm for dual graded graphs, say $\gamma = 1243$. Below are $M_{\phi,\sigma}$; γ and σ:

$$\begin{pmatrix}
\emptyset & & \emptyset & & \emptyset & & \emptyset & & \emptyset \\
 & \mathbf{0} & & \mathbf{0} & & \mathbf{0} & & \mathbf{1} & \\
\emptyset & & \emptyset & & \emptyset & & \emptyset & & 1 \\
 & \mathbf{0} & & \mathbf{0} & & \mathbf{1} & & \mathbf{0} & \\
\emptyset & & \emptyset & & \emptyset & & 1 & & 2 \\
 & \mathbf{1} & & \mathbf{0} & & \mathbf{0} & & \mathbf{0} & \\
\emptyset & & 1 & & 1 & & 2 & & 3 \\
 & \mathbf{0} & & \mathbf{1} & & \mathbf{0} & & \mathbf{0} & \\
\emptyset & & 1 & & 11 & & 21 & & 31
\end{pmatrix} ; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} ; \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$\gamma = 1243$ and $\sigma = 3421$, so the permutation produced by the Schensted-Fomin algorithm in **Sym** and *QSym* using the adaptation (see **4.1.2**) of Fomin's natural r-correspondence, differs from the initial permutation. Once more, an observation can be made on the relation between them: *one is obtained from the other by reflexion in a central vertical line.*

PROPOSITION 4.4. *The natural choice of an r-correspondence in $\Gamma_{R_1}^{right}$ and Γ_{F_1} converts the Schensted-Fomin algorithm for dual graded graphs into a parallel version of the hypoplactic insertion algorithm.*

5. Conclusion

As suggested by the three examples studied in this paper, which are not isolated cases since numerous other examples can be found in some other algebras, there seems to be a strong connection between dual graded graphs and the construction of some dual Hopf algebras. For more examples, in the Hopf algebra of free quasi-symmetric functions (**FQSym**) we've consider the two pairs of graded graphs describing the operations $\mathbb{F}_1\mathbb{F}_\sigma$ and $\mathbb{G}_\sigma\mathbb{G}_1$ for the first pair, $\mathbb{F}_\sigma\mathbb{F}_1$ and $\mathbb{G}_\sigma\mathbb{G}_1$ for the second pair. From explicit computations on finite realizations of those graphs, we believe that they are also examples of graph duality arising from multiplication in dual Hopf algebras. Finally, another interesting example is the algebra of free symmetric functions denoted **FSym**, providing a realization of the algebra of tableaux introduced by Poirier and Reutenauer [**9**] as a subalgebra of the free associative algebra. In this case, the Hopf algebra duality may be identified with the duality of two graphs the duality of which is established in [**7**]: the Schensted graph and the SYT-Tree. Their vertices are Young tableaux.

So dual graded graphs could be viewed as the description of the multiplication rules for products of basis elements by the ones of rank 1, in two dual Hopf algebras constructed by means of a congruence relation on words. The congruence itself could be obtained by the Schensted-Fomin algorithm using a certain r-correspondence in those graphs. This is clearly observed in the construction of the algebra of symmetric functions (*Sym*), the algebra of noncommutative symmetric functions (**Sym**), the algebra of quasi-symmetric functions (*QSym*) and the algebra of planar binary trees (**PBT**), using the plactic, hypoplactic and syslvester congruences respectively.

References

[1] *D. E. Knuth*, *The art of computer programming*, vol.3: *Searching and sorting* (Addison-Wesley, 1973).
[2] *D. Krob* and *J.-Y. Thibon*, *Noncommutative symmetric function IV: Quantum linear groups and Hecke algebras at q=0*, J. Alg. Comb. **6** (1997), 339-376.
[3] *F. Hivert*, *J. C. Novelli*, and *J.-Y. Thibon*, *The Algebra of Binary Search Trees*, Theoretical Computer Science **339** (2005), 129-165.
[4] *F. Hivert*, *G. Duchamp* and *J.-Y. Thibon*, *Noncommutative symmetric function VI: free quasi-symmetric functions and related algebras*, Int. J. Alg. Comp. **12** (2002), 357-404.
[5] *J.-L. Loday* and *M. O. Ronco*, *Hopf Algebra of the Planar Binary Trees*, Adv. Math, **139** (1998) n. **2**, 293-309.
[6] *I. G. MacDonald*, *Symmetric functions and Hall Polynomials*, 2nd ed, Clarendon Press, Oxford Science Publications, **139** (1995).
[7] *S. Fomin*, *Duality of Graded Graphs*, J. Alg. Comb. **3** (1994), 357-404.
[8] *S. Fomin*, *Schensted Algorithms for Dual Graded Graphs*, J. Alg. Comb. **4** (1995), 5-45.
[9] *S. Poirier* and *C. Reutenauer*, *Algèbre de Hopf des tableaux*, Ann. Sci. Math. Qébec **19** (1995), 79-90.

INSTITUT GASPARD-MONGE, LABORATOIRE D'INFORMATIQUE, UMR 8049, 77454 MARNE-LA-VALLÉE CEDEX 2, FRANCE
E-mail address: `jnzeutch@etudiant.univ-mlv.fr`
URL: `http://monge.univ-mlv.fr/~nzeutcha`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Pieri's Formula for Generalized Schur Polynomials

NUMATA, Yasuhide

Abstract. We define a generalization of Schur polynomials as a expansion coefficient of generalized Schur operators. We generalize the Pieri's formula to the generalized Schur polynomials.

Résumé. Nous définissons une généralisation de polynômes de Schur comme un coefficient de l'expansion d'opérateurs de Schur généralisés. Nous généralisons la formule du Pieri aux polynômes de Schur généralisés.

1. Introduction

Young's lattice is a prototypical example of differential posets defined by Stanley [**9**]. Young's lattice has so called the Robinson correspondence, the correspondence between permutations and pairs of standard tableaux whose shapes are the same Young diagram. This correspondence is generalized for differential posets or dual graphs (that is a generalization of differential posets) by Fomin [**3**].

Young's lattice also has the Robinson-Schensted-Knuth correspondence, the correspondence between certain matrices and pairs of semi-standard tableaux. Fomin generalizes the method of the Robinson correspondence to that of the Robinson-Schensted-Knuth correspondence in his paper [**4**]. The operators in Fomin [**4**] are called generalized Schur operators. We can define a generalization of Schur polynomials by generalized Schur operators.

A complete symmetric polynomial is a Schur polynomial associated with a Young diagram consisting of only one row. Schur polynomials satisfy the Pieri's formula, the formula describing products of a complete symmetric polynomial and a Schur polynomial as sums of Schur polynomials like the following;

$$h_i(t_1,\dots,t_n)s_\lambda(t_1,\dots,t_n) = \sum_{\mu} s_\mu(t_1,\dots,t_n),$$

where the sum is over all μ's that are obtained from λ by adding i boxes, with no two in the same column, h_i is the i-th complete symmetric polynomial and s_λ is the Schur polynomial associated with λ.

We generalize the Pieri's formula to generalized Schur polynomials (Theorem 3.2 and Proposition 3.3).

2. Definition

We introduce two types of polynomials in this section. One of them is a generalization of Schur polynomials. The other is a generalization of complete symmetric polynomials. We will show Pieri's formula for these polynomials in Section 3.

2.1. Schur Operators. First we recall generalized Schur operators defined by Fomin [**4**]. We define a generalization of Schur function as expansion coefficients of generalized Schur operators.

Let K be a field of characteristic zero that contains all formal power series of variables $t, t', t_1, t_2, \dots$. Let V_i be finite dimensional K-vector spaces for all $i \in \mathbb{Z}$. Fix a basis Y_i of each V_i so that $V_i = KY_i$ and

2000 *Mathematics Subject Classification.* Primary 05E10; Secondary 05C78.

Key words and phrases. Generalized Schur operators; Pieri's formula; Dual graded graph; Differential poset; Symmetric functions; Binary trees.

$V = KY$ where $Y = \bigcup_i Y_i$. The *rank function* on V which maps $\lambda \in V_i$ to i is denoted by ρ. We say that V *has the minimum* $\varnothing$ if $Y_i = \emptyset$ for $i < 0$ and $Y_0 = \{\varnothing\}$.

For a sequence $\{A_i\}$ and a formal variable x, we write $A(x)$ for the generating function $\sum_{i \geq 0} A_i x^i$.

Hereafter, for $i > 0$, let D_i and U_i be linear operators on V satisfying $\rho(U_i\lambda) = \rho(\lambda) + i$ and $\rho(D_i\lambda) = \rho(\lambda) - i$ for $\lambda \in Y$. In other words, the images $D_j(V_i)$ and $U_j(V_i)$ of V_i by D_j and U_j are contained in V_{i-j} and V_{i+j} for $i \in \mathbb{Z}$ and $j \in \mathbb{N}$ respectively. We call D_i or $D(t)$ and U_i or $U(t)$ down operators and up operators.

Definition 2.1. Let $\{a_i\}$ be a sequence of elements of K. Down and up operators $D(t_1)\cdots D(t_n)$ and $U(t_n)\cdots U(t_1)$ are said to be *generalized Schur operators* if the equation $D(t')U(t) = a(tt')U(t)D(t')$ holds.

We write $*$ for the conjugation with respect to the natural pairing $\langle\ ,\ \rangle$ in KY. For all i, U_i^* and D_i^* act as down and up operators, respectively. By definition, $U^*(t')D^*(t) = a(tt')D^*(t)U^*(t')$ if $D(t')U(t) = U(t)D(t')a(tt')$. Hence down and up operators $U^*(t_n)\cdots U^*(t_1)$ and $D^*(t_1)\cdots D^*(t_n)$ are also generalized Schur operators when $D(t_1)\cdots D(t_n)$ and $U(t_n)\cdots U(t_1)$ are generalized Schur operators.

Let down and up operators $D(t)$ and $U(t)$ be generalized Schur operators with $\{a_i\}$ where $a_0 \neq 0$. Since $a_0 \neq 0$, there exists $\{b_i\}$ such that $a(t)b(t) = 1$. Hence the equation $D(t')U(t) = a(tt')U(t)D(t')$ implies

$$U(t)D(t') = b(tt')D(t')U(t) \tag{2.1}$$

and

$$D^*(t')U^*(t) = b(tt')U^*(t)D^*(t'). \tag{2.2}$$

Let ρ' be $-\rho$. We take ρ' as rank function for the same vertex set V. For this rank function ρ' and the vector space V, D_i^* and U_i^* act as down and up operators, respectively. Since they satisfy the equation (2.2), down and up operators $D^*(t)$ and $U^*(t)$ are generalized Schur operators with $\{b_i\}$. Similarly, it follows from the equation (2.1) that down and up operators $U(t)$ and $D(t)$ are also generalized Schur operators with $\{b_i\}$ for ρ' and V.

Definition 2.2. Let $D(t_1)\cdots D(t_n)$ and $U(t_n)\cdots U(t_1)$ be generalized Schur operators. For $\lambda \in V$ and $\mu \in Y$, we write $s^D_{\lambda,\mu}(t_1,\ldots,t_n)$ and $s_U^{\mu,\lambda}(t_1,\ldots,t_n)$ for the coefficient of μ in $D(t_1)\cdots D(t_n)\lambda$ and $U(t_n)\cdots U(t_1)\lambda$, respectively. We call these polynomials $s^D_{\lambda,\mu}(t_1,\ldots,t_n)$ and $s_U^{\mu,\lambda}(t_1,\ldots,t_n)$ *generalized Schur polynomials*.

Generalized Schur polynomials $s^D_{\lambda,\mu}(t_1,\ldots,t_n)$ are symmetric in the case when $D(t)D(t') = D(t')D(t)$ but not symmetric in general. It follows by definition that

$$\begin{aligned} s^D_{\lambda,\mu}(t_1,\ldots,t_n) &= \langle D(t_1)\cdots D(t_n)\lambda, \mu\rangle \\ &= \langle \lambda, D^*(t_n)\cdots D^*(t_1)\mu\rangle \\ &= s_{D^*}^{\lambda,\mu}(t_1,\ldots,t_n) \end{aligned}$$

for $\lambda, \mu \in Y$.

Example 2.3. Our prototypical example is Young's lattice $\mathbb{Y}$ that consists of all Young diagrams. Let a basis Y, K-vector space V and rank function ρ be Young lattice $\mathbb{Y}$, the K-vector space $K\mathbb{Y}$ and the ordinal rank function ρ which maps Young diagram λ to the number of boxes in λ. Young's lattice $\mathbb{Y}$ has the minimum $\varnothing$ the Young diagram with no boxes. Define U_i and D_i by $U_i(\mu) = \sum_\lambda \lambda$, where the sum is over all λ's that are obtained from μ by adding i boxes, with no two in the same column; and by $D_i(\lambda) = \sum_\mu \mu$, where the sum is over all μ's that are obtained from λ by removing i boxes, with no two in the same column. Then the operators $D(t_1)\cdots D(t_n)$ and $U(t_n)\cdots U(t_1)$ are generalized Schur operators with $\{a_i = 1\}$. In this case, both $s^D_{\lambda,\mu}(t_1,\ldots,t_n)$ and $s_U^{\lambda,\mu}(t_1,\ldots,t_n)$ are equal to the skew Schur polynomial $s_{\lambda/\mu}(t_1,\ldots,t_n)$ for λ and $\mu \in \mathbb{Y}$.

2.2. Weighted Complete Symmetric Polynomials. Next we introduce a generalization of complete symmetric polynomials. We define weighted symmetric polynomials inductively.

Definition 2.4. Let $\{a_m\}$ be a sequence of elements of K. We define the i-th weighted complete symmetric polynomial $h_i^{\{a_m\}}(t_1,\ldots,t_n)$ by

$$h_m^{\{a_n\}}(t_1,\ldots,t_n)=\begin{cases}\sum_{j=0}^{i} h_j^{\{a_m\}}(t_1,\ldots,t_{n-1})h_{i-j}^{\{a_m\}}(t_n), & (\text{for } n>1)\\ h_i^{\{a_m\}}(t_1)=a_i t_1^i & (\text{for } n=1).\end{cases}\tag{2.3}$$

By definition, the i-th weighted complete symmetric polynomial $h_i^{\{a_m\}}(t_1,\ldots,t_n)$ is a homogeneous symmetric polynomial of degree i.

Example 2.5. When a_i equal 1 for all i, $h_j^{\{1,1,\ldots\}}(t_1,\ldots,t_n)$ equals the complete symmetric polynomial $h_j(t_1,\ldots,t_n)$. In this case, the formal power series $\sum_i h_i(t)$ equals the generating function $a(t)=\sum_i t^i=\frac{1}{1-t}$.

Example 2.6. When a_i equal $\frac{1}{i!}$ for all i, $h_j^{\{\frac{1}{m!}\}}(t_1,\ldots,t_n)=\frac{1}{i!}(t_1+\cdots+t_n)^i$ and $\sum_j h_j^{\{\frac{1}{m!}\}}(t)=\exp(t)=a(t)$.

In general, the formal power series $\sum_i h_i^{\{a_m\}}(t)$ equals the generating function $a(t)=\sum a_i t^i$ by the definition of weighted complete symmetric polynomials. It follows from the equation (2.3) that $a(t_1)a(t_2)=\sum_i h_i^{\{a_m\}}(t_1)\sum_j h_j^{\{a_m\}}(t_2)=\sum_j h_j^{\{a_m\}}(t_1,t_2)$. Since the weighted complete symmetric polynomials satisfy the equation (2.3),

$$\begin{aligned}a(t_1)\cdots a(t_{n-1})a(t_n) &= \sum_i h_i^{\{a_m\}}(t_1,\ldots,t_{n-1})\sum_j h_j^{\{a_m\}}(t_n)\\ &= \sum_i\sum_{k=0}^{i} h_{i-k}^{\{a_m\}}(t_1,\ldots,t_{n-1})h_k^{\{a_m\}}(t_n)\\ &= \sum_i h_i^{\{a_m\}}(t_1,\ldots,t_n)\end{aligned}$$

if $a(t_1)\cdots a(t_{n-1})=\sum h_i^{\{a_m\}}(t_1,\ldots,t_{n-1})$. Hence

$$a(t_1)\cdots a(t_n)=\sum_i h_i^{\{a_m\}}(t_1,\ldots,t_n)$$

as in the case when $a_i=1$ for all i. It follows from this relation that $h_0^{\{a_m\}}(t_1,\ldots,t_n)=a_0^n$.

3. Main Theorem

We show some properties of generalized Schur polynomials and weighted complete symmetric polynomials in this section. We show Pieri's formula (Theorem 3.2 and Proposition 3.3) generalized to our polynomials, the main results in this paper.

First we describe the commuting relation of U_i and $D(t_1)\cdots D(t_n)$. This relation implies Pieri's formula for our polynomials. It also follows from this relation that the weighted complete symmetric polynomials are written as linear combinations of generalized Schur polynomials when V has the minimum.

Proposition 3.1. *Generalized Schur operators $D(t_1)\cdots D(t_n)$ and $U(t_n)\cdots U(t_1)$ with $\{a_i\}$ satisfy*

$$D(t_1)\cdots D(t_n)U_i=\sum_{j=0}^{i} h_{i-j}^{\{a_m\}}(t_1,\ldots,t_n)U_jD(t_1)\cdots D(t_n).$$

In the case when the K-vector space V has the minimum $\varnothing$, weighted complete symmetric polynomials are written as linear combinations of generalized Schur polynomials.

Proposition 3.2. *For generalized Schur operators $D(t_1)\cdots D(t_n)$ and $U(t_n)\cdots U(t_1)$ with $\{a_i\}$ on V with the minimum $\varnothing$, the following equations hold for all $i\geq 0$;*

$$s^D_{U_i\varnothing,\varnothing}(t_1,\ldots,t_n)=h_i^{\{a_m\}}(t_1,\ldots,t_n)d_0^n u_0,$$

where u_0 and $d_0\in K$ satisfy $D_0\varnothing=d_0\varnothing$ and $U_0\varnothing=u_0\varnothing$.

EXAMPLE 3.1. In the prototypical example $\mathbb{Y}$, Proposition 3.2 means that the Schur polynomial $s_{(i)}$ corresponding to Young diagram with only one row equals the complete symmetric polynomial h_i.

Next we consider the case when Y may not have a minimum. It follows from Proposition 3.1 that

$$\langle D(t_1)\cdots D(t_n)U_i\lambda,\mu\rangle = \langle \sum_{j=0}^{i} h_{i-j}^{\{a_m\}}(t_1,\ldots,t_n)U_jD(t_1)\cdots D(t_n)\lambda,\mu\rangle$$

for $\lambda\in V$ and $\mu\in Y$. This equation implies Theorem 3.2, the main result in this paper.

THEOREM 3.2 (Pieri's formula). *For any $\mu\in Y_k$ and any $\lambda\in V$, generalized Schur operators satisfy*

$$s^D_{U_i\lambda,\mu}(t_1,\ldots,t_n) = \sum_{j=0}^{i} h_{i-j}^{\{a_m\}}(t_1,\ldots,t_n) \sum_{\nu(\in Y_{k-j})} \langle U_j\nu,\mu\rangle s^D_{\lambda,\nu}(t_1,\ldots,t_n).$$

If Y has the minimum $\varnothing$, this theorem implies the following proposition.

PROPOSITION 3.3. *For all $\lambda\in V$, the following equations hold;*

$$\begin{aligned} s^D_{U_i\lambda,\varnothing}(t_1,\ldots,t_n) &= h_i^{\{a_m\}}(t_1,\ldots,t_n)u_0 s^D_{\lambda,\varnothing}(t_1,\ldots,t_n)\\ &= s^D_{U_i\varnothing,\varnothing}(t_1,\ldots,t_n)u_0 s^D_{\lambda,\varnothing}(t_1,\ldots,t_n), \end{aligned}$$

where $U_0\varnothing = u_0\varnothing$.

EXAMPLE 3.3. In the prototypical example $\mathbb{Y}$, for any $\lambda\in\mathbb{Y}$, $U_i\lambda$ means the sum of all Young diagrams obtained from λ by adding i boxes, with no two in the same column. Thus Proposition 3.3 is nothing but the classical Pieri's formula. Theorem 3.2 means Pieri's formula for skew Schur polynomials; for a skew Young diagram λ/μ and $i\in\mathbb{N}$,

$$\sum_{\kappa} s_{\kappa/\mu}(t_1,\ldots,t_n) = \sum_{j=0}^{i}\sum_{\nu} h_{i-j}(t_1,\ldots,t_n)s_{\lambda/\nu}(t_1,\ldots,t_n),$$

where the first sum is over all κ's that are obtained from λ by adding i boxes, with no two in the same column; the last sum is over all ν's that are obtained from μ by removing j boxes, with no two in the same column.

4. More Examples

In this section, we see some examples of generalized Schur operators.

4.1. Shifted Shapes. This example is the same as Fomin [**4**, Example 2.1].
Let Y be the set of all shifted shapes. (i.e., $Y = \{\{(i,j)|i\le j<\lambda_i+i\}|\lambda=(\lambda_1>\lambda_2>\cdots),\lambda_i\in\mathbb{N}\}$.)
Down operators D_i are defined for $\lambda\in Y$ by

$$D_i\lambda = \sum_{\nu} 2^{cc_0(\lambda\setminus\nu)}\nu,$$

where $cc_0(\lambda\setminus\nu)$ is the number of connected components of $\lambda\setminus\nu$ which do not intersect the main diagonal; and the sum is over all ν's that are satisfying $\nu\subset\lambda$, $\rho(\nu)=\rho(\lambda)-i$ and $\lambda\setminus\nu$ contains at most one box on each diagonal.

Up operators U_i are defined for $\lambda\in Y$ by

$$U_i\lambda = \sum_{\mu} 2^{cc(\mu\setminus\lambda)}\mu,$$

where $cc(\lambda\setminus\mu)$ is the number of connected components of $\lambda\setminus\nu$; and the sum is over all μ's that are satisfying $\lambda\subset\mu$, $\rho(\mu)=\rho(\lambda)+i$ and $\lambda\setminus\mu$ contains at most one box on each diagonal.

In this case, since down and up operators $D(t)$ and $U(t)$ satisfy

$$D(t')U(t) = \frac{1+tt'}{1-tt'}U(t)D(t'),$$

down and up operators $D(t_1)\cdots D(t_n)$ and $U(t_n)\cdots U(t_1)$ are generalized Schur operators with $a_0 = 1$, $a_i = 2$ for $i \geq 1$. In this case, for λ, $\mu \in Y$, generalized Schur polynomials $s^D_{\lambda,\mu}$ and $s^{\lambda,\mu}_U$ are respectively $Q_{\lambda/\mu}(t_1,\dots,t_n)$ and $P_{\lambda/\mu}(t_1,\dots,t_n)$, where $P\cdots$ and $Q\cdots$ are the shifted skew Schur polynomials.

In this case, Proposition 3.2 means

$$h_i^{\{1,2,2,2,\dots\}}(t_1,\dots,t_n) = \begin{cases} 2Q_{(i)}(t_1,\dots,t_n) & i>0 \\ Q_{\varnothing}(t_1,\dots,t_n) & i=0 \end{cases}.$$

It also follows that

$$h_i^{\{1,2,2,2,\dots\}}(t_1,\dots,t_n) = P_{(i)}(t_1,\dots,t_n).$$

Proposition 3.3 means

$$h_i^{\{1,2,2,2,\dots\}}Q_\lambda(t_1,\dots,t_n) = \sum_\kappa 2^{cc(\lambda\setminus\mu)}Q_\kappa(t_1,\dots,t_n),$$

where $cc(\lambda\setminus\mu)$ is the number of connected components of $\lambda\setminus\nu$; and the sum is over all μ's that are satisfying $\lambda\subset\mu$, $\rho(\mu)=\rho(\lambda)+i$ and $\lambda\setminus\mu$ contains at most one box on each diagonal.

4.2. Young's Lattice: Dual Identities. This example is the same as Fomin [**4**, Example 2.4]. We take Young's lattice $\mathbb{Y}$ for Y. Up operators U_i are the same as in the prototypical example, (i.e., $U_i\lambda = \sum_\mu \mu$, where the sum is over all μ's that are obtained from λ by adding i boxes, with no two in the same column.) Down operators D'_i are defined by $D'_i = \sum_\mu \mu$, where the sum is over all μ's that are obtained from λ by adding i boxes, with no two in the same row. (In other words, down operators D'_i remove a vertical strip, while up operators U_i add a horizontal strip.)

In this case, since down and up operators $D'(t)$ and $U(t)$ satisfy

$$D'(t')U(t) = (1+tt')U(t)D'(t'),$$

down and up operators $D'(t_1)\cdots D'(t_n)$ and $U(t_n)\cdots U(t_1)$ are generalized Schur operators with $a_0 = a_1 = 1$ $a_i = 0$ for $i\geq 2$. In this case, for λ, $\mu\in Y$, generalized Schur polynomials $s^{D'}_{\lambda,\mu}$ equal $s_{\lambda'/\mu'}(t_1,\dots,t_n)$, where λ' and μ' are the transposes of λ and μ, $s_{\lambda'/\mu'}(t_1,\dots,t_n)$ are the shifted Schur polynomials.

In this case, Proposition 3.2 means

$$h_i^{\{1,1,0,0,0,\dots\}}(t_1,\dots,t_n) = s_{(1^i)}(t_1,\dots,t_n) = e_i(t_1,\dots,t_n),$$

where $e_i(t_1,\dots,t_n)$ stands for the i-th elementally symmetric polynomials.

Proposition 3.3 means

$$e_i(t_1,\dots,t_n)s_\lambda(t_1,\dots,t_n) = \sum_\mu s_\mu(t_1,\dots,t_n),$$

where the sum is over all μ's that are obtained from λ by adding i boxes, with no two in the same row.

For a skew Young diagram λ/μ and $i\in\mathbb{N}$, Theorem 3.2 means

$$\sum_\kappa s_{\kappa/\mu}(t_1,\dots,t_n) = \sum_{j=0}^{i}\sum_\nu h_{i-j}(t_1,\dots,t_n)s_{\lambda/\nu}(t_1,\dots,t_n),$$

where the first sum is over all κ's that are obtained from λ by adding i boxes, with no two in the same row; the last sum is over all ν's that are obtained from μ by removing j boxes, with no two in the same row.

4.3. Planar Binary Trees. Let F be the monoid of words generated by the alphabet $\{1,2\}$ and 0 denotes the word of length 0. We identify F with a poset by $v\leq vw$ for $v,w\in F$. We call an ideal of poset F a *planar binary tree* or shortly a *tree*. An element of a tree is called a *node* of the tree. We write $\mathbb{T}$ for the set of trees and $\mathbb{T}_i$ for the set of trees of i nodes. For $T\in\mathbb{T}$ and $v\in F$, we define T_v by $T_v := \{w\in T|v\leq w\}$.

DEFINITION 4.1. Let T be a tree and m a positive integer. We call a map $\varphi : T\to\{1,\dots,m\}$ a *left-strictly-increasing labeling* if

- $\varphi(w)<\varphi(v)$ for $w\in T$ and $v\in T_{w1}$ and
- $\varphi(w)\leq\varphi(v)$ for $w\in T$ and $v\in T_{w2}$.

We call a map $\varphi : T\to\{1,\dots,m\}$ a *right-strictly-increasing labeling* if

- $\varphi(w) \leq \varphi(v)$ for $w \in T$ and $v \in T_{w1}$ and
- $\varphi(w) < \varphi(v)$ for $w \in T$ and $v \in T_{w2}$.

We call a map $\varphi : T \to \{1, \dots, m\}$ a *binary-searching labeling* if

- $\varphi(w) \geq \varphi(v)$ for $w \in T$ and $v \in T_{w1}$ and
- $\varphi(w) < \varphi(v)$ for $w \in T$ and $v \in T_{w2}$.

First we consider a presentation of increasing labelings as sequences of trees. For a tree $T \in \mathbb{T}$, we call a node $w \in T$ an *l-node* in T if $T_w \subset \{w1^n | n \in \mathbb{N}\}$. A node $w \in T$ is called an *r-node* in T if $T_w \subset \{w2^n | n \in \mathbb{N}\}$. By the definition of increasing labelings φ, the inverse image $\varphi^{-1}(\{1, \dots, n\})$ is a tree for each n. For a right-strictly-increasing labeling φ, $\varphi^{-1}(\{1, \dots, n+1\}) \setminus \varphi^{-1}(\{1, \dots, n\})$ consists of some l-nodes in $\varphi^{-1}(\{1, \dots, n+1\})$. Conversely, for a left-strictly-increasing labeling φ, $\varphi^{-1}(\{1, \dots, n+1\}) \setminus \varphi^{-1}(\{1, \dots, n\})$ consists of some r-nodes in $\varphi^{-1}(\{1, \dots, n+1\})$. Hence we respectively identify right-strictly-increasing and left-strictly-increasing labelings φ with sequences $(\emptyset = T^0, T^1, \dots, T^m)$ of $m+1$ trees such that $T^{i+1} \setminus T^i$ consists of some l-nodes and r-nodes in T^{i+1} for all i.

We define linear operators D and D' on $K\mathbb{T}$ by

$$DT := \sum_{T' \subset T;\ T \setminus T' \text{ consists of some l-nodes}} T',$$

$$D'T := \sum_{T' \subset T;\ T \setminus T' \text{ consists of some r-nodes}} T'.$$

Next we consider binary-searching trees. For $T \in \mathbb{T}$, let s_T be $\{w \in T |$ If $w = v1w'$ then $v2 \notin T$. $w2 \notin T.\}$. The set s_T is a chain. We define S_T by the set of ideals of s_T. For $s \in S_T$, we define $T \ominus s$ by

$$T \ominus s := \begin{cases} T & (s = \emptyset) \\ (T - \max(s)) \ominus (s \setminus \{\max(s)\}) & (s \neq \emptyset), \end{cases}$$

where

$$T - w = (T \setminus T_w) \cup \{wv | w1v \in T_w\}$$

for $w \in T$ such that $w2 \notin T$. There exists the natural inclusion ν from $T - w$ to T defined by

$$\nu(v') = \begin{cases} w1v & v' = wv \in T_w \\ v' & v' \notin T_w. \end{cases}$$

This inclusion induces the inclusion $\nu : T \ominus s \to T$. For a binary-searching labeling φ from $T \in \mathbb{T}$ to $\{1, \dots, m\}$, by the definition of binary-searching labeling, the inverse image $\varphi^{-1}(\{m\})$ is in S_T. The map $\varphi \circ \nu$ induced from φ by the natural inclusion $\nu : T \ominus \varphi^{-1}(\{m\}) \to T$ is a binary-searching labeling from $T \ominus \varphi^{-1}(\{m\})$ to $\{1, \dots, m-1\}$. Hence we identify binary-searching labelings φ with sequences $(\emptyset = T^0, T^1, \dots, T^m)$ of $m+1$ trees such that there exists $s \in S_{T^{i+1}}$ satisfying $T^i = T^{i+1} \ominus s$ for each i.

We define linear operators U on $K\mathbb{T}$ by

$$UT := \sum_{s \in S_T} T \ominus s.$$

These operators $D(t')$, $D'(t')$ and $U(t)$ satisfy the following equations;

$$D(t')U(t) = \frac{1}{1 - tt'} U(t) D(t'),$$
$$D'(t')U(t) = (1 + tt') U(t) D'(t').$$

Hence the generalized Schur polynomials for these operators satisfy the same Pieri's formula as in the case of the classical Young's lattice and its dual construction.

In this case, generalized Schur polynomials are not symmetric in general. For example, since

$$\begin{aligned} &U^*(t_1)U^*(t_2)\{0, 1, 12\} \\ &= U^*(t_1)(\{0, 1, 12\} + t_2\{0, 2\} + t_2^2\{0\}) \\ &= (\{0, 1, 12\} + t_1\{0, 2\} + t_1^2\{0\}) + t_2(\{0, 2\} + t_1\{0\}) + t_2^2(\{0\} + t_1\emptyset), \end{aligned}$$

$s^{U^*}_{\{0,1,12\},\emptyset}(t_1,t_2) = s^{\{0,1,12\},\emptyset}_U(t_1,t_2) = t_1t_2^2$ is not symmetric.

For a labeling φ from T to $\{1,\dots,m\}$, we define $t^\varphi = \prod_{w\in T} t_{\varphi(w)}$. For a tree T, it follows that

$$s^{T,\emptyset}_U(t_1,\dots,t_n) = \sum_{\varphi;\text{ a binary-searching labeling}} t^\varphi,$$

$$s^{D}_{T,\emptyset}(t_1,\dots,t_n) = \sum_{\varphi;\text{ a right-strictly-increasing labeling}} t^\varphi,$$

$$s^{D'}_{T,\emptyset}(t_1,\dots,t_n) = \sum_{\varphi;\text{ a left-strictly-increasing labeling}} t^\varphi.$$

These generalized Schur polynomials $s^{T,\emptyset}_U(t_1,\dots,t_n)$ in this case are the commutativizations of the basis elements $\mathbf{P}_T$ of **PBT** in Hivert-Novelli-Thibon [**7**].

References

[1] Fomin,S., Generalized Robinson-Schensted-Knuth correspondence, Zar. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 155(1986), 156–175,195 (Russian); English transl., J.Soviet Math. 41(1988), 979–991.
[2] Fomin,S., Duality of graded graphs, J. Algebra Combin. 3 (1994) 357-404.
[3] Fomin,S., Schensted algorithms for dual graded graphs, J. Algebraic Combin. 4(1995),5–45.
[4] Fomin,S., Schur Operators and Knuth Correspondences, J. of Combinatoric theory, Series A. 72(1995),277-292.
[5] Fulton,W., Young Tableaux; with applications to representation theory and geometry, volume 35 of London Mathematical Society Student Texts, Cambridge University Press,New York, 1997.
[6] Gessel,Ira M., Counting paths in Young's lattice, J. Statistical planning and inference. 34(1993),125–134.
[7] Hivert, F., Novelli, J., and Thibon, J. 2005. The algebra of binary search trees. Theor. Comput. Sci. 339, 1 (Jun. 2005), 129-165.
[8] Roby,T., Applications and extensions of Fomin's generalization of the Robinson-Schensted correspondence to differential posets, Ph.D.thesis, M.I.T., 1991.
[9] Stanley,R., Differential posets, J.American Math. Soc, 1(1988), 919–961.
[10] Stanley,R., Variations on differential posets, Invariant theory and tableaux (Stanton,D.,ed.), IMA volumes in mathematics and its applications, Springer-Verlag, New York, pp.145–165.

Department of Mathematics, Hokkaido University, Japan
E-mail address: nu@math.sci.hokudai.ac.jp
URL: http://www.math.sci.hokudai.ac.jp/~nu/

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Pattern avoiding doubly alternating permutations

Erik Ouchterlony

Abstract. We study pattern avoiding doubly alternating (DA) permutations, i.e., alternating (or zigzag) permutations whose inverse is also alternating. We exhibit a bijection between the 1234-avoiding permutations and the 1234-avoiding DA permutations of twice the size using the Robinson-Schensted correspondence. Further, we present a bijection between the 1234- and 2134-avoiding DA permutations and we prove that the 2413-avoiding DA permutations are counted by the Catalan numbers.

Résumé. Nous étudions les permutations qui évitent les motifs double-alternant (DA), c'est à dire, les permutations alternantes dont l'inverse est alternante. Nous montrons, en utilisant le correspondance de Robinson-Schensted, une bijection entre les permutations de longueur n évitant 1234 et les permutations DA de longueur $2n$ évitant 1234. Nous montrons aussi, une bijection entre les permutations DA évitant 1234 et celles évitant 2134, et que les permutations DA évitant 2413 sont dénombrées par les nombres de Catalan.

1. Introduction

A permutation $\sigma \in \mathcal{S}_n$ is said to *contain* the pattern $\tau \in \mathcal{S}_m$ if there is a subsequence of (the word representation of) σ which is order equivalent to (the word representation of) τ. To distinguish between patterns and other permutations, we will use slightly different notation. For example, the permutation $(1, 3, 2, 4)$ will be written as 1324 if it is used as a pattern. We will often use the matrix representation of σ, which is the $n \times n$ 0-1-matrix having ones in the positions with matrix coordinates $(i, \sigma(i))$. It can also be written as $([\![\sigma(i) = j]\!])_{i,j=1}^{n}$, using Iverson's bracket notation [**9**] for the characteristic function, $[\![S]\!] = 1$ if S is true and 0 otherwise. In the figures we will use dots instead of ones and leave the zeroes empty, as in Figure 1, to make the picture clearer. In this notation σ contains the pattern τ if some submatrix of (the matrix representation of) σ is equal to (the matrix representation of) τ. The permutations not containing τ are called τ-*avoiding*, and we write

$$\mathcal{S}_n(\tau_1, \tau_2, \ldots, \tau_t) = \{\sigma \in \mathcal{S}_n : \sigma \text{ is } \tau_i\text{-avoiding for all } i = 1, \ldots, t\}.$$

A word, e.g., a permutation, $w = (w_i)_{i=1}^{n}$, is *(up-down)-alternating* if $w_{2i-1} < w_{2i}$ and $w_{2i} \geqslant w_{2i+1}$ for all applicable i. This means that the word alternates between *rises* and *descents*, beginning with a rise. If it instead starts with a descent, it is called *down-up-alternating*.

A permutation σ is *doubly alternating* (DA) if both σ and σ^{-1} are alternating. The set of pattern avoiding doubly alternating permutations is denoted by

$$\mathrm{DA}_n(\tau_1, \ldots, \tau_t) = \{\sigma \in \mathcal{S}_n(\tau_1, \ldots, \tau_t) : \sigma \text{ is doubly alternating }\}.$$

Pattern avoiding permutations have been subject to much attention since the pioneering work by Knuth's [**10**], where he used them for studying stack sortable permutations. For a thorough summary of the current status of research, see Bóna's book [**4**]. Alternating permutations have a long history, they were studied already in the 19th century by André [**1**], and it is well know that they are counted by the tangent and secant numbers, also known as Euler numbers, E_k, and thus, their exponential generating function

2000 *Mathematics Subject Classification.* Primary: 05A05.

Key words and phrases. permutations, pattern avoidance, doubly alternating.

The author was supported by the European Commission's IHRP Programme, grant HPRN-CT-2001-00272, "Algebraic Combinatorics in Europe".

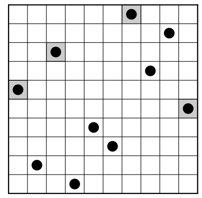


Figure 1. The permutation $(7, 9, 3, 8, 1, 10, 5, 6, 2, 4) \in \mathrm{DA}(1234)$ contains the pattern 3214, but avoids 1234.

is $\tan(x) + \sec(x)$. Alternating permutation avoiding patterns have been studied by Mansour [**11**], but there are still many open questions remaining.

The doubly alternating permutations were first counted by Foulkes [**6**], up to $n = 10$, using a result which we state as Theorem 5.1. The only formula known is due to Stanley [**13**]:

$$\sum_{n \text{ odd}} \mathrm{DA}_n\, x^n = \sum_{k \text{ odd}} E_k^2((\log(1+x)/(1-x))/2)^k/k!,$$
$$\sum_{n \text{ even}} \mathrm{DA}_n\, x^n = (1-x^2)^{-1/2} \sum_{k \text{ even}} E_k^2((\log(1+x)/(1-x))/2)^k/k!,$$

from which we get that the first few numbers for $\mathrm{DA}_n, n \geqslant 1$, are

$$1, 1, 1, 2, 3, 8, 19, 64, 880, 3717, 18288, 92935, \ldots$$

The motivation for studying doubly alternating permutations came from work by Guibert and Linusson [**8**] who showed that doubly alternating Baxter permutations are counted by the Catalan numbers. It was a natural step to study other restrictions to see whether interesting results could be found.

Using computer enumerations Guibert came up with several conjectures that indicated there are surprising connections between doubly alternating permutations and ordinary permutations. Some of these are proved in this paper, see proposition 4.1 and Theorems 5.2 and 6.2, whereas others still remain unproved and are listed in conjecture 7.1.

In this paper we study doubly alternating permutations avoiding patterns of lengths three and four. The patterns of length three are covered in Section 3. In Section 4, we show that doubly alternating permutations avoiding 2413 are counted by Catalan numbers, and are closely related to the doubly alternating Baxter permutations. Section 5 contains a bijection between $\mathrm{DA}_{2n}(1234)$ and $\mathcal{S}_n(1234)$ and in Section 6 we use a result by Babson and West [**2**] to construct a bijection between $\mathrm{DA}_n(12\tau)$ and $\mathrm{DA}_n(21\tau)$, where τ is any permutation of $\{3, 4, \ldots, m\}$, $m \geqslant 3$. In Section 7 other patterns giving the same sequence are investigated and in the final section some remarks on a few DA permutations avoiding two patterns of length four are given.

I like to thank Olivier Guibert for introducing me to the problem and for interesting discussions. Thanks also to Svante Linusson, Bruce Sagan and Mark Dukes for numerous comments and suggestions.

2. Notation and basic facts

First we define the *reverse*, the *complement* and the *rotation* of a permutation σ,

$$\sigma^r = (\sigma(n+1-i))_{i=1}^n$$
$$\sigma^c = (n+1-\sigma(i))_{i=1}^n$$
$$\sigma^\# = (\sigma^c)^r = (\sigma^r)^c = (n+1-\sigma(n+1-i))_{i=1}^n$$

In terms of matrices, the first two correspond to flipping the matrix vertically and horizontally, respectively, whereas the last operation rotates the matrix 180 degrees. However, these bijections do not in general preserve the doubly alternating property, which means that we lose some symmetry compare with ordinary permutations, so that more genuinely different patterns need to be examined. However, it is obvious from the definition that inverting and, if n is even, rotating a permutation does preserve the property of being doubly alternating.

Lemma 2.1.

(a) $\sigma \in \mathrm{DA}_n \iff \sigma^{-1} \in \mathrm{DA}_n$
(b) $\sigma \in \mathrm{DA}_{2n} \iff \sigma^{\#} \in \mathrm{DA}_{2n}$

Another simple, but very useful, property that follows from the DA condition is that some areas on the border of the matrix can never have a dot, see Figure 2.

Lemma 2.2.

(a) *Let* $\sigma \in \mathrm{DA}_{2n}$, *then*
  (i) $\sigma(1)$ *is odd,*
  (ii) $\sigma(2) \in \{3, 5, 7, \dots, 2n-1, 2n\}$,
  (iii) $\sigma(2) = 2n$ *iff* $\sigma(1) = 2n-1$,
  (iv) $\sigma(2n)$ *is even,*
  (v) $\sigma(2n-1) \in \{1, 2, 4, 6, \dots, 2n-2\}$,
  (vi) $\sigma(2n-1) = 1$ *iff* $\sigma(2n) = 2$.

(b) *Let* $\sigma \in \mathrm{DA}_{2n+1}$, *then*
  (i) $\sigma(1)$ *is odd and less than* $2n+1$,
  (ii) $\sigma(2)$ *is odd and greater than* 1,
  (iii) $\sigma(2n+1)$ *is even,*
  (iv) $\sigma(2n) \in \{4, 6, 8, \dots, 2n, 2n+1\}$,
  (v) $\sigma(2n) = 2n+1$ *iff* $\sigma(2n+1) = 2n$.

Proof. First for the case a(i), if $\sigma(1) = k > 1$, then $\sigma^{-1}(k) = 1$, so $\sigma^{-1}(k-1) > 1$, which implies that k is odd, since $\sigma \in \mathrm{DA}$. For a(ii), assume $\sigma(2) = m < 2n$, $m > \sigma(1) \geqslant 1$. Then $\sigma^{-1}(m-1) > 2$ or $\sigma^{-1}(m+1) > 2$, so m is odd. The equivalence a(iii) is a direct consequence of the definition of doubly alternating, since $\sigma(2) > \sigma(1)$ and $\sigma^{-1}(2n-1) < \sigma^{-1}(2n)$. The other cases are similar. ☐

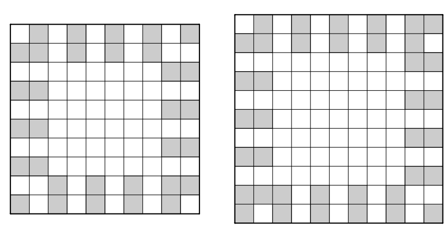


Figure 2. Illustration of Lemma 2.2. Shaded areas are forbidden.

Note that this lemma could be applied to σ^{-1} as well, because of Lemma 2.1. From the two lemmas it is clear that there is a difference between odd and even sizes, so they require separate treatment in many of the proofs. This disparity is also reflected in the fact that $\mathrm{DA}_n(\tau_1, \tau_2, \dots, \tau_t)$ is not an increasing function of n for all patterns. Some counterexamples are, $\mathrm{DA}_4(321) = 2 > 1 = \mathrm{DA}_5(321)$, $\mathrm{DA}_6(2431) = 6 > 5 = \mathrm{DA}_7(2431)$ and, if Conjecture 8.1 is true, $\mathrm{DA}_{27}(1234, 2134) = 2681223 > 2674440 = \mathrm{DA}_{28}(1234, 2134)$.

3. Patterns of length three

For normal permutations, patterns of length three are the first non-trivial cases; they are all counted by the Catalan numbers. However, for the doubly alternating permutations, it turns out that all the patterns of length three are (more or less) trivial.

Proposition 3.1.

(i) $|\,\mathrm{DA}_n(123)| = |\,\mathrm{DA}_n(213)| = |\,\mathrm{DA}_n(231)| = |\,\mathrm{DA}_n(312)| = 1$
(ii) $|\,\mathrm{DA}_n(132)| = [\![n \text{ even or } n = 1]\!]$
(iii) $|\,\mathrm{DA}_n(321)| = 1 + [\![n \text{ even and } n \geqslant 4]\!]$

Proof. Omitted in the extended abstract. ☐

4. 2413-avoiding doubly alternating permutations

The doubly alternating 2413-avoiding permutations were conjectured by Guibert to be counted by the Catalan numbers. We prove this by showing them to possess a fairly simple block structure. First we need a technical lemma.

Lemma 4.1. *Let $\sigma \in \mathrm{DA}_n(2413)$, then*

(i) n *odd* $\implies$ $\sigma(1) = 1$

(ii) n *even* $\implies$ $\sigma(1) = 1$ *and* $\sigma(n) = n$ *or there is a* k, $2 < k \leqslant n-1$, *such that* $\sigma(i) > \sigma(j)$ *for all* $i < k \leqslant j$.

Proof. We can assume that $n \geqslant 3$, the smaller cases are trivial. Let $a = \sigma(1)$, and assume $\sigma(1) \neq 1$. By Lemma 2.2, a must be odd. Now let $b = \sigma(\beta)$, where β is the smallest number such that $\sigma(\beta) < a$. Note that $\beta \geqslant 3$, since $\sigma(2) > a$. Also, β is odd since $\sigma(\beta - 1) > \sigma(\beta)$.

Let c be the largest number such that $\gamma = \sigma^{-1}(c) < \beta$, see Figure 3. Thus $c > a$ and c must be odd or $c = n$, since $\sigma^{-1}(c+1) > \beta > \sigma^{-1}(c) = \gamma$ if $c < n$.

If $\kappa > \beta$, then $\sigma(\kappa) < a$ or $\sigma(\kappa) > c$, otherwise we get the 2413 pattern. Therefore the rectangle, with NW corner $(2, a+1)$ and SE corner $(\beta - 1, c)$ contains exactly one dot in each row and column, so it is square, and hence $c = a + \beta - 2$ is even and thus $c = n$ is the only possibility. This proves the first assertion.

If n is even, $\sigma(1) = 1$ implies $\sigma(n)$, by rotational symmetry, so the second assertion follows from $i < \beta \leqslant j \Rightarrow \sigma(j) < a \leqslant \sigma(i)$. □

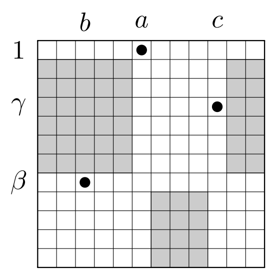


Figure 3. Illustration of Lemma 4.1, the shaded areas are empty.

As a direct consequence, we get the following corollary, which gives a very explicit description of what the DA 2413-avoiding permutations look like.

Corollary 4.2.

(i) $\sigma \in \mathrm{DA}_{2n+1}(2413)$ *iff* $\sigma = (1, \tilde{\sigma})$, *where* $(\tilde{\sigma}^r)^{-1} \in \mathrm{DA}_{2n}(2413)$.

(ii) $\sigma \in \mathrm{DA}_{2n}(2413)$ *iff the permutation matrix of* σ *is a block matrix, where all but the anti-diagonal blocks are empty. Any non-empty block* ν *has even size,* $2k$, *and can be written* $\nu = (1, \tilde{\nu}, 2k)$, *where* $(\tilde{\nu}^r)^{-1} \in \mathrm{DA}_{2k-2}(2413)$.

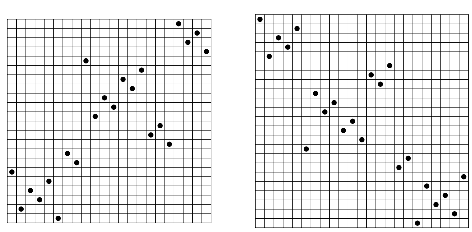


Figure 4. Example of the block structure of 2413-avoiding DA permutations.

The block structure condition in the corollary is in fact invariant under taking inverses, even though the pattern 2413 is not, therefore we get the following, slightly surprising result: $\mathrm{DA}_n(2413)$ and $\mathrm{DA}_n(3142)$ are not only the same size, but are actually the same sets. Therefore also $\mathrm{DA}_n(2413, 3142)$ is the same set. They are all counted by the Catalan numbers.

Proposition 4.1. $|\mathrm{DA}_n(2413)| = C_{\lfloor n/2 \rfloor}$.

Proof. First if n is odd, we have as a direct consequence of Corollary 4.2(i) that $|\mathrm{DA}_n(2413)| = |\mathrm{DA}_{n-1}(2413)|$. If n is even, then Corollary 4.2(ii) tells us that $\sigma \in \mathrm{DA}_n(2413)$ can be factored into blocks $\sigma_1, \sigma_2, \ldots, \sigma_m$, where $\sigma_i = (1, \tilde{\sigma}_i, |\sigma_i|)$ and $\tilde{\sigma}_i^r \in \mathrm{DA}(2413)$.

Let $D(x)$ be the generating function $D(x) = \sum_k |\mathrm{DA}_{2k}(2413)|x^k$. Then

$$D(x) = \sum_{i=0}^{\infty}(xD(x))^i = \frac{1}{1 - xD(x)}$$

which implies $xD(x)^2 - D(x) + 1 = 0$, i.e., the well know equation for the generating function of the Catalan numbers. Since $D(0) = 1$, we get $|\mathrm{DA}_{2k}(2413)| = C_k$. $\square$

Another way to prove this is to construct a bijection with Dyck paths. We define $\Theta : \mathrm{DA}_{2n}(2413) \leftrightarrow \{\text{Dyck paths of length } 2n\}$ recursively, by using Corollary 4.2.

(i) $\Theta(\emptyset) = \emptyset$

(ii) If σ consists of a single block, so that $\sigma = (1, \tilde{\sigma}, 2n)$, then $\Theta(\sigma)$ is the Dyck path starting with a rise, ending with a descent and having the Dyck path $\Theta(\tilde{\sigma}^r)$ as the middle part.

(iii) If σ can be factored into k blocks, $\sigma_1, \ldots, \sigma_k$ (starting with the leftmost block), then $\Theta(\sigma)$ is the concatenation of the Dyck paths $\Theta(\sigma_1), \ldots, \Theta(\sigma_k)$.

The inverse is similarly defined, using recursion.

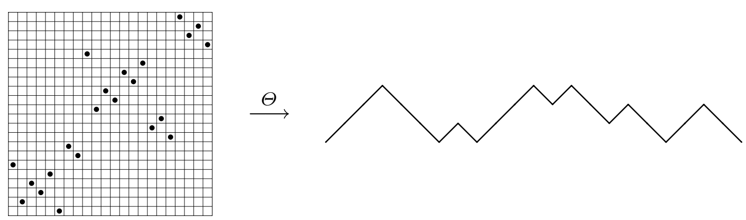


Figure 5. Example of the bijection between $\mathrm{DA}_{2n}(2413)$ and Dyck paths.

4.1. Doubly alternating Baxter permutations.

A *Baxter permutation* is defined to be a permutation, $\sigma = (\sigma_i)_{i=1}^n$, such that for all $1 \leqslant i < j < k < l \leqslant n$,

$$\begin{aligned} \sigma_i + 1 = \sigma_l \text{ and } \sigma_j > \sigma_l &\implies \sigma_k > \sigma_l \quad \text{and} \\ \sigma_l + 1 = \sigma_i \text{ and } \sigma_k > \sigma_i &\implies \sigma_j > \sigma_i. \end{aligned}$$

It is clear from this definition that if σ avoids both 2413 and 3142 then it is a Baxter permutation, so we have

$$\mathrm{DA}_n(2413, 3142) \subset \{\sigma \in \mathrm{DA}_n : \sigma \text{ is Baxter}\}.$$

However, in [**8**], Guibert and Linusson showed that the doubly alternating Baxter permutations are counted by the Catalan numbers, so the sets must in fact be the same:

Corollary 4.3.

$$\{\sigma \in \mathrm{DA}_n : \sigma \textit{ is Baxter}\} = \mathrm{DA}_n(2413, 3142) = \mathrm{DA}_n(2413) = \mathrm{DA}_n(3142).$$

It is also possible to prove this directly, without referring to the result by Guibert and Linusson.

Lemma 4.4. $\{\sigma \in \mathrm{DA}_n : \sigma \textit{ is Baxter}\} \subset \mathrm{DA}_n(2413)$.

Proof. Assume σ is Baxter, but not 2413-avoiding, and d_1, d_2, d_3, d_4, with $d_k = (i_k, j_k)$, constitute a 2413 pattern, such that $j_4 - j_1$ is as small as possible and given j_1 and j_4, $i_3 - i_2$ is as small as possible , as in Figure 6. The four areas shaded in the figure are empty, otherwise we would use one of those dots for the 2413-pattern. Now, let ν be the permutation having as permutation matrix the submatrix of σ, consisting of the rows $i_2 + 1, i_2 + 2, \ldots, i_3 - 1$ and columns $j_1 + 1, j_1 + 2, \ldots, j_4 - 1$. Since σ is Baxter, ν cannot be empty. The DA condition implies that ν is up-down-alternating, ν^{-1} is down-up-alternating and $|\nu|$ is even.

Also ν is 2413-avoiding, since otherwise this occurrence of 2413 would have been used instead of $d_1, \ldots, d_4$, and $\nu(1) \neq 1$, since ν^{-1} is down-up-alternating.

Rehashing the argument for Lemma 4.1, we can see that in fact no such permutation ν can exist. Defining a, b, c, β and γ in the same way as in the proof of Lemma 4.1 we get once again Figure 3. The difference is that now a and c must be even, whereas β is still odd. But then $c = a + \beta - 2$ is odd, a contradiction. $\square$

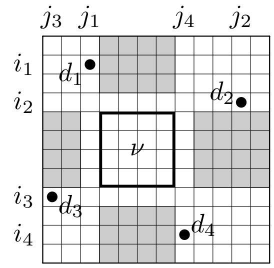


Figure 6. Illustration of Lemma 4.4. The shaded areas do not contain any dots.

5. 1234-avoiding doubly alternating permutations

In this section we construct a bijection between the doubly alternating 1234-avoiding permutations of size $2n$ and the ordinary 1234-avoiding permutations of size n by using the Robinson-Schensted correspondence. Let λ be a *Young diagram*, and denote by $\mathrm{SYT}(\lambda)$ the set of standard Young tableaux of shape λ. For a standard Young tableaux, T, let $\mathsf{row}_k(T)$ ($\mathsf{col}_k(T)$) denote the number of the row (column) for the entry k, counting from the top row (leftmost column), which is given the number one. The vector $\mathsf{row}(T)$ ($\mathsf{col}(T)$) is called the *row (column) reading* of T.

We define the set of *alternating standard tableaux* as

$$\begin{aligned}\mathrm{Alt}(\lambda) =& \{T \in \mathrm{SYT}(\lambda) : \mathsf{col}(T) \text{ is up-down-alternating}\}\\ =& \{T \in \mathrm{SYT}(\lambda) : \mathsf{row}(T) \text{ is down-up-alternating}\},\end{aligned}$$

where the second equality is a consequence of the relative positions of two consecutive entries in a standard tableau, as shown in Figure 7.

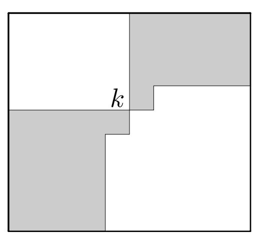


Figure 7. The shaded area denotes the possible positions for the entry $k+1$, relative to the entry k, in a standard Young tableau.

The Robinson-Schensted correspondence, RSK, is a well known bijection between a permutation and a pair of standard Young tableaux of the same shape, see for example the book by Fulton [**7**]. An interesting fact is that doubly alternating permutations can be recognised by their RSK tableaux: They are both alternating if and only if the permutation is DA. In fact, Foulkes [**5**] proved a more general theorem, in which he counts the number of permutations with any given sequences of ups and downs for the permutation and its inverse. The following lemma is the key for proving Foulkes theorem.

Lemma 5.1. *Let $\sigma \in \mathcal{S}_n$, $\mathrm{RSK}(\sigma) = (P, Q)$ and $1 \leqslant k < n$, then*

$$k \text{ comes before } k+1 \text{ in } \sigma \quad \Longleftrightarrow \quad \mathsf{row}_k(P) \geqslant \mathsf{row}_{k+1}(P).$$

Proof. First assume k is inserted before $k+1$ in the RSK bumping process. This means that $k+1$ can never end up below k, since whenever they are in the same row only k can be bumped down.

For the converse, assume $k+1$ is inserted before k. Then k will always be strictly above $k+1$, since if k is in the row exactly above $k+1$ and is being bumped down (or $k+1$ is in the first row and k is about to be inserted), k has to bump down $k+1$ in the next step and thus stay above. □

Let the *signature* of a word, $w = w_1 w_2 \dots w_n$, be a sequence of $+$'s and $-$'s which has a $+$ in position i iff $w_i < w_{i+1}$. For example, signature$(4,1,5,5,6,2,2) = (-,+,-,+,-,-)$. We now get Foulkes theorem as a consequence of Lemma 5.1 and the fact that $\mathrm{RSK}(\sigma) = (P,Q)$ iff $\mathrm{RSK}(\sigma^{-1}) = (Q,P)$:

theorem 5.1 (Foulkes). *Let $\sigma \in \mathcal{S}_n$ and $\mathrm{RSK}(\sigma) = (P,Q)$. Then*

$$\begin{aligned} \mathrm{signature}(\sigma^{-1}) &= \mathrm{signature}(\mathsf{col}(P)) = -\,\mathrm{signature}(\mathsf{row}(P)), \\ \mathrm{signature}(\sigma) &= \mathrm{signature}(\mathsf{col}(Q)) = -\,\mathrm{signature}(\mathsf{row}(Q)). \end{aligned}$$

The doubly alternating permutations are a special case:

Corollary 5.2. *Let $\sigma \in \mathcal{S}_n$ and $\mathrm{RSK}(\sigma) = (P,Q)$. Then*

$$\sigma \in \mathrm{DA}_n \quad \Longleftrightarrow \quad P, Q \in \mathrm{Alt}(\lambda).$$

Let T be an alternating standard tableau with $2n$ entries and at most three columns. We define the *pair column reading* $\mathsf{colpair}(T) = (w_i)_{i=1}^n$, where $w_i = \mathsf{col}_{2i-1}(T) + \mathsf{col}_{2i}(T) - 2$, i.e, $(1,2) \mapsto 1, (1,3) \mapsto 2$ and $(2,3) \mapsto 3$, since Lemma 5.1 tells us that the only possibilities for the pairs are $(1,2), (1,3)$ and $(2,3)$.

Let $w = (w_i)_{i=1}^l$ be a word, and $\mathsf{weight}(w) \stackrel{\mathrm{def}}{=} (|\{i : w_i = k\}|)_{k \geqslant 1}$ be the *weight* vector of w. We call w *Yamanouchi* (or a *ballot sequence*) if the weight of each prefix of w is a partition, i.e., it is weakly decreasing.

Lemma 5.3. $\mathsf{colpair}$ *is a bijection between alternating standard tableaux with $2n$ elements and at most three columns and Yamanouchi words of length n on three letters.*

Proof. Let T be an alternating standard tableaux, with at most three columns. Then $\mathsf{colpair}(T) = (w_i)_{i=1}^n$ is, as noted above, a word on the letters 1,2 and 3, so we need to show that $\mathsf{colpair}(T)$ is Yamanouchi iff $\mathsf{col}(T)$ is an alternating Yamanouchi word.

First assume $\mathsf{colpair}(T)$ is Yamanouchi and let $v = (w_i)_{i=1}^k$ be an arbitrary prefix of $\mathsf{colpair}(T)$. Then $\mathsf{weight}(v) = (a,b,c)$, is a partition, i.e., $a \geqslant b \geqslant c$. Hence $\mathsf{weight}((\mathsf{col}_j(T)_{j=1}^{2k}) = (a+b, a+c, b+c)$ is also a partition. Since $\mathsf{col}(T)$ is alternating and $\mathsf{weight}((\mathsf{col}_i(T))_{i=1}^{2k+1})$ is a partition if $\mathsf{weight}((\mathsf{col}_i(T))_{i=1}^{2k+2})$ is, it follows that $\mathsf{col}(T)$ is an alternating Yamanouchi word.

For the converse, assume $\mathsf{col}(T)$ is an alternating Yamanouchi word and let $u = (\mathsf{col}_i(T))_{i=1}^{2k}$ be a prefix of $\mathsf{col}(T)$. Then $\mathsf{weight}(u) = (d,e,f)$ is a partition, so $\mathsf{weight}((\mathsf{colpair}_i(T))_{i=1}^k) = \frac{1}{2}(d+e-f, d+f-e, e+f-d)$ is a partition, which proves that $\mathsf{colpair}(T)$ is Yamanouchi. □

Now we are ready to combine the bijections to get the bijection $\Phi : \mathcal{S}_n(1234) \to \mathrm{DA}_{2n}(1234)$, defined by

$$\Phi(\sigma) = \mathrm{RSK}^{-1}(\mathsf{colpair}^{-1}(\mathsf{col}(P)), \mathsf{colpair}^{-1}(\mathsf{col}(Q)),$$

where $\mathrm{RSK}(\sigma) = (P,Q)$. See Figure 8 for an illustrative example.

theorem 5.2. *Φ is a bijection, hence*

$$|\,\mathrm{DA}_{2n}(1234)| = |\,\mathcal{S}_n(1234)|.$$

Proof. Let $\sigma \in \mathcal{S}_n(1234)$ and $\mathrm{RSK}(\sigma) = (P,Q)$. RSK is a bijection between permutations and pairs of standard tableaux of the same shape such that if the permutation is 1234-avoiding iff the shape does not have more than three columns. From the definitions we know that $\mathsf{col}(P)$ and $\mathsf{col}(Q)$ are Yamanouchi words, so, by Lemma 5.3, $\mathsf{colpair}^{-1}(\mathsf{col}(P))$ and $\mathsf{colpair}^{-1}(\mathsf{col}(Q))$ are alternating standard tableaux with at most three columns. Their shapes are the same since the weights of $\mathsf{col}(P)$ and $\mathsf{col}(Q)$ are the same, which is a consequence of P and Q having the same shape. Applying the inverse of RSK and using Corollary 5.2, we get that $\Phi(\sigma) \in \mathrm{DA}_{2n}(1234)$.

The converse is similar. □

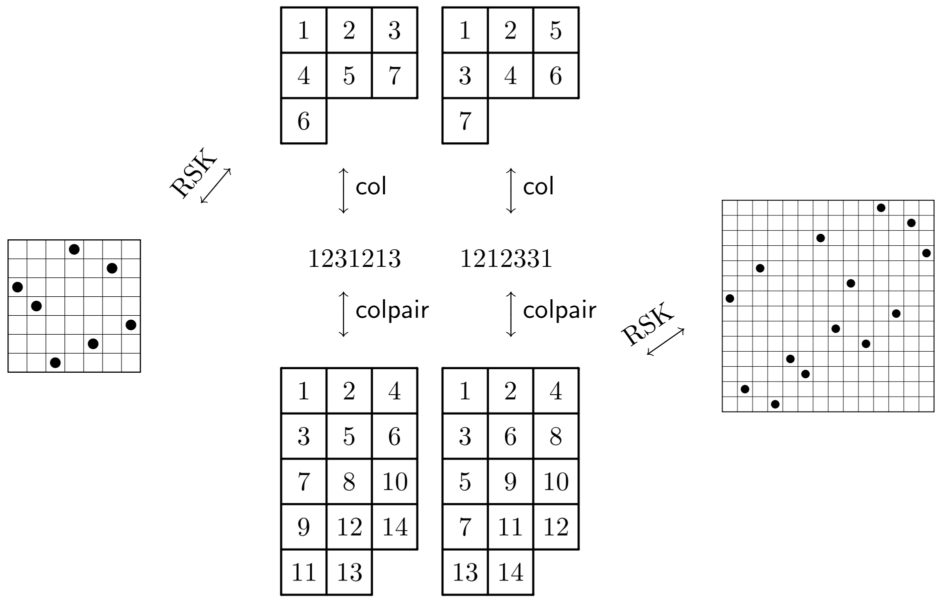


FIGURE 8. Example of the bijection $\Phi : \mathcal{S}_n(1234) \to \mathrm{DA}_{2n}(1234)$.

6. 21τ-avoiding doubly alternating permutations

The goal of this section is to find a bijection between $\mathrm{DA}_n(12\tau)$ and $\mathrm{DA}_n(21\tau)$, where τ is any permutation of $\{3, 4, \dots, m\}$, $m \geqslant 3$, by using a well known bijection due to Babson and West [**2**]. The problem is that it is far from obvious that it will preserve the property of being doubly alternating. To show this we need a few definitions. During this section we assume τ to be fixed.

A dot, d, is called *active* if d is the 1 or 2 in any 12τ or 21τ pattern in σ and other dots are called *inactive*. Also the pair of dots, (d_1, d_2), is called an *active pair* if $d_1 d_2$ is the 12 in a 12τ-pattern or the 21 in a 21τ-pattern.

LEMMA 6.1. *Assume $\sigma \in \mathrm{DA}_n(12\tau) \cup \mathrm{DA}_n(21\tau)$ and $d = (i, j)$ is any active dot. Then i and j are odd.*

PROOF. First assume $\sigma \in \mathrm{DA}_n(12\tau)$ and that σ has a 21τ-pattern, otherwise there are no active dots. By inversion symmetry, we can assume that d is the 1 in a 21τ pattern. If $j = 1$, i.e. $\sigma^{-1}(1) = j$, then j is odd by Lemma 2.2, and if $j > 1$, then $\sigma^{-1}(j-1) > \sigma^{-1}(j)$, since a dots to the north-west of d would give a 12τ pattern. Hence j is odd. Also, to avoid the 12τ, $\sigma(i-1) > \sigma(i)$, so i is odd as well.

Now assume instead $\sigma \in \mathrm{DA}_n(21\tau)$. Let $d_1, d_2, \dots, d_m$, be the dots in a 12τ pattern, with $d_k = (i_k, j_k)$. If i_1 is even then there is a descent from i_1 to $i_1 + 1$ and so the corresponding points along with tau will make the forbidden pattern. So i_1 is odd and the same argument applies to i_2, j_1, and j_2. $\square$

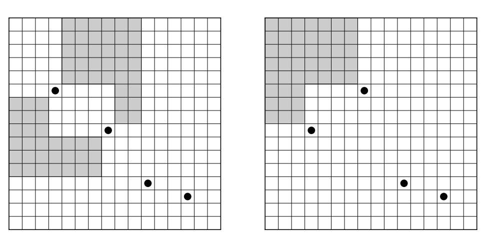


FIGURE 9. Illustration of the proof for Lemma 6.1, with $\tau = (3, 4)$. Shaded areas are forbidden.

We now define a Young diagram, λ_σ, consisting of the part of the board which contains the active dots. For a pair of dots, d_1, d_2, let R_{d_1,d_2} to be the smallest rectangle with top left coordinates $(1, 1)$, such that

$d_1, d_2 \in R_{d_1,d_2}$. Define

$$\lambda_\sigma \stackrel{\text{def}}{=} \bigcup R_{d_1,d_2},$$

where the union is over all active pairs (d_1, d_2). It is clear from the definition that λ_σ is indeed a Young diagram (see Figure 10).

A *rook placement* (also known as *traversal* or *transversal*) of a Young diagram, λ, is a placement of dots, such that all rows and columns contain exactly one dot. If some of the rows or columns are empty we call it a *partial rook placement*. Furthermore, we say that a rook placement on λ avoids the pattern τ if no rectangle, $R \subset \lambda$, contain τ.

The definition of λ_σ implies the following useful fact:

LEMMA 6.2. *Let $\sigma \in \mathrm{DA}_n$ and $\mathrm{rp}(\lambda_\sigma)$ be the partial rook placement on λ_σ induced by σ. Then*

$$\begin{aligned} \sigma \in \mathrm{DA}_n(12\tau) \quad &\Longleftrightarrow \quad \mathrm{rp}(\lambda_\sigma) \textit{ is } 12\textit{-avoiding,} \\ \sigma \in \mathrm{DA}_n(21\tau) \quad &\Longleftrightarrow \quad \mathrm{rp}(\lambda_\sigma) \textit{ is } 21\textit{-avoiding.} \end{aligned}$$

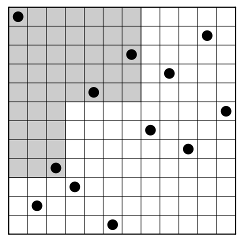


FIGURE 10. Example of λ_σ for $\sigma = (1, 11, 7, 9, 5, 12, 8, 10, 3, 4, 2, 6)$ and $\tau = (3, 4)$.

The bijection we will use is due to Babson and West [**2**], which built on work by Simion and Schmidt [**12**] and West [**14**]. But we give here the more general result by Backelin, West, and Xin [**3**].

THEOREM 6.1 (Backelin, West, Xin). *Let τ be any permutation of $\{t+1, t+2, \ldots, m\}$. Then for every Young diagram λ, the number of $(t, t-1, \ldots, 1, \tau)$-avoiding rook placements on λ equals the number of $(1, 2, \ldots, t, \tau)$-avoiding rook placements on λ.*

We call two permutations of the same size *a-equivalent* if all the inactive dots are the same, and write $\sigma_1 \sim_a \sigma_2$. We shall see in Lemma 6.4 that this implies $\lambda_{\sigma_1} = \lambda_{\sigma_2}$.

LEMMA 6.3. *If $\sigma \in \mathrm{DA}_n(12\tau) \cup \mathrm{DA}_n(21\tau)$ and $\nu \sim_a \sigma$, then ν is doubly alternating.*

PROOF. Let $d = (i, j) \in \nu$ be a dot in an odd row, so that, by Lemma 6.1, both the dots in row $i-1$ and row $i+1$ are inactive (if they exists). If d is inactive then all three dots also belong to σ so $\nu(i-1) > \nu(i) < \nu(i+1)$. If d is active there is a τ-pattern to the SE of d, so if either of the dots in row $i-1$ or row $i+1$ are to the left of d, then this dot is active, since it creates either a 12τ or 21τ pattern together with d and the τ pattern, giving a contradiction, so again $\nu(i-1) > \nu(i) < \nu(i+1)$. The same applies, by symmetry, to dots in odd columns. $\square$

LEMMA 6.4. *If $\sigma, \nu \in \mathrm{DA}_n$, then*

$$\sigma \sim_a \nu \quad \Longrightarrow \quad \lambda_\sigma = \lambda_\nu.$$

PROOF. Let $\sigma \sim_a \nu$ be two arbitrary a-equivalent DA permutations and assume $s = (i, j)$ is a SE corner of λ_σ. We need to show that $s \in \lambda_\nu$, so that $\lambda_\sigma \subseteq \lambda_\nu$ and thus, since $\sim_a$ is reflexive, $\lambda_\nu = \lambda_\sigma$.

Let $R_{d_1,d_2} \subset \lambda_\sigma$ be a rectangle, such that $s \in R_{d_1,d_2}$. Such a rectangle must exist, otherwise could not s belong to λ_σ. Hence there is a τ-pattern to the SE of s and one of the d_k is in row i and one (possibly the same one) is in column j. But, as $\nu \sim_a \sigma$, they have the same inactive dots, so there must also exist a dot $d'_1 \in \lambda_\nu$ in row i and a dot $d'_2 \in \lambda_\nu$ in column j. If d'_1 is east of s or if d'_2 is south of s then $s \in \lambda_\nu$. Hence we can assume d'_1 and d'_2 to be weakly NW of s. If $d'_1 \neq d'_2$, then $s \in R_{d'_1,d'_2} \subset \lambda_\nu$, since the τ-pattern is still SE of s, and if $d'_1 = d'_2 = s$ then clearly $s \in \lambda_\nu$. $\square$

Now we are ready to construct a bijection $\varPsi : \mathrm{DA}_n(12\tau) \to \mathrm{DA}_n(21\tau)$. Let $\sigma \in \mathrm{DA}_n(12\tau)$, so that the restriction of σ to λ_σ is a partial 12-avoiding rook placement. By Theorem 6.1 (ignoring the empty rows and columns) and Lemma 6.4, there exists a unique 21-avoiding (partial) rook placement on λ_σ, with the same rows and columns empty, which we combine with the inactive dots of σ to get $\varPsi(\sigma)$. By Lemma 6.3, $\varPsi(\sigma)$ is DA, and Lemma 6.2 says that it avoids 21τ. It is also clear from Theorem 6.1 that it is indeed a bijection. We have thus bijectively shown:

Theorem 6.2. *Let τ be any permutation of $\{3, 4, \ldots, m\}$, $m \geqslant 3$. Then*

$$|\,\mathrm{DA}_n(21\tau)| = |\,\mathrm{DA}_n(12\tau)|.$$

As a special case we have

Corollary 6.5. $\qquad |\,\mathrm{DA}_n(2134)| = |\,\mathrm{DA}_n(1234)|.$

7. Other patterns with the same number sequence as $\mathcal{S}_n(1234)$

By examining all the patterns of length four with computer, Guibert found 15 different cases that all seem to give rise to the same sequence, $|\,\mathcal{S}_n(1234)|$. Using Theorems 5.2 and 6.2, inversion, rotation and Proposition 7.1 below, we get altogether ten bijections, see Figure 11. However, to prove that all of them are indeed the same we would need five more bijections. In fact, we conjecture that the number of permutations are the same in all the cases given below.

Conjecture 7.1 (Guibert).

$$\begin{aligned}
|\,\mathrm{DA}_{2n}(1234)| &= |\,\mathrm{DA}_{2n+1}(1243)| \\
&= |\,\mathrm{DA}_{2n}(1432)| \\
&= |\,\mathrm{DA}_{2n+1}(1432)| \\
&= |\,\mathrm{DA}_{2n}(2341)| \\
&= |\,\mathrm{DA}_{2n}(3421)|
\end{aligned}$$

One can note that many of the patterns in the conjecture are of the same type as treated in Theorem 6.1, but the proof does not work here, except for 2134, since the bijections destroy the DA property.

Proposition 7.1. $\qquad |\,\mathrm{DA}_{2n}(2143)| = |\,\mathrm{DA}_{2n+1}(3412)| = |\,\mathrm{DA}_{2n+2}(3412)|.$

Proof. Let $\sigma \in \mathrm{DA}_n(3412)$, with $n \geqslant 4$. If $\sigma(1) > 1$, we get the forbidden pattern on the rows 1, 2, $\sigma^{-1}(1)$, $\sigma^{-1}(2)$, so $\sigma(1) = 1$. Let $\tilde{\sigma}$ be the permutation with the first row and column of σ removed. It is clear that if n is odd then $\tilde{\sigma}^c \in \mathrm{DA}_{n-1}(2143)$ iff $\sigma \in \mathrm{DA}_n(3412)$, and if n is even then $\tilde{\sigma}^{\#} \in \mathrm{DA}_{n-1}(3412)$ iff $\sigma \in \mathrm{DA}_n(3412)$. $\square$

$\mathcal{S}_n(1234)$	$\xleftrightarrow{\text{Th. 5.2}}$	$\mathrm{DA}_{2n}(1234) \xleftrightarrow{\text{Th. 6.2}} \mathrm{DA}_{2n}(2134) \xleftrightarrow{\#} \mathrm{DA}_{2n}(1243)$
	$\xleftrightarrow{\text{Th. 6.2}}$	$\mathrm{DA}_{2n}(2143) \xleftrightarrow{\text{Pr. 7.1}} \mathrm{DA}_{2n+1}(3412) \xleftrightarrow{\text{Pr. 7.1}} \mathrm{DA}_{2n+2}(3412)$
$\mathrm{DA}_{2n+1}(1243)$	$\xleftrightarrow{\text{Th. 6.2}}$	$\mathrm{DA}_{2n+1}(2143)$
$\mathrm{DA}_{2n}(1432)$	$\xleftrightarrow{\#}$	$\mathrm{DA}_{2n}(3214)$
$\mathrm{DA}_{2n}(2341)$	$\xleftrightarrow{-1}$	$\mathrm{DA}_{2n}(4123)$
$\mathrm{DA}_{2n}(3421)$	$\xleftrightarrow{\#}$	$\mathrm{DA}_{2n}(4312)$
$\mathrm{DA}_{2n+1}(1432)$		

Figure 11. Known bijections between the sequences conjectured to be $|\,\mathcal{S}_n(1234)|$.

8. Avoiding pairs of patterns of length four

When we have two patterns of length four, there are a huge number of cases. We have not yet studied many of these, but would like to give a flavour of what can happen by presenting one result and two conjectures. Combining the results in Sections 4 and 5 we get

PROPOSITION 8.1.

$$|\operatorname{DA}_n(1234,2413)| = \begin{cases} F_{n/2}, & \text{if } n \text{ is even,} \\ 2, & \text{if } n = 5, \\ 1, & \text{otherwise,} \end{cases}$$

where the F_n are the Fibonacci numbers.

PROOF. Let $\sigma \in \operatorname{DA}_{2n}(1234, 2413)$. By Corollary 4.2, σ can be factored into blocks, $\sigma_1, \sigma_2, \dots \sigma_k$. As σ avoids 1234 must each block be either 12 or 1324, since each of them have a dot in the NW corner and the SE corner. Hence

$$|\operatorname{DA}_{2n}(1234,2413)| = |\operatorname{DA}_{2n-2}(1234,2413)| + |\operatorname{DA}_{2n-4}(1234,2413)|,$$

and since $|\operatorname{DA}_0(1234,2413)| = |\operatorname{DA}_2(1234,2413)| = 1$, we get the Fibonacci numbers.

If $\sigma \in \operatorname{DA}_{2n+1}(1234,2413)$, then $\sigma(1) = 1$. Let $\tilde{\sigma} = (2n+1-\sigma(i))_{i=2}^{2n+1}$ be the permutation constructed from σ by removing the first row and column and then flipping horizontally. Then $\tilde{\sigma} \in \operatorname{DA}_{2n}(321, 2413)$, which by Proposition 3.1(iii) gives two possibilities if $2n \geqslant 4$, namely $(1,3,2,5,4,\dots,2n-1,2n-2,2n)$ and $(3,5,1,7,2,9,4,\dots,n,n-4,n-2)$. However, only the former avoids 2413 if $n \geqslant 6$, so we get the desired result. $\square$

The following two conjectures have been verified by computer calculations up to $n = 23$.

CONJECTURE 8.1.

$$|\operatorname{DA}_n(1234,3214)| = \begin{cases} F_{n-1}, & \text{if } n \text{ is even,} \\ 1, & \text{if } n = 1 \text{ or } n = 3, \\ F_{n-1} - F_{n-7}, & \text{otherwise.} \end{cases}$$

CONJECTURE 8.2.

$$|\operatorname{DA}_n(1234,2134)| = \begin{cases} C_{n/2}, & \text{if } n \text{ is even,} \\ 1, & \text{if } n = 1 \text{ or } n = 3, \\ C^{(4)}_{(n-5)/2}, & \text{otherwise.} \end{cases}$$

Here $C^{(4)}_n$ is the fourth difference of the Catalan numbers, defined recursively by $C^{(0)}_n = C_n$ and $C^{(i+1)}_n = C^{(i)}_{n+1} - C^{(i)}_n$. By collecting the terms and simplifying we get

$$\begin{aligned} C^{(4)}_n &= C_{n+4} - 4C_{n+3} + 6C_{n+2} - 4C_{n+1} + C_n \\ &= 9C_n \frac{9n^4 + 54n^3 + 135n^2 + 122n + 40}{(n+2)(n+3)(n+4)(n+5)}. \end{aligned}$$

References

[1] D. André. Sur les permutations alternées. *J. Math. Pures Appl.*, **7**:167–184, 1881.

[2] E. Babson and J. West. The permutations $123p_1 \dots p_t$ and $321p_1 \dots p_t$ are Wilf equivalent. *Graphs Comb.*, **16**:373–380, 2000.

[3] J. Backelin, J. West, and G. Xin. Wilf equivalence for singleton classes. In *Prceedings of the 13th Conferens on Formal Power Series and Algebraic Combinatorics*, pages 29–38, 2001.

[4] M. Bóna. *Combinatorics of Permutations.* CRC Press, 2004.

[5] H. O. Foulkes. Enumeration of permutations with prescribed up-down and inversion sequences. *Discrete Math.*, **15**:235–252, 1976.

[6] H. O. Foulkes. Tangent and secant numbers and representations of symmetric groups. *Discrete Math.*, **15**:311–324, 1976.

[7] W. Fulton. *Young Tableaux*, volume **35** of *L.M.S. Student Texts*. Cambridge University Press, 1997.

[8] O. Guibert and S. Linusson. Doubly alternating Baxter permutations are Catalan. *Discrete Math.*, **217**:157–166, 2000.

[9] K. E. Iverson. *A Programming Language.* Wiley, 1962.

[10] D. Knuth. *The art of computer programming. Vol 3: Sorting and searching.* Addison-Wesley, 1973.

[11] T. Mansour. Restricted 132-alternating permutations and Chebyshev polynomials. *Annals of Combinatorics.*, **7:2**:201–227, 2003.
[12] R. Simion and F. W. Schmidt. Restricted permutations. *European J. Combin.*, **6**:383–406, 1985.
[13] R. Stanley. private communications, December 2005.
[14] J. West. Permutations with restricted subsequences and stack-sortable permutations. Ph. D. thesis, MIT, 1990.

Linköpings Universitet, 581 83 Linköping, Sweden
E-mail address: `erouc@mai.liu.se`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Enriched P-partitions and peak algebras of types A and B

T. Kyle Petersen

Abstract. We generalize Stembridge's enriched P-partitions and use this theory to outline the structure of peak algebras for the symmetric group and the hyperoctahedral group. Whereas Stembridge's enriched P-partitions are related to quasisymmetric functions (the coalgebra dual to Solomon's type A descent algebra), our generalized enriched P-partitions are related to type B quasisymmetric functions (the coalgebra dual to Solomon's type B descent algebra). Using these functions, we explore three different peak algebras: the "interior" and "left" peak algebras of type A, and a new type B peak algebra. Our results specialize to results for commutative peak algebras as well.

Résumé. Nous généralisons les P-partitions enrichies de Stembridge et employons cette théorie pour décrire la structure des algèbres de pics du groupe symétrique et du groupe hyperoctaédral. Tandis que les P-partitions enrichies sont liés aux fonctions quasisymmetriques (la coalgèbre duale de l'algèbre des descentes du type A de Solomon), nos généralisations des P-partitons enrichies sont liés aux fonctions quasisymmetriques de type B (la coalgèbre duale de l'algèbre des descentes de type B de Solomon). En utilisant ces fonctions, nous présentons trois différentes algèbres de pics : les algébres de pics "intérieures" et "gauches" de type A, et une nouvelle algèbre du pics du type B. Nous en déduisons des résultats reliés à des algèbres de pics commutatives.

1. Introduction

Much attention has been given to the so-called *descent algebras*; see [**3, 4, 5, 10, 11, 12, 14, 15, 18, 19, 21, 25**]. Here we add a chapter to the story of the more recently introduced *peak algebras*. Our approach expands on the one taken in [**13, 16**] and [**21**], where descents were studied using Richard Stanley's P-partitions, or modified versions thereof. This paper is a condensed version of [**22**], which contains several results not mentioned here, as well as any omitted proofs.

Generically, a *peak* of a permutation $\pi \in \mathfrak{S}_n$ is a position i such that $\pi(i-1) < \pi(i) > \pi(i+1)$. The only difference between the various types of peak sets we will study is the values of i that we allow. The *interior peak set* and the *left peak set* are, respectively:

$$\mathrm{Pk}(\pi) := \{i \in [2, n-1] \,|\, \pi(i-1) < \pi(i) > \pi(i+1)\}$$
$$\mathrm{Pk}^{(\ell)}(\pi) := \{i \in [1, n-1] \,|\, \pi(i-1) < \pi(i) > \pi(i+1)\},$$

where we take $\pi(0) = 0$. For example, the permutation $\pi = (2, 1, 4, 3, 5)$ has $\mathrm{Pk}(\pi) = \{3\}$, $\mathrm{Pk}^{(\ell)}(\pi) = \{1, 3\}$. We will also study the peak set of signed permutations $\pi \in \mathfrak{B}_n$, defined by

$$\mathrm{Pk}_B(\pi) := \{i \in [0, n-1] \,|\, \pi(i-1) < \pi(i) > \pi(i+1)\},$$

where $\pi(0) = 0$ and we say there is a peak in position 0 if $\pi(1) < 0$. For example, if $\pi = (-2, 3, 4, -5, 1)$, then $\mathrm{Pk}_B(\pi) = \{0, 3\}$. One can study the suitably defined *right* and *exterior* peaks as well, but the algebraic implications are more limited. See sections 5 and 6, and also [**22**].

The study of algebras related to peaks began with John Stembridge's paper [**26**] on *enriched* P-partitions, followed by others, including [**1, 2, 6, 7, 8, 9, 17, 24**]. While [**26**] explores "the algebra of peaks" related

1991 *Mathematics Subject Classification.* Primary 05E99; Secondary 20C05.

Key words and phrases. peak algebra, enriched P-partition, quasisymmetric function, peak function.

to quasisymmetric functions, it does not use enriched P-partitions for the study of subalgebras of the group algebra $\mathbb{Z}[\mathfrak{S}_n]$ as we will here, and the only notion of peak that it uses is that of an interior peak. Kathryn Nyman [**20**] built on [**26**] to show that there is a subalgebra of the group algebra of the symmetric group, akin to Solomon's descent algebra [**25**], formed by the linear span of

$$v_I := \sum_{\substack{\pi \in \mathfrak{S}_n \\ \mathrm{Pk}(\pi)=I}} \pi,$$

which we call the *interior peak algebra*, denoted $\mathfrak{P}_n$. Later, without the use of enriched P-partitions, Marcelo Aguiar, Nantel Bergeron, and Nyman [**1**] showed that left peaks also give a subalgebra in this sense. We will denote by $\mathfrak{P}_n^{(\ell)}$ the linear span of sums of permutations with the same set of left peaks. In [**1**], the authors also examined commutative subalgebras of the peak algebras—the "Eulerian" peak algebras formed by sums of permutations with the same number of peaks. One goal of this work is to derive some of the results of [**1**] as a natural application of enriched P-partitions. In doing so, we are led to the type B enriched P-partitions and to the type B peak algebra, $\mathfrak{P}_{B,n}$.

The link between peak algebras and enriched P-partitions is through quasisymmetric generating functions. Let $\mathrm{Qsym} := \bigoplus_{n\geq 0} \mathrm{Qsym}_n$ denote the space of quasisymmetric functions, where Qsym_n denotes the quasisymmetric functions homogeneous of degree n. Ira Gessel [**16**] showed how generating functions for ordinary P-partitions give a natural basis for Qsym, and moreover, he defined a coproduct on Qsym_n that makes it the coalgebra dual to Solomon's descent algebra for the Coxeter group of type A_{n-1}. Stembridge [**26**] defined generating functions for enriched P-partitions that form a subring of the ring of quasisymmetric functions, called the *peak functions*. Let $\mathbf{\Pi} := \bigoplus_{n\geq 0} \mathbf{\Pi}_n$ denote the space of peak functions, with $\mathbf{\Pi}_n$ the n-th graded component. We will use an approach similar to Gessel's to give a coproduct on $\mathbf{\Pi}_n$ that makes it dual to Nyman's interior peak algebra.

Just as Stembridge's enriched P-partitions connect with quasisymmetric functions (the coalgebra dual to Solomon's type A descent algebra), the new types of enriched P-partitions we present here connect to the type B quasisymmetric functions, $\mathrm{BQsym} := \bigoplus_{n\geq 0} \mathrm{BQsym}_n$ (the coalgebra dual to Solomon's type B descent algebra), as defined by Chak-On Chow [**13**] using type B P-partitions. We will define the *type B peak functions* $\mathbf{\Pi}_B := \bigoplus_{n\geq 0} \mathbf{\Pi}_{B,n}$ and the *left peak functions* $\mathbf{\Pi}^{(\ell)} := \bigoplus_{n\geq 0} \mathbf{\Pi}_n^{(\ell)}$, and give a natural coproduct that makes $\mathbf{\Pi}_n^{(\ell)}$ dual to $\mathfrak{P}_n^{(\ell)}$ and $\mathbf{\Pi}_{B,n}$ dual to $\mathfrak{P}_{B,n}$.

Remark 1.1. It is known that the quasisymmetric functions form a Hopf algebra, and Stembridge's peak functions are a Hopf subalgebra [**8**]. A natural question is whether the type B quasisymmetric functions form a Hopf algebra, and they do. As of this writing, it is known that the left peak functions do *not* form a Hopf subalgebra, but an unresolved question is whether type B peak functions form a Hopf subalgebra. This topic is part of ongoing work.

2. Enriched P-partitions

The "P" in P-partition stands for a partially ordered set, or poset. For our purposes, we assume that all posets P, with partial order $<_P$, are finite. And unless otherwise noted, if $|P| = n$, then the elements of P are labeled distinctly with the numbers $1, 2, \ldots, n$. We will sometimes describe a poset by its Hasse diagram, as in Figure 1. We can think of any permutation $\pi \in \mathfrak{S}_n$ as a poset with the total order $\pi(s) <_\pi \pi(s+1)$.

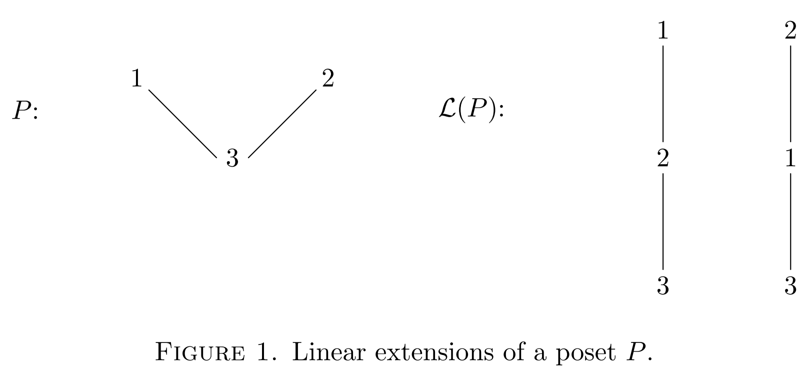


Figure 1. Linear extensions of a poset P.

For a poset P with n elements, let $\mathcal{L}(P)$ denote its Jordan-Hölder set: the set of all permutations of $[n]$ which extend P to a total order. This set is also called the set of "linear extensions" of P. For example let P be the poset defined by $1 >_P 3 <_P 2$. In linearizing P we form a total order by retaining all the relations of P but introducing new relations so that any element is comparable to any other. In this case, 1 and 2 are not comparable, so we have exactly two ways of linearizing P: $3 < 2 < 1$ or $3 < 1 < 2$. These correspond to the permutations $(3, 2, 1)$ and $(3, 1, 2)$. Let us make the following observation.

Observation 2.1. *A permutation π is in $\mathcal{L}(P)$ if and only if $i <_P j$ implies $\pi^{-1}(i) < \pi^{-1}(j)$.*

In other words, if i is "below" j in the Hasse diagram of the poset P, it must be below j in any linear extension of the poset.

We now introduce the basic theory of enriched P-partitions, building on Stembridge's work [**26**]. To begin, Stembridge defines $\mathbb{P}'$ to be the set of nonzero integers with the following total order:

$$-1 < 1 < -2 < 2 < -3 < 3 < \cdots$$

We will have use for this set, but we view it as a subset of a similar set. Define $\mathbb{P}^{(\ell)}$ to be the integers with the following total order:

$$0 < -1 < 1 < -2 < 2 < -3 < 3 < \cdots$$

Then $\mathbb{P}'$ is simply the set of all $i \in \mathbb{P}^{(\ell)}$, $i > 0$. In general, for any countable totally ordered set $S = \{s_1, s_2, \ldots\}$ we define $S^{(\ell)}$ to be the set

$$\{s_0, -s_1, s_1, -s_2, s_2, \ldots\},$$

with total order

$$s_0 < -s_1 < s_1 < -s_2 < s_2 < \cdots$$

(so we can think of $S^{(\ell)}$ as two interwoven copies of S along with a zero element) and define S' to be the set $\{s \in S^{(\ell)} \,|\, s > s_0\}$. For any $s_i \in \{s_0\} \cup S$, we say s_i is *nonnegative*. On the other hand, if $i \neq 0$ we say $-s_i$ is *negative*. The absolute value removes any minus signs: $|\pm s| = s$ for any $s \in \{s_0\} \cup S$.

For s and t in $S^{(\ell)}$, we write $s \leq^+ t$ to mean either $s < t$ in $S^{(\ell)}$, or $s = t \geq 0$. Similarly we define $s \leq^- t$ to mean either $s < t$ in $S^{(\ell)}$, or $s = t < 0$. For example, on $\mathbb{P}^{(\ell)}$, we have $\{s \,|\, s \leq^+ 3\} = \{0, \pm 1, \pm 2, \pm 3\}$, $\{s \,|\, s \leq^- 3\} = \{0, \pm 1, \pm 2, -3\} = \{s \,|\, s \leq^- -3\}$, $\{s \,|\, 0 \leq^+ s \leq^+ 2\} = \{0, \pm 1, \pm 2\}$ and $\{s \,|\, 0 \leq^- s \leq^+ 2\} = \{\pm 1, \pm 2\}$.

Definition 2.2 (Enriched P-partition). An *enriched P-partition* (resp. *left enriched P-partition*) is an order-preserving map $f : P \to S'$ (resp. $S^{(\ell)}$) such that for all $i <_P j$ in P,

(1) $f(i) \leq^+ f(j)$ only if $i < j$ in $\mathbb{N}$,
(2) $f(i) \leq^- f(j)$ only if $i > j$ in $\mathbb{N}$.

It is helpful to remember that Stembridge's enriched P-partitions are the nonzero left enriched P-partitions, i.e., those for which $f(i) \neq s_0$ for any i. We let $\mathcal{E}(P; S)$ denote the set of all enriched P-partitions $f : P \to S'$; $\mathcal{E}^{(\ell)}(P; S)$ denotes the set of left enriched P-partitions $f : P \to S^{(\ell)}$. If S is irrelevant or understood, we simply write $\mathcal{E}(P)$ or $\mathcal{E}^{(\ell)}(P)$. For example, if our poset is $1 >_P 3 <_P 2$, then

$$\mathcal{E}^{(\ell)}(P) = \{f : P \to S^{(\ell)} \,|\, f(1) \geq^- f(3) \leq^- f(2)\},$$

which we can see actually splits into the two following disjoint subsets:

$$\{f(3) \leq^- f(1) \leq^+ f(2)\} \sqcup \{f(3) \leq^- f(2) \leq^- f(1)\} = \mathcal{E}^{(\ell)}(312) \sqcup \mathcal{E}^{(\ell)}(321).$$

This example leads us to the following, which, by analogy with a similar result for ordinary P-partitions, is referred to as the fundamental lemma of enriched P-partitions. It follows by induction on the number of incomparable pairs of elements in the poset.

Lemma 2.3. *For any poset P, the set of all (left) enriched P-partitions is the disjoint union of all (left) enriched π-partitions for linear extensions π of P. Equivalently,*

$$\mathcal{E}(P) = \coprod_{\pi \in \mathcal{L}(P)} \mathcal{E}(\pi),$$

$$\mathcal{E}^{(\ell)}(P) = \coprod_{\pi \in \mathcal{L}(P)} \mathcal{E}^{(\ell)}(\pi).$$

Therefore when studying enriched P-partitions it is enough to consider the case where P is a totally ordered chain, i.e., a permutation π. It is easy to describe the set of all enriched π-partitions in terms of descent sets. For any $\pi \in \mathfrak{S}_n$ we have

$$\begin{aligned} \mathcal{E}(\pi) = \{\, f : [n] \to S' \mid & \; f(\pi(1)) \le f(\pi(2)) \le \cdots \le f(\pi(n)), \\ & i \notin \mathrm{Des}(\pi) \Rightarrow f(\pi(i)) \le^{+} f(\pi(i+1)) \\ & i \in \mathrm{Des}(\pi) \Rightarrow f(\pi(i)) \le^{-} f(\pi(i+1)) \,\}, \end{aligned} \tag{2.1}$$

and the analogous description for $\mathcal{E}^{(\ell)}(\pi)$ where we replace S' with $S^{(\ell)}$.

From (2.1) it is clear that enriched π-partitions depend on the descent set of π. The connection to peaks is less obvious. In section 3 we will establish this link, and also show how left enriched π-partitions are related to left peaks. First, we present our main theorem.

Let S and T be any two countable totally ordered sets, and let $S' \times T' = \{(s,t) \,|\, s \in S', t \in T'\}$ be the cartesian product of S' and T' with the *up-down order* defined as follows: $(s,t) < (u,v)$ if and only if

(1) $s < u$, or
(2) $s = u > 0$ and $t < v$, or
(3) $s = u < 0$ and $t > v$.

In other words, we read up the nonnegative columns, down the negative ones. Here we write $(s,t) \le^{+} (u,v)$ in one of three cases: if $s < u$, or if $s = u > 0$ and $t \le^{+} v$, or if $s = u < 0$ and $t \ge^{-} v$. Similarly, $(s,t) \le^{-} (u,v)$ if $s < u$, or if $s = u > 0$ and $t \le^{-} v$, or if $s = u < 0$ and $t \ge^{+} v$. We define $S^{(\ell)} \times T^{(\ell)}$ in the same way. See Figure 2.

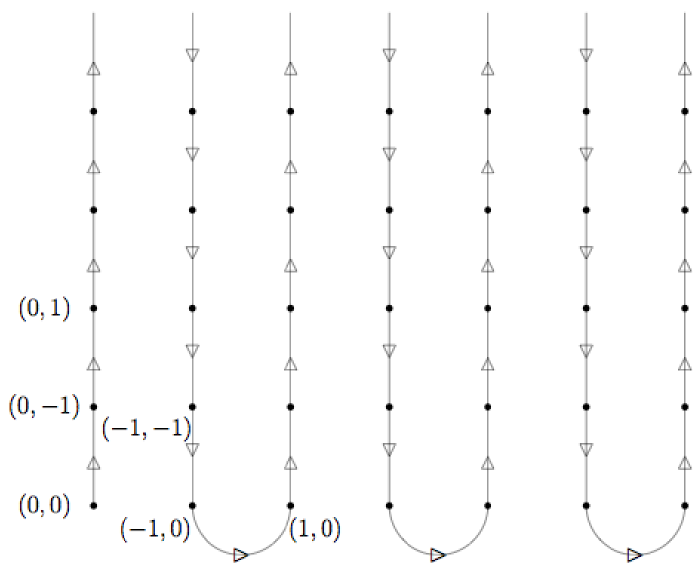


Figure 2. The up-down order for $\mathbb{P}^{(\ell)} \times \mathbb{P}^{(\ell)}$.

Theorem 2.4. *We have the following bijections:*

$$\mathcal{E}(\pi; S \times T) \longleftrightarrow \coprod_{\sigma\tau=\pi} \mathcal{E}(\tau; S) \times \mathcal{E}(\sigma; T) \tag{2.2}$$

$$\mathcal{E}^{(\ell)}(\pi; S \times T) \longleftrightarrow \coprod_{\sigma\tau=\pi} \mathcal{E}^{(\ell)}(\tau; S) \times \mathcal{E}(\sigma; T) \tag{2.3}$$

Proof. We will provide proof for (2.2) and remark that the proof of (2.3) is nearly identical.

For $\pi \in \mathfrak{S}_n$, we can write the set of all enriched π-partitions $f : \pi \to S' \times T'$ as follows:

$$\begin{aligned} \mathcal{E}(\pi) = \{\, F = ((s_1, t_1), \ldots, (s_n, t_n)) \in (S' \times T')^n \mid (s_1, t_1) \le (s_2, t_2) \le \cdots \le (s_n, t_n), \\ i \notin \mathrm{Des}(\pi) \Rightarrow (s_i, t_i) \le^+ (s_{i+1}, t_{i+1}) \\ i \in \mathrm{Des}(\pi) \Rightarrow (s_i, t_i) \le^- (s_{i+1}, t_{i+1}) \,\}. \end{aligned} \tag{2.4}$$

We will now sort the points F into distinct cases. For any $i = 1, 2, \ldots, n-1$, if $\pi(i) < \pi(i+1)$, then $(s_i, t_i) \le^+ (s_{i+1}, t_{i+1})$, which falls into one of two mutually exclusive cases:

$$s_i \le^+ s_{i+1} \text{ and } t_i \le^+ t_{i+1}, \text{ or} \tag{2.5}$$

$$s_i \le^- s_{i+1} \text{ and } t_i \ge^- t_{i+1}. \tag{2.6}$$

If $\pi(i) > \pi(i+1)$, then $(s_i, t_i) \le^- (s_{i+1}, t_{i+1})$, which we split as:

$$s_i \le^+ s_{i+1} \text{ and } t_i \le^- t_{i+1}, \text{ or} \tag{2.7}$$

$$s_i \le^- s_{i+1} \text{ and } t_i \ge^+ t_{i+1}, \tag{2.8}$$

also mutually exclusive. Define I_F to be the set of all i such that either (2.6) or (2.8) holds for F. Notice that in both cases, $s_i \le^- s_{i+1}$. Now for any $I \subset [n-1]$, let A_I be the set of all F satisfying $I_F = I$. We have $\mathcal{E}(\pi; S \times T) = \coprod_{I \subset [n-1]} A_I$.

For any particular $I \subset [n-1]$, form the poset P_I of the elements $1, 2, \ldots, n$ by $\pi(s) <_{P_I} \pi(s+1)$ if $s \notin I$, $\pi(s) >_{P_I} \pi(s+1)$ if $s \in I$. We form a "zig-zag" poset (see Figure 3) of n elements labeled consecutively by $\pi(1), \pi(2), \ldots, \pi(n)$ with downward zigs corresponding to the elements of I.

P_I:

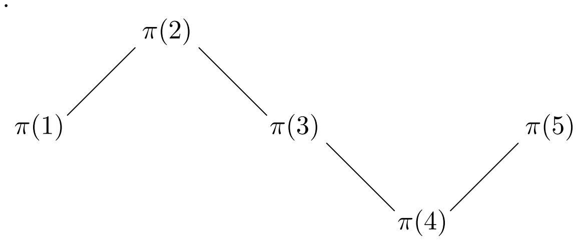


FIGURE 3. The zig-zag poset P_I for $I = \{2, 3\} \subset [5]$.

For any F in A_I, let $f : [n] \to T'$ be defined by $f(\pi(i)) = t_i$. It is straightforward to verify that f is an enriched P_I-partition. Conversely, any enriched P_I-partition f gives a point F in A_I since by cases (2.5)–(2.8) above, if $t_i = f(\pi(i))$, then

$$((s_1, t_1), \ldots, (s_n, t_n)) \in A_I$$

if and only if $s_1 \le \cdots \le s_n$ and $s_i \le^- s_{i+1}$ for all $i \in I$, $s_i \le^+ s_{i+1}$ for $i \notin I$. We can therefore turn our attention to enriched P_I-partitions.

Let $\sigma \in \mathcal{L}(P_I)$. Recall by Observation 2.1 that $\sigma^{-1}\pi(i) < \sigma^{-1}\pi(i+1)$ if $\pi(i) <_{P_I} \pi(i+1)$, i.e., if $i \notin I$. If $\pi(i) >_{P_I} \pi(i+1)$ then $\sigma^{-1}\pi(i) > \sigma^{-1}\pi(i+1)$ and $i \in I$. We get that $\mathrm{Des}(\sigma^{-1}\pi) = I$ if and only if $\sigma \in \mathcal{L}(P_I)$. Set $\tau = \sigma^{-1}\pi$. We have

$$\mathcal{E}(\tau; S) = \{s_1 \le \cdots \le s_n \,|\, s_i \le^- s_{i+1} \text{ if } i \in \mathrm{Des}(\tau), s_i \le^+ s_{i+1} \text{ otherwise}\},$$

and since $\mathrm{Des}(\tau) = I$, we can write A_I as

$$\coprod_{\substack{\sigma \in \mathcal{L}(P_I) \\ \sigma\tau = \pi}} \{F \in (S' \times T')^n \,|\, (s_1, \ldots, s_n) \in \mathcal{E}(\tau; S), (t_{\pi^{-1}\sigma(1)}, \ldots, t_{\pi^{-1}\sigma(n)}) \in \mathcal{E}(\sigma; T)\}.$$

Running over all subsets $I \subset [n-1]$, we obtain

$$\mathcal{E}(\pi; S \times T) = \coprod_{\sigma\tau = \pi} \{F \in (S' \times T')^n \,|\, (s_1, \ldots, s_n) \in \mathcal{E}(\tau; S), (t_{\pi^{-1}\sigma(1)}, \ldots, t_{\pi^{-1}\sigma(n)}) \in \mathcal{E}(\sigma; T)\}.$$

(Note that $\pi^{-1}\sigma = \tau^{-1}$.) Now we can see the obvious bijection $\coprod_{\sigma\tau=\pi} \mathcal{E}(\tau; S) \times \mathcal{E}(\sigma; T) \to \mathcal{E}(\pi; S \times T)$ given by

$$((s_1, \ldots, s_n), (t_1, \ldots, t_n)) \mapsto ((s_1, t_{\tau(1)}), \ldots, (s_n, t_{\tau(n)})).$$

$\square$

Now we present the type B enriched P-partitions. When working with signed permutations, we need to change our notion of a poset slightly. See Chow [**13**]; this definition is a simpler version of the notion due to Vic Reiner [**23**].

DEFINITION 2.5. A *type B poset*, or $\mathfrak{B}_n$ *poset*, is a poset P whose elements are $0, \pm 1, \pm 2, \ldots, \pm n$ such that if $i <_P j$ then $-j <_P -i$.

Note that if we are given a poset with $n+1$ elements labeled by $0, a_1, \ldots, a_n$ where $a_i = i$ or $-i$, then we can extend it to a $\mathfrak{B}_n$ poset of $2n+1$ elements. For example, the P in Figure 4 could be specified by the relations $0 >_P 1 <_P -2$. In the same way, any signed permutation $\pi \in \mathfrak{B}_n$ is a $\mathfrak{B}_n$ poset under the total order $\pi(s) <_\pi \pi(s+1)$, $0 \le s \le n-1$. If P is a type B poset, let $\mathcal{L}_B(P)$ denote the set of linear extensions of P that are themselves type B posets. Then $\mathcal{L}_B(P)$ is naturally identified with some set of signed permutations. See Figure 4.

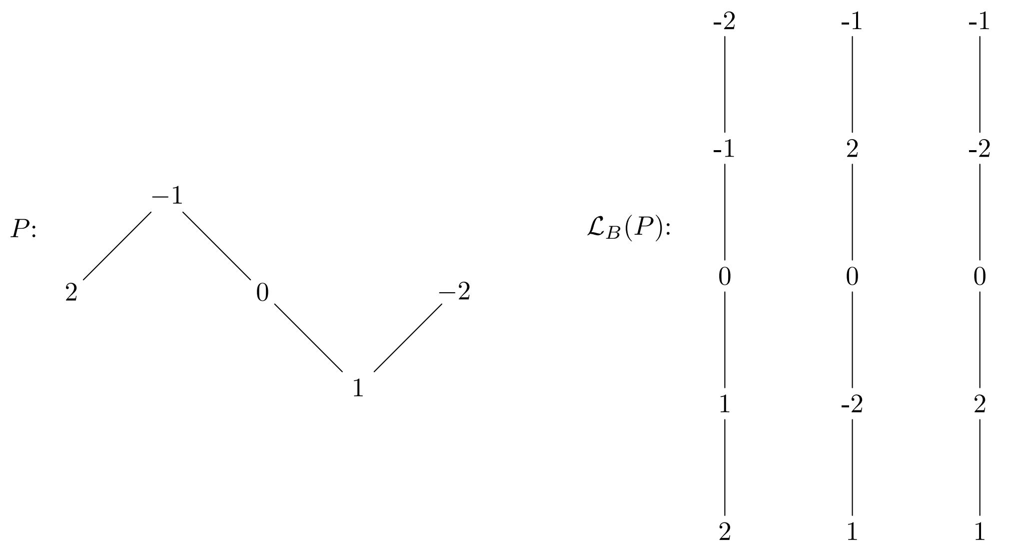


FIGURE 4. A $\mathfrak{B}_2$ poset and its linear extensions.

We will present some alternate notation for the set S' introduced above. Let $S = \{s_1, s_2, \ldots\}$ be any countable totally ordered set. Then we define the set S' to be the set

$$\{s_1^{-1}, s_1, s_2^{-1}, s_2, \ldots\},$$

with total order

$$s_1^{-1} < s_1 < s_2^{-1} < s_2 < \cdots$$

We introduce this new notation because we want to avoid confusion in defining the set

$$\mathbb{Z}' = \{\ldots, -2, -2^{-1}, -1, -1^{-1}, 0, 1^{-1}, 1, 2^{-1}, 2, \ldots\},$$

with the total order

$$\cdots - 2 < -2^{-1} < -1 < -1^{-1} < 0 < 1^{-1} < 1 < 2^{-1} < 2 < \cdots$$

In general, if we define $\pm S = \{\ldots, -s_2, -s_1, s_0, s_1, s_2, \ldots\}$, we have the total order on $\pm S'$ given by

$$\cdots - s_2 < -s_2^{-1} < -s_1 < -s_1^{-1} < s_0 < s_1^{-1} < s_1 < s_2^{-1} < s_2 < \cdots$$

For any s in $\pm S'$, let $\varepsilon(s)$ be the exponent on s, and let $|s|$ be a map $\pm S' \to S$ that forgets signs and exponents. For example, if $s = -s_i^{-1}$, then $\varepsilon(s) = -1 < 0$ and $|s| = s_i$, while if $s = s_i$, then $\varepsilon(s) = 1 > 0$

and $|s| = s_i$. For $i = 0$, we require $\varepsilon(s_0) = 1 > 0$, $|s_0| = s_0$, and $-s_0 = s_0$. Let $s \leq^+ t$ mean that $s < t$ in $\pm S'$ or $s = t$ and $\varepsilon(s) > 0$. Similarly define $s \leq^- t$ to mean that $s < t$ in $\pm S'$ or $s = t$ and $\varepsilon(s) < 0$.

DEFINITION 2.6 (Type B enriched P-partition). For any $\mathfrak{B}_n$ poset P, an *enriched P-partition of type B* is an order-preserving map $f : \pm[n] \to \pm S'$ such that for every $i <_P j$ in P,

(1) $f(i) \leq^+ f(j)$ only if $i < j$ in $\mathbb{Z}$,
(2) $f(i) \leq^- f(j)$ only if $i > j$ in $\mathbb{Z}$,
(3) $f(-i) = -f(i)$.

This definition differs from type A enriched P-partitions only in the last condition. It forces $f(0) = s_0$, and if we know where to map $a_1, a_2, \ldots, a_n$, where $a_i = i$ or $-i$, then it tells us where to map everything else. Let $\mathcal{E}_B(P; S)$ denote the set of all type B enriched P-partitions $f : P \to \pm S'$. There is a fundamental lemma for type B.

LEMMA 2.7. *We have,*

$$\mathcal{E}_B(P) = \coprod_{\pi \in \mathcal{L}_B(P)} \mathcal{E}_B(\pi).$$

We can easily characterize the type B enriched π-partitions in terms of descent sets, keeping in mind that if we know where to map i, then we know where to map $-i$ by the symmetry property: $f(-i) = -f(i)$. For any signed permutation $\pi \in \mathfrak{B}_n$ we have

$$\begin{aligned} \mathcal{E}_B(\pi) = \{\, f : [n] \to \pm S' \mid {} & s_0 \leq f(\pi(1)) \leq f(\pi(2)) \leq \cdots \leq f(\pi(n)) \\ & i \notin \mathrm{Des}_B(\pi) \Rightarrow f(\pi(i)) \leq^+ f(\pi(i+1)), \\ & i \in \mathrm{Des}_B(\pi) \Rightarrow f(\pi(i)) \leq^- f(\pi(i+1)) \,\}. \end{aligned} \tag{2.9}$$

Notice that since $\varepsilon(s_0) = 1$, then $s_0 \leq^- f(\pi(1))$ is the same as saying $s_0 < f(\pi(1))$, and $s_0 \leq^+ f(\pi(1))$ is the same as $s_0 \leq f(\pi(1))$. We will show that the set of type B enriched P-partitions relates to the set of type B peaks. First, we present the main theorem for type B enriched P-partitions. Its proof varies only slightly from that of Theorem 2.4 and is omitted. Let $\mathcal{E}_B(P; S \times T)$ denote the set of all enriched P-partitions $f : P \to \pm S' \times \pm T'$ with the up-down order.

THEOREM 2.8. *We have the following bijection:*

$$\mathcal{E}_B(\pi; S \times T) \longleftrightarrow \coprod_{\sigma\tau = \pi} \mathcal{E}_B(\tau; S) \times \mathcal{E}_B(\sigma; T) \tag{2.10}$$

3. Generating functions

Recall that a quasisymmetric function is a formal series

$$Q(x_1, x_2, \ldots) \in \mathbb{Z}[[x_1, x_2, \ldots]]$$

of bounded degree such that for any composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_k^{\alpha_k}$ is the same as the coefficient of $x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_k}^{\alpha_k}$ for all $i_1 < i_2 < \cdots < i_k$. Recall that a composition of n, written $\alpha \models n$, is an ordered tuple of positive integers $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ such that $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_k = n$. In this case we say that α has k parts, or $l(\alpha) = k$. We can put a partial order on the set of all compositions of n by refinement. The covering relations are of the form

$$(\alpha_1, \ldots, \alpha_i + \alpha_{i+1}, \ldots, \alpha_k) \prec (\alpha_1, \ldots, \alpha_i, \alpha_{i+1}, \ldots, \alpha_k).$$

Let Qsym_n denote the set of all quasisymmetric functions homogeneous of degree n. Then $\mathrm{Qsym} := \bigoplus_{n \geq 0} \mathrm{Qsym}_n$ denotes the graded ring of all quasisymmetric functions, where $\mathrm{Qsym}_0 = \mathbb{Z}$.

The most obvious basis for Qsym_n is the set of *monomial* quasisymmetric functions, defined for any composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k) \models n$,

$$M_\alpha := \sum_{i_1 < i_2 < \cdots < i_k} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_k}^{\alpha_k}.$$

There are 2^{n-1} compositions of n, and hence, the graded component Qsym_n has dimension 2^{n-1} as a vector space. We can form another natural basis with the *fundamental* quasisymmetric functions, also indexed by compositions,

$$F_\alpha := \sum_{\alpha \leq \beta} M_\beta,$$

since, by inclusion-exclusion we can express the M_α in terms of the F_α:

$$M_\alpha = \sum_{\alpha \leq \beta} (-1)^{l(\beta)-l(\alpha)} F_\beta.$$

There is a well-known bijection between compositions of n and subsets of $[n-1]$ given by

$$\alpha \mapsto I(\alpha) = \{\alpha_1, \alpha_1+\alpha_2, \ldots, \alpha_1+\cdots+\alpha_{k-1}\},$$

and so we can also write $M_\alpha = M_{I(\alpha)}$ or $F_\alpha = F_{I(\alpha)}$ when convenient.

Define the generating function for enriched P-partitions $f : P \to \mathbb{P}'$ by

$$\Lambda(P) = \sum_{f \in \mathcal{E}(P)} \prod_{i=1}^{n} z_{|f(i)|}.$$

Then clearly $\Lambda(P)$ is a quasisymmetric function. By the fundamental Lemma 2.3, we have that

$$\Lambda(P) = \sum_{\pi \in \mathcal{L}(P)} \Lambda(\pi).$$

For any subset of the integers I, define the set $I+1 = \{i+1 \mid i \in I\}$. From [**26**] we see that the generating function for enriched π-partitions depends only on the peak set of π.

THEOREM 3.1 (Stembridge [**26**], Proposition 2.2). *For $\pi \in \mathfrak{S}_n$, we have the following equality:*

$$\Lambda(\pi) = \sum_{\substack{E \subset [n-1] \\ \mathrm{Pk}(\pi) \subset E \cup (E+1)}} 2^{|E|+1} M_E.$$

For any sets I and J, let $I \bigtriangleup J = (I \cup J) \setminus (I \cap J)$ denote the symmetric difference of sets. The generating functions are also F-positive.

THEOREM 3.2 (Stembridge [**26**], Proposition 3.5). *For $\pi \in \mathfrak{S}_n$, we have the following equality:*

$$\Lambda(\pi) = 2^{|\mathrm{Pk}(\pi)|+1} \sum_{\substack{D \subset [n-1] \\ \mathrm{Pk}(\pi) \subset D \bigtriangleup (D+1)}} F_D. \tag{3.1}$$

For interior peak sets I, let K_I be the quasisymmetric function defined by

$$K_I := \Lambda(\pi),$$

where π is any permutation such that $\mathrm{Pk}(\pi) = I$. Let $\mathbf{\Pi}_n$ denote the space of quasisymmetric functions spanned by the K_I, where I runs over all interior peak sets of $[n-1]$. Stembridge then defines the set of *peak functions* $\mathbf{\Pi} := \bigoplus_{n \geq 0} \mathbf{\Pi}_n$, which is a graded subring of Qsym. He proved that the functions K_I are linearly independent, and so the rank of $\mathbf{\Pi}_n$ is the the number of distinct interior peak sets, which happens to be the Fibonacci number f_{n-1}, defined by $f_0 = f_1 = 1$ and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$.

Before discussing generating functions for left enriched P-partitions and type B enriched P-partitions, we need to introduce Chow's type B quasisymmetric functions [**13**]. Define a *pseudo-composition* of n, written $\alpha \Vdash n$, to be an ordered tuple of nonnegative integers $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ whose sum $|\alpha| = \alpha_1 + \cdots + \alpha_k$ is n, where $\alpha_1 \geq 0$, $\alpha_i > 0$ for $i > 1$. In other words, given any ordinary composition $\alpha \models n$, we have two corresponding pseudo-compositions: α and $0\alpha = (0, \alpha_1, \ldots, \alpha_k)$. The partial order on the set of all pseudo-compositions of n is again by refinement.

Now we can define a type B quasisymmetric function to be a formal series

$$Q(x_0, x_1, x_2, \ldots) \in \mathbb{Z}[[x_0, x_1, x_2, \ldots]]$$

of bounded degree such that for any pseudo-composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, the coefficient of $x_0^{\alpha_1} x_1^{\alpha_2} \cdots x_{k-1}^{\alpha_k}$ is the same as the coefficient of $x_0^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_k}^{\alpha_k}$ for all $0 < i_2 < \cdots < i_k$. Let BQsym_n denote the set of all quasisymmetric functions homogeneous of degree n. Then $\mathrm{BQsym} := \bigoplus_{n \geq 0} \mathrm{BQsym}_n$ is the ring of type

B quasisymmetric functions. As before we have a monomial and fundamental basis for BQsym_n. For any pseudo-composition $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k) \Vdash n$, the monomial functions are

$$M_{B,\alpha} := \sum_{i_2 < \cdots < i_k} x_0^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_k}^{\alpha_k}.$$

There are 2^n pseudo-compositions of n, so the dimension of BQsym_n is 2^n. The fundamental basis is

$$F_{B,\alpha} := \sum_{\alpha \le \beta} M_{B,\beta}.$$

There is a bijection between pseudo-compositions of n and subsets of $[0, n-1]$ given by the same map

$$\alpha \mapsto I(\alpha) = \{\alpha_1, \alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{k-1}\},$$

and so we can also write $M_{B,\alpha} = M_{B,I(\alpha)}$ or $F_{B,\alpha} = F_{B,I(\alpha)}$ when convenient.

Define the generating functions for left enriched P-partitions $f : P \to \mathbb{P}^{(\ell)}$, and type B enriched P-partitions $f : P \to \mathbb{Z}'$,

$$\Lambda^{(\ell)}(P) = \sum_{f \in \mathcal{E}^{(\ell)}(P)} \prod_{i=1}^{n} z_{|f(i)|},$$

$$\Lambda_B(P) = \sum_{f \in \mathcal{E}_B(P)} \prod_{i=1}^{n} z_{|f(i)|}.$$

The fundamental lemma gives that

$$\Lambda^{(\ell)}(P) = \sum_{\pi \in \mathcal{L}(P)} \Lambda^{(\ell)}(\pi),$$

$$\Lambda_B(P) = \sum_{\pi \in \mathcal{L}_B(P)} \Lambda_B(\pi).$$

We can relate $\Lambda^{(\ell)}(\pi)$ and $\Lambda_B(\pi)$ to the monomial and fundamental quasisymmetric functions of type B. Notice that for a permutation $\pi \in \mathfrak{S}_n \subset \mathfrak{B}_n$, left peaks coincide with the type B peaks. Therefore we can view left enriched P-partitions as a special case of type B enriched P-partitions. Furthermore, since Stembridge's enriched P-partitions are simply those left enriched P-partitions that are nonzero, we have

$$\Lambda(P)(z_1, z_2, \ldots) = \Lambda^{(\ell)}(P)(0, z_1, z_2, \ldots),$$

so the results for $\Lambda(P)$ can be obtained from our results for $\Lambda^{(\ell)}(P)$ by setting $z_0 = 0$.

Theorem 3.3. *For $\pi \in \mathfrak{B}_n$, we have the following equations:*

$$\begin{aligned} \Lambda_B(\pi) &= \sum_{\substack{E \subset [0,n-1] \\ \mathrm{Pk}_B(\pi) \subset E \cup (E+1)}} 2^{|E|} M_{B,E}, \\ &= 2^{|\mathrm{Pk}_B(\pi)|} \sum_{\substack{D \subset [0,n-1] \\ \mathrm{Pk}_B(\pi) \subset D \triangle (D+1)}} F_{B,D}. \end{aligned}$$

We omit the proof of this theorem, but remark that it follows the same lines of reasoning as in Stembridge's proofs of Theorems 3.1 and 3.2.

Corollary 3.4. *For $\pi \in \mathfrak{S}_n$, we have the following equations:*

$$\begin{aligned} \Lambda^{(\ell)}(\pi) &= \sum_{\substack{E \subset [0,n-1] \\ \mathrm{Pk}^{(\ell)}(\pi) \subset E \cup (E+1)}} 2^{|E|} M_{B,E}, \\ &= 2^{|\mathrm{Pk}^{(\ell)}(\pi)|} \sum_{\substack{D \subset [0,n-1] \\ \mathrm{Pk}^{(\ell)}(\pi) \subset D \triangle (D+1)}} F_{B,D}. \end{aligned}$$

Corollary 3.5. *The function $\Lambda_B(\pi)$ depends only on the type B peak set of $\pi \in \mathfrak{B}_n$, the function $\Lambda^{(\ell)}(\pi)$ depends only on the left peak set of $\pi \in \mathfrak{S}_n$.*

We define the functions $K_{B,I}$ by

$$K_{B,I} := \Lambda_B(\pi),$$

where π is any permutation such that $\mathrm{Pk}_B(\pi) = I$. Note that for a permutation $\pi \in \mathfrak{S}_n$, if $\Lambda^{(\ell)}(\pi) = K_{B,I}$ then $0 \notin I$.

Let $\mathbf{\Pi}_{B,n}$ denote the span of the $K_{B,I}$, where I ranges over all type B peak sets of $[0, n-1]$. It is not hard to see that the $K_{B,I}$ are linearly independent, and so by counting the number of type B peak sets we see $\mathbf{\Pi}_{B,n}$ has rank f_{n+1}. If we define the *type B peak functions*, $\mathbf{\Pi}_B := \bigoplus_{n\geq 0} \mathbf{\Pi}_{B,n}$, then we can see it is a subring of BQsym, as an argument identical to that of [**26**] Theorem 3.1 shows.

Similarly, let $\mathbf{\Pi}_n^{(\ell)}$ denote the span of all $K_{B,I}$, where I ranges over the left peak sets in $[1, n-1]$. Then $\mathbf{\Pi}_n^{(\ell)}$ has rank f_n and the *left peak functions*, $\mathbf{\Pi}^{(\ell)} := \bigoplus_{n\geq 0} \mathbf{\Pi}_n^{(\ell)}$, form a subring of $\mathbf{\Pi}_B$.

4. Duality

Let $X = \{x_1, x_2, \ldots\}$ and $Y = \{y_1, y_2, \ldots\}$ be two sets of commuting indeterminates. Define the set $XY = \{xy : x \in X, y \in Y\}$. Then we define the bipartite generating function,

$$\Lambda(P)(XY) = \sum_{F\in\mathcal{E}(P;\mathbb{P}\times\mathbb{P})} x_{s_1}\cdots x_{s_n} y_{t_1}\cdots y_{t_n}.$$

The functions $\Lambda^{(\ell)}(P)(XY)$ and $\Lambda_B(P)(XY)$ are defined similarly. Then the following are consequences of Theorem 2.4 and Theorem 2.8.

THEOREM 4.1. *For any $\pi \in \mathfrak{S}_n$, we have the following equations:*

$$\Lambda(\pi)(XY) = \sum_{\sigma\tau=\pi} \Lambda(\tau)(X)\,\Lambda(\sigma)(Y), \tag{4.1}$$

$$\Lambda^{(\ell)}(\pi)(XY) = \sum_{\sigma\tau=\pi} \Lambda^{(\ell)}(\tau)(X)\,\Lambda^{(\ell)}(\sigma)(Y). \tag{4.2}$$

THEOREM 4.2. *For any $\pi \in \mathfrak{B}_n$, we have the following equation:*

$$\Lambda_B(\pi)(XY) = \sum_{\sigma\tau=\pi} \Lambda_B(\tau)(X)\,\Lambda_B(\sigma)(Y). \tag{4.3}$$

The formulas above imply duality between $\mathbf{\Pi}_n$ and $\mathfrak{P}_n$, $\mathbf{\Pi}_n^{(\ell)}$ and $\mathfrak{P}_n^{(\ell)}$, and $\mathbf{\Pi}_{B,n}$ and $\mathfrak{P}_{B,n}$. Moreover, they give an explicit combinatorial description for the structure constants of the algebras. We will show how this works for the case of interior peaks. The steps of the construction are the same for the other cases.

First, notice that equation (4.1) implies that

$$K_C(XY) = \sum_{A,B} c_{A,B}^C K_A(X) K_B(Y),$$

where the sum is over all pairs of interior peak subsets A and B of $[2, n-1]$, and if $\pi \in \mathfrak{S}_n$ is any permutation with $\mathrm{Pk}(\pi) = C$, then the integer $c_{A,B}^C$ is the number of pairs of permutations σ, τ such that $\mathrm{Pk}(\sigma) = B$, $\mathrm{Pk}(\tau) = A$, and $\sigma\tau = \pi$. We now use this formula to define $\mathbf{\Pi}_n$ as a coalgebra with coproduct $\Delta : \mathbf{\Pi}_n \to \mathbf{\Pi}_n \otimes \mathbf{\Pi}_n$ defined as

$$\Delta(K_C) = \sum_{A,B} c_{A,B}^C K_A \otimes K_B.$$

We can define a coalgebra $\mathbb{Z}[\mathfrak{S}_n]^*$ dual to the group algebra with coproduct defined as

$$\Delta(\pi) = \sum_{\sigma\tau=\pi} \tau \otimes \sigma.$$

Define the map $\varphi^* : \mathbb{Z}[\mathfrak{S}_n]^* \to \mathbf{\Pi}_n$ by $\varphi^*(\pi) = K_{\mathrm{Pk}(\pi)}$, which, by (4.1), is a surjective homomorphism of coalgebras. Now we dualize.

Let $\mathbf{\Pi}_n^*$ be the algebra dual to $\mathbf{\Pi}_n$, with basis elements K_I^*. By definition, multiplication in this basis is

$$K_A^* K_B^* = \sum_C c_{A,B}^C K_C^*,$$

where the sum is over all interior peak subsets C. The dual of φ^* is now an injective homomorphism of algebras, $\varphi : \mathbf{\Pi}_n^* \to \mathbb{Z}[\mathfrak{S}_n]$ defined by

$$\varphi(K_I^*) = \sum_{\mathrm{Pk}(\pi)=I} \pi = v_I.$$

Thus the interior peak algebra can be defined as the image of φ, and the structure constants carry through:

$$v_A v_B = \sum_C c_{A,B}^C v_C.$$

We describe the structure constants for the left and type B peak algebras:

- Let $d_{A,B}^C$, over triples of left peak sets A, B, C, be the structure constants for $\mathfrak{P}_n^{(\ell)}$. Then for any $\pi \in \mathfrak{S}_n$ such that $\mathrm{Pk}^{(\ell)}(\pi) = C$, the integer $d_{A,B}^C$ is the number of pairs of permutations σ, τ such that $\mathrm{Pk}^{(\ell)}(\sigma) = B$, $\mathrm{Pk}^{(\ell)}(\tau) = A$, and $\sigma\tau = \pi$.
- Let $e_{A,B}^C$, over triples of type B peak sets A, B, C, be the structure constants for $\mathfrak{P}_{B,n}$. Then for any $\pi \in \mathfrak{B}_n$ such that $\mathrm{Pk}_B(\pi) = C$, the integer $e_{A,B}^C$ is the number of pairs of permutations σ, τ such that $\mathrm{Pk}_B(\sigma) = B$, $\mathrm{Pk}_B(\tau) = A$, and $\sigma\tau = \pi$.

The type B peak algebra is something new, defined as the linear span of sums of signed permutations in $\mathfrak{B}_n$ with common type B peak set.

THEOREM 4.3. *The space $\mathfrak{P}_{B,n}$ is a subalgebra of $\mathbb{Z}[\mathfrak{B}_n]$ of dimension f_{n+1}. (In fact it is a subalgebra of Solomon's descent algebra for type B_n.)*

REMARK 4.4. While $\mathfrak{P}_n^{(\ell)}$ was introduced in [**1**], the authors had no combinatorial description for its structure constants (see [**1**], Remark 4.4), and neither were the structure constants for $\mathfrak{P}_n$ known. Independently, Nantel Bergeron and Christophe Hohlweg [**7**] recently gave the same description we give here.

REMARK 4.5. Theorem 2.4 can be modified to combine left enriched P-partitions and interior enriched P-partitions. When translated to generating functions, it implies that $\mathbf{\Pi}_n$ is a two-sided ideal in $\mathbf{\Pi}_n^{(\ell)}$, and hence $\mathfrak{P}_n$ is an ideal in $\mathfrak{P}_n^{(\ell)}$.

5. Specializations

Define the polynomial $\Omega(P; x)$, called the *enriched order polynomial*, over all positive integers k by

$$\Omega(P; k) := \Lambda(P)(\underbrace{1, 1, \ldots, 1}_{k}, 0, 0, \ldots),$$

meaning we set $z_i = 1$ for $i = 1, \ldots, k$, and $z_i = 0$ for $i > k$. It turns out that for $\pi \in \mathfrak{S}_n$, $\Omega(\pi; x)$ is an even or odd polynomial of degree n that only depends on the *number* of interior peaks of π (see [**22**], Proposition 4.4). We can use order polynomials to study commutative peak algebras, spanned by sums of permutations with the same *number* of peaks. We sketch the idea for the interior peaks case.

Let E_i be the sum of all permutations with i interior peaks, and let $\Omega(i; x)$ denote the order polynomial for any such permutation. Now we define:

$$\rho(x) := \sum_{\pi \in \mathfrak{S}_n} \Omega(\pi; x/2)\pi = \sum_{i=1}^{\lfloor \frac{n+1}{2} \rfloor} \Omega(i; x/2) E_i = \begin{cases} \displaystyle\sum_{i=1}^{n/2} e_i x^{2i} & \text{if } n \text{ is even,} \\ \displaystyle\sum_{i=1}^{(n+1)/2} e_i x^{2i-1} & \text{if } n \text{ is odd.} \end{cases} \tag{5.1}$$

The function $\rho(x)$ is a polynomial in x with coefficients in the group algebra $\mathbb{Q}[\mathfrak{S}_n]$ (we now need to work over the rational numbers). From Theorem 2.4 we can obtain the following.

THEOREM 5.1. *As polynomials in x and y with coefficients in the group algebra $\mathbb{Q}[\mathfrak{S}_n]$, we have:*

$$\rho(x)\rho(y) = \rho(xy).$$

$(\delta_{ij} e_i^* e_j^*)$	e_j	$\overline{e}_j$	$e_j^{(\ell)}$	$e_j^{(r)}$
e_i	e_i	e_i	e_i	e_i
$\overline{e}_i$	$\overline{e}_i$	$\overline{e}_i$	$\overline{e}_i$	$\overline{e}_i$
$e_i^{(\ell)}$	e_i	$\overline{e}_i$	$e_i^{(\ell)}$	$e_i^{(r)}$
$e_i^{(r)}$	$\overline{e}_i$	e_i	$e_i^{(r)}$	$e_i^{(\ell)}$

TABLE 1. Multiplication table for type A coefficients.

What the theorem tells us is that the coefficients e_i as defined by (5.1) are mutually orthogonal idempotents. With a little more work, we see that the span of the E_i is the same as the span of the e_i, so that the sums of permutations with common peak numbers span a commutative $\lfloor \frac{n+1}{2} \rfloor$-dimensional subalgebra of the group algebra. This algebra and its left peak variant were introduced in [**1**].

We can use the same approach to get similar results for other commutative peak algebras, given by the span of sums of permutations with the same number of right peaks, exterior peaks, and type B peaks. Table 1 summarizes how the different type A peak idempotents interact (though $e_i^{(r)}$ is not technically idempotent). We remark that while the number of right peaks does not give a basis on its own, its multiplicative closure is still a proper subalgebra.

6. Negative results

Before anything was proved, the type B peak algebra was found experimentally, and along the road to its discovery there were several dead-end definitions. To save others the trouble of these detours, we finish with a list of some subalgebras that *do not* exist in general. For the symmetric group, the sums of permutations with the same set of right peaks do not form an algebra, nor do the sums of permutations with the same set of exterior peaks. For the hyperoctahedral group, we do not get a proper subalgebra by taking the sums of permutations with the same: interior peak set (i.e., ignoring peaks at 0), number of interior peaks, exterior peak set, or exterior peak number.

References

[1] M. Aguiar, N. Bergeron, and K. Nyman, *The peak algebra and the descent algebras of types B and D*, Transactions of the American Mathematical Society **356** (2004), 2781–2824.

[2] M. Aguiar, K. Nyman, and R. Orellana, *New results on the peak algebra*, math.CO/0406418.

[3] F. Bergeron and N. Bergeron, *A decomposition of the descent algebra of the hyperoctahedral group. I.* Journal of Algebra **148** (1992), 86–97.

[4] F. Bergeron and N. Bergeron, *Orthogonal idempotents in the descent algebra of B_n and applications*, Journal of Pure and Applied Algebra **79** (1992), 109–129.

[5] N. Bergeron, *A decomposition of the descent algebra of the hyperoctahedral group. II.* Journal of Algebra **148** (1992), 98–122.

[6] N. Bergeron, F. Hivert, and J.-Y. Thibon, *The peak algebra and the Hecke-Clifford algebras at q=0*, math.CO/0304191.

[7] N. Bergeron and C. Hohlweg, *Coloured peak algebras and Hopf algebras*, math.AC/0505612.

[8] N. Bergeron, S. Mykytiuk, F. Sottile, and S. van Willigenburg, *Shifted quasi-symmetric functions and the Hopf algebra of peak functions*, Discrete Mathematics **256** (2002), 57–66.

[9] L. Billera, S. Hsiao, and S. van Willigenburg, *Peak quasisymmetric functions and Eulerian enumeration*, Advances in Mathematics **176** (2003), 248–276.

[10] P. Cellini, *A general commutative descent algebra*, Journal of Algebra **175** (1995), 990–1014.

[11] P. Cellini, *A general commutative descent algebra. II. The case C_n*, Journal of Algebra **175** (1995), 1015–1026.

[12] P. Cellini, *Cyclic Eulerian elements*, European Journal of Combinatorics **19** (1998), 545–552.

[13] C.-O. Chow, *Noncommutative symmetric functions of type B*, Ph.D. thesis (2001).

[14] J. Fulman, *Affine shuffles, shuffles with cuts, the Whitehouse model, and patience sorting*, Journal of Algebra **231** (2000), 614–639.

[15] A. Garsia and C. Reutenauer, *A decomposition of Solomon's descent algebra*, Advances in Mathematics **77** (1989), 189–262.

[16] I. Gessel, *Multipartite P-partitions and inner products of skew Schur functions*, Contemporary Mathematics **34** (1984), 289–317.

[17] D. Krob, J.-Y. Thibon *Higher order peak algebras*, math.CO/0411407.

[18] J.-L. Loday, *Opeérations sur l'homologie cyclique des algèbres commutatives*, Inventiones mathematicae **96** (1989), 205–230.

[19] S. Mahajan, *Shuffles on Coxeter groups*, math.CO/0108094.

[20] K. Nyman, *The peak algebra of the symmetric group*, Journal of Algebraic Combinatorics **17** (2003), 309–322.
[21] T. K. Petersen, *Cyclic descents and P-partitions*, Journal of Algebraic Combinatorics **22** (2005), 343–375.
[22] T. K. Petersen, *Enriched P-partitions and peak algebras*, Advances in Mathematics, to appear.
[23] V. Reiner, *Signed posets*, Journal of Combinatorial Theory Series A **62** (1993), 324–360.
[24] M. Schocker, *The peak algebra of the symmetric group revisited*, Advances in Mathematics **192** (2005), 259–309.
[25] L. Solomon, *A Mackey formula in the group ring of a finite Coxeter group*, Journal of Algebra **41** (1976), 255–264.
[26] J. Stembridge, *Enriched P-partitions*, Transactions of the American Mathematical Society **349** (1997), 763–788.

Department of Mathematics, Brandeis University, Waltham, MA, USA, 02454
E-mail address: tkpeters@brandeis.edu
URL: http://people.brandeis.edu/~tkpeters

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A new construction of the Loday-Ronco algebra

Maxime Rey

Abstract. We provide a new construction of the Loday-Ronco algebra by realizing it in terms of non-commutative polynomials in infinitely many variables. This construction relies on a bijection between words and labeled binary trees which can be regarded as a kind of degenerate Robinson-Schensted correspondence and leads to a new Knuth type correspondence involving binary trees.

Résumé. Nous donnons une nouvelle construction de l'algèbre de Loday et Ronco en termes de polynômes non-commutatifs en une infinité de variables. Cette construction repose sur une bijection entre les mots et les arbres binaires étiquetés qui permet de définir une correspondence de type Robinson-Schensted dégénérée et aboutit à la construction d'une nouvelle correspondence de type Knuth mettant en jeu les arbres binaires.

1. Introduction

We give a new construction of the Loday-Ronco algebra of the plane binary trees, also known as the free dendriform dialgebra on one generator (see [**8**]). We first use, in Section 3, the argument given in [**9**] on dendriform trialgebras in order to prove that the algebra of non-commutative polynomials in infinitely variables can be endowed with the structure of a dendriform dialgebra. We then state that the sub-dialgebra generated by the sum of the letters is free as a dendriform dialgebra. To prove this statement, we introduce in Section 4 a bijection between words and labeled binary trees, which leads to a degenerate kind of Robinson-Schensted correspondence, reminiscent of the degenerate correspondence with ribbons and quasi-ribbon diagrams in [**5**], and dual to the Sylvester Schensted Algorithm of [**2**] as explained in Section 5. This leads, in Section 6, to a new Knuth type correspondence between integer matrices and some pairs of labeled binary trees. In Section 7 we define a family of elements indexed by binary trees that permits to prove that our dendriform dialgebra on one generator is free, using a bijection between binary trees and its elements.

2. Preliminaries and Notations

In this paper, $\mathbb{K}$ stands for a field of any characteristic. Let $A = \{a_1, a_2, ...\}$ be a totally ordered (infinite) alphabet and denote by A^* the free monoid on A. The map $max : A^* \to A$ maps a word w to its greatest letter, according to the total order of the alphabet A. We denote by $Std(w)$ the standardized word of $w \in A^*$ defined as follows.

Definition 2.1. Let $w = w_1 \cdots w_n \in A^*$ and $Std(w) = w'_1 \cdots w'_n$. Then, $\forall i, j \in [1, n]$ with $i \neq j$:

- if $w_i > w_j$ then $w'_i > w'_j$,
- if $w_i = w_j$ with $i > j$, then $w'_i > w'_j$,

such that $Std(w)$ is a permutation.

For example $Std(abcadbcaa) = 157296834$. For a word $w \in A^*$ and a subset B of A, $w|_B$ stands for the subword of w obtained by erasing the letters which are not in B. The evaluation of a word w is the vector $ev(w) = (|w|_{a_1}, |w|_{a_2}, ...)$.

2000 *Mathematics Subject Classification.* Primary 05-06; Secondary 05E99.

Key words and phrases. Algebraic combinatorics, Robinson-Schensted correspondence, Knuth correspondence, words, dendriform dialgebras.

We will denote by BT the set of all plane binary trees and LBT stands for the set of labeled plane binary trees. We denote by $Shape$ the map that for a labeled binary tree forgets its labels and returns a binary tree of the same shape.

Let w be an element of A^* without repetition. Its decreasing tree $\mathcal{T}(w)$ is an element of LBT obtained as follows: the root is labeled by the greatest letter, n of w, and if $w = unv$, where u and v are words without repetition, the left subtree is $\mathcal{T}(u)$ and the right subtree is $\mathcal{T}(v)$. Moreover, we associate the empty tree to the empty word.

Let w be an element of A^*, we will denote by $\mathcal{B}(w)$ its associated binary search tree. It is obtained by reading w from *right-to-left*, each letter being inserted into a binary search tree in the following way: if the tree is empty, one creates a node labeled by the letter; otherwise, this letter is recursively inserted into the left (resp. right) subtree if it is smaller or equal than (resp. greater than) the root. Exemples will be given further.

A biletter on A is a pair $(a, b) \in A \times A$ which we will write $\begin{bmatrix} a \\ b \end{bmatrix}$ for convenience. A biword $\begin{bmatrix} u \\ v \end{bmatrix}$ on A^* is a concatenation of biletters $\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \end{bmatrix}\begin{bmatrix} u_2 \\ v_2 \end{bmatrix}...\begin{bmatrix} u_n \\ v_n \end{bmatrix}$. We denote by $\begin{bmatrix} u' \\ v' \end{bmatrix}$ the nondecreasing rearrangement of $\begin{bmatrix} u \\ v \end{bmatrix}$ for the lexicographic order with priority on the top row, and by $\begin{bmatrix} u'' \\ v'' \end{bmatrix}$ the nondecreasing rearrangement for the lexicographic order with priority on the bottom row. Let $\langle\rangle$ denote the linear map from $\mathbb{K}[[A, B]]$ to $\mathbb{K}\langle\langle A\rangle\rangle \otimes \mathbb{K}\langle\langle B\rangle\rangle$, defined by $\langle\begin{pmatrix} u \\ v \end{pmatrix}\rangle = u'' \otimes v'$, with $u \in A$ and $v \in B$.

3. The free dendriform dialgebra embedded in words

Following a suggestion of [**9**], we define the following operations on words.

Definition 3.1. For all $u, v \in A^+$,

$$u \leftharpoonup v := \begin{cases} uv & \text{if } max(u) > max(v) \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

$$u \rightharpoondown v := \begin{cases} uv & \text{if } max(u) \leq max(v) \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Clearly, the usual operation of concatenation $\cdot$ on A^* can be written this way:

$$\cdot \;=\; \leftharpoonup + \rightharpoondown \quad . \tag{3.3}$$

Proposition 3.1. *$(\bigoplus_{n\geq 0} \mathbb{K}[A], \leftharpoonup, \rightharpoondown)$ is a dendriform dialgebra, in the sense of* [**7**].

Proof – Since (3.3) holds by definition, we only have to check the following three relations:

$$\begin{cases} (u \leftharpoonup v) \leftharpoonup w &= u \leftharpoonup (v \cdot w)\,, & (i) \\ (u \rightharpoondown v) \leftharpoonup w &= u \rightharpoondown (v \leftharpoonup w)\,, & (ii) \\ (u \cdot v) \rightharpoondown w &= u \rightharpoondown (v \rightharpoondown w)\,, & (iii) \end{cases} \tag{3.4}$$

whith $u, v, w \in A^+$. Notice first that for all these relations, there are only two possible values for each side, which are 0 and uvw.

(i) We first prove that

$$(u \leftharpoonup v) \leftharpoonup w = uvw \iff u \leftharpoonup (v \cdot w) = uvw\,. \tag{3.5}$$

By definition, we have $(u \leftharpoonup v) \leftharpoonup w = uvw$ if and only if $max(u) > max(v)$ and $max(uv) > max(w)$. Since if $max(u) > max(v)$ then $max(uv) = max(u)$, we have a necessary and sufficient condition

$$max(u) > max(v) \;\wedge\; max(u) > max(w)\,. \tag{3.6}$$

On the right-hand side, we have $u \leftharpoonup (v \cdot w) = uvw$ if and only if

$$max(u) > max(vw)\,. \tag{3.7}$$

Since (3.6) and (3.7) are clearly equivalent, we have proved assertion (3.5).

(ii) We have $(u \rightharpoondown v) \leftharpoonup w = uvw$ if and only if

$$max(u) \leq max(v) \ \wedge \ max(w) < max(uv)\,. \tag{3.8}$$

But since $max(u) \leq max(v)$, assertion (3.8) is equivalent to

$$max(u) \leq max(v) \ \wedge \ max(w) < max(v)\,. \tag{3.9}$$

Moreover, $u \rightharpoondown (v \leftharpoonup w) = uvw$ if and only if

$$max(w) < max(v) \ \wedge \ max(u) \leq max(vw)\,,$$

which can be rewritten as

$$max(w) < max(v) \ \wedge \ max(u) \leq max(v)\,, \tag{3.10}$$

due to $max(w) < max(v)$. It results that $(u \rightharpoondown v) \leftharpoonup w = uvw$ if and only if $u \rightharpoondown (v \leftharpoonup w) = uvw$, by equivalence of assertions (3.9) and (3.10).

(iii) We have $u \rightharpoondown (v \rightharpoondown w) = uvw$ if and only if

$$max(u) \leq max(vw) \ \wedge \ max(v) \leq max(w)\,,$$

which can immediately be rewritten as

$$max(u) \leq max(w) \ \wedge \ max(v) \leq max(w)\,, \tag{3.11}$$

due to $max(v) \leq max(w)$. Moreover, we have $(u \cdot v) \rightharpoondown w = uvw$ if and only if $max(uv) \leq max(w)$, which is equivalent to (3.11). Hence it results $(u \cdot v) \rightharpoondown w = uvw$ if and only if $u \rightharpoondown (v \rightharpoondown w) = uvw$. ■

Consider now the sub-dialgebra $\mathfrak{D}$ of $(\bigoplus_{n\geq 0} \mathbb{K}[A], \leftharpoonup, \rightharpoondown)$ generated by

$$P_\bullet := \sum_{a\in A} a\,.$$

There are two basis elements in the homogeneous component of degree 2 of $\mathfrak{D}$:

$$P_\bullet \leftharpoonup P_\bullet = \sum_{a<b} ba\,,$$

$$P_\bullet \rightharpoondown P_\bullet = \sum_{a\leq b} ab\,.$$

There are only five independent basis elements in the homogeneous component of degree 3 of $\mathfrak{D}$:

$$P_\bullet \leftharpoonup (P_\bullet \leftharpoonup P_\bullet) = \sum_{a<b<c} cba\,, \tag{3.12}$$

$$P_\bullet \rightharpoondown (P_\bullet \leftharpoonup P_\bullet) = \sum_{a<b;\ a'\leq b} a'ba\,, \tag{3.13}$$

$$P_\bullet \leftharpoonup (P_\bullet \rightharpoondown P_\bullet) = \sum_{a\leq b<c} cab\,, \tag{3.14}$$

$$(P_\bullet \leftharpoonup P_\bullet) \rightharpoondown P_\bullet = \sum_{a<b\leq c} bac\,, \tag{3.15}$$

$$(P_\bullet \rightharpoondown P_\bullet) \rightharpoondown P_\bullet = \sum_{a\leq b\leq c} abc\,, \tag{3.16}$$

since following equalities hold:

$$(P_\bullet \leftharpoonup P_\bullet) \leftharpoonup P_\bullet = \sum_{a<b;\ a'<b} baa' = (3.12) + (3.14)\,,$$

$$(P_\bullet \rightharpoondown P_\bullet) \leftharpoondown P_\bullet = \sum_{a \leq b;\ a' < b} aba' = (3.13)\,,$$
$$P_\bullet \rightharpoondown (P_\bullet \rightharpoondown P_\bullet) = \sum_{a \leq b;\ a' \leq b} a'ab = (3.15) + (3.16)\,,$$

which are exactly relations (3.4). We can now state our main result.

Theorem 3.2. *$\mathfrak{D}$ is the free dendriform dialgebra on one generator.*

The remainder of this article will provide appropriate tools to prove this statement and exhibit some interesting remarks about them.

4. Algorithm Ψ

We first describe an algorithm Ψ that associates with a word a labelled plane binary tree. Then it will be possible to associate a plane binary tree with a word by considering only the shape of the labelled tree produced by algorithm Ψ. To this purpose we introduce a map

$$\Gamma : LBT \times A \longrightarrow LBT\,,$$

which can be recursively defined as follows:

$$\Gamma\Big(\begin{smallmatrix} & y & \\ \alpha & & \beta\end{smallmatrix}, x\Big) = \begin{cases} \begin{smallmatrix} & & x\\ & y & \\ \alpha & & \beta \end{smallmatrix} & \text{if } x \geq y \\ \begin{smallmatrix} & y & \\ \alpha & & \Gamma(\beta,x)\end{smallmatrix} & \text{if } x < y\end{cases}, \tag{4.1}$$

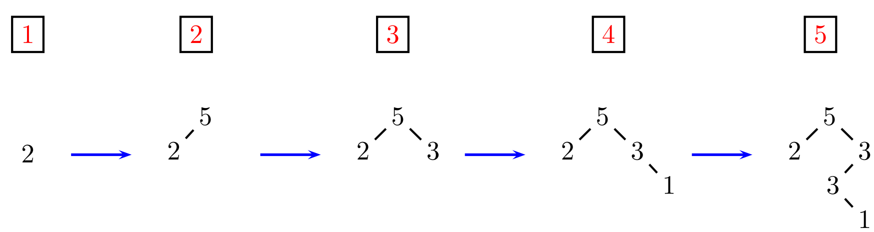
$$\Gamma(\circ, x) = x\,,$$

where $\circ$ stands for the empty tree and $\alpha, \beta \in LBT$.

Definition 4.1. Consider the following function Ψ:

$$\begin{array}{cccc} \Psi : & A^* & \longrightarrow & LBT\,, \\ & w = w_1 w_2 \dots w_n & \longmapsto & \begin{cases} \Gamma(\Psi(w_1 \dots w_{n-1}), w_n) & \text{if } n \geq 2\,, \\ \Gamma(\circ, w_1) & \text{if not}\,. \end{cases} \end{array} \tag{4.2}$$

For example, using the alphabet $\mathbb{N}_{>0}$ with the natural order on integers, we apply Ψ to the word 25313 as follows:



$$= \underbrace{\Gamma(\dots(\Gamma}_{5}(\circ, 2), 5), 3), 1), 4)\,,$$

starting with $\Gamma(\circ, 2)$.

Proposition 4.1. *Let $w \in A^*$. Then, $Shape(\Psi(w)) = Shape(\Psi(Std(w)))$.*

Proof – We proceed by induction. The initial case, considering a word of size 1, is obvious since there is only one tree of size 1. Assuming that this property is satisfied for words of length $n-1$, we check that it is true for words of length n due to Definition 2.1 and the inductive definition (4.1). ■

We recall that $\mathcal{T}$ stands for the decreasing tree algorithm which associates to a permutation its decreasing tree.

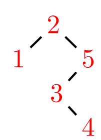
PROPOSITION 4.2. $\forall \sigma \in \mathfrak{S}, \Psi(\sigma) = \mathcal{T}(\sigma)$.

Proof – We proceed by induction. This property is obvious for the permutation of size 1. Assume that this property is satisfied for permutations of all sizes smaller than n and $\sigma = \sigma_1 \cdot max(\sigma) \cdot \sigma_2$ is a permutation of size n. By the inductive Definition 4.1 it is clear that the root of $\Psi(\sigma)$ is labeled by $max(\sigma)$, and that the left subtree of the root of $\Psi(\sigma)$ will be $\mathcal{T}(\sigma_1)$ by the inductive hypothesis, and similarly the right subtree of the root of $\Psi(\sigma)$ will be $\mathcal{T}(\sigma_2)$. ■

Hence algorithm Ψ is clearly a generalization on words of the well-known decreasing tree algorithm for permutations. From Propositions 4.2 and 4.1 the following result is immediate.

PROPOSITION 4.3. *The algorithm* Ψ *is injective.*

Let $w \in A^*$, we now consider the labeled binary tree having the same shape as $\Psi(w)$ and for which the label of each node is the step of its insertion in the tree $\Psi(w)$. We will denote it by $\psi(w)$. From the previous calculation of $\Psi(25313)$ we get the following tree:



.

We notice that $\psi(25313)$ is the binary search tree of $Std(25313)^{-1} = 41352$. We develop a new Schensted-like correspondence and a new Schensted-Knuth-like correspondence from this consideration.

5. The Co-Sylvester Schensted Algorithm ($CSSA$)

The Sylvester Schensted Algorithm (SSA) has been introduced in [**2**]. From Algorithms Ψ and ψ we give a dual correspondence of SSA.

DEFINITION 5.1. We note $CSSA$ the *Co-Sylvester Schensted Algorithm* which sends a word $w \in A^*$ to the pair

$$(\Psi(w), \psi(w)) .$$

Algorithm $CSSB$ sends this pair to the word obtained by reading the labels of $\Psi(w)$ in the order of the corresponding labels in $\psi(w)$.

We know that

LEMMA 5.2 ([**2**]). *Let* w *be a word and* $\sigma = Std(w)$. *Then*

$$Shape(\mathcal{B}(w)) = Shape(\mathcal{B}(\sigma)) = Shape(\mathcal{T}(\sigma^{-1})) .$$

We generalize this result in terms of biwords.

LEMMA 5.3. *For any biword* $\begin{bmatrix} u \\ v \end{bmatrix}$, $Shape(\Psi(v')) = Shape(\mathcal{B}(u''))$.

Proof – Using notations of Section 2, the standardization being compatible under transposition of two letters it follows that $Std(v)' = Std(v')$ and $Std(u)'' = Std(u'')$. It is well-known that the inverse of a permutation $\begin{bmatrix} Id \\ \sigma \end{bmatrix}$ is the biword $\begin{bmatrix} \sigma^{-1} \\ \sigma^{\uparrow} = Id \end{bmatrix}$ where $w^{\uparrow}$ denotes the nondecreasing rearrangement of a word w. Then, since $Std(v)' = (Std(u)'')^{-1}$, we have

$$Std(v') = (Std(u''))^{-1} .$$

Hence, using Lemma 5.2 and Proposition 4.1 we obtain these equivalences:

$$Shape(\Psi(v')) = Shape(\mathcal{T}(Std(v'))) = Shape(\mathcal{B}((Std(v'))^{-1})) = Shape(\mathcal{B}(u''))\,.$$

■

This allows us to state an analogous of the Schützenberger Theorem on tableaux for binary trees.

Theorem 5.4. *$\forall \sigma \in \mathfrak{S}$, $\psi(\sigma) = \mathcal{B}(\sigma^{-1})$.*

Proof – By definition of ψ, $\psi(\sigma)$ have the same shape as $\Psi(\sigma) = \mathcal{T}(\sigma)$. Moreover, by induction it is also clear that $\psi(\sigma)$ is a binary search tree. Indeed, assuming the inductive hypothesis for trees with $k-1$ nodes, the k-th node is, by definition of the algorithm ψ, labeled by k whereas the $k-1$ remaining nodes are labeled by elements of $[1,(k-1)]$. Moreover, by recursive definition (4.1) two cases arise:

- The k-th node is inserted at the root and then it is still a binary search tree since the $k-1$ remaining nodes are in its left subtree.
- The k-th node is inserted somewhere in the right subtree, and so it is still a binary search tree since it is greater than the root, and by inductive hypothesis.

Nevertheless, in general there is not a single permutation σ' such that $\mathcal{B}(\sigma') = T$, with T a given binary search tree (see [**2**] for the exact description of such sets of permutations). But it is clear, by definition of ψ, that at each step k of the insertion algorithm $CSSA$, $\Psi(\sigma_1\sigma_2...\sigma_k)$ and $\psi(\sigma_1\sigma_2...\sigma_k)$ have the same shape. Moreover setting

$$\begin{bmatrix} u(k)' \\ v(k)' \end{bmatrix} = \begin{bmatrix} 1 & 2 & \cdots & k \\ \sigma_1 & \sigma_2 & \cdots & \sigma_k \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} u(k)'' \\ v(k)'' \end{bmatrix} = \begin{bmatrix} \sigma^{-1}|_{[1,k]} \\ 1\,2\,\cdots\,k \end{bmatrix}\ ,$$

from Lemma 5.3 we have that at each step of the insertion algorithm, $\Psi(\sigma_1\sigma_2\cdots\sigma_k)$ and $\mathcal{B}(\sigma^{-1}|_{[1,k]})$ have the same shape. Hence, at the last step, $\psi(\sigma) = \mathcal{B}(\sigma^{-1})$. ■

Proposition 5.1. *$\forall\, w \in A^*, \psi(w) = \psi(Std(w))$.*

Proof – From Definition 2.1 and Proposition 4.1, it is immediate since at each step k, $\Psi(w_1\cdots w_k)$ and $\Psi(Std(w_1\cdots w_k))$ have the same shape. ■

Hence from Theorem 5.4 and Proposition 5.1 we immediately obtain the following result on words.

Corollary 5.5. *$\forall\, w \in A^*$, $\psi(w) = \mathcal{B}(Std(w)^{-1})$.*

From Theorem 5.4 and Proprosition 4.2 it is straightforward that $CSSA$ is the dual Schensted-like correspondence of SSA of [**2**] in the following sense.

Proposition 5.2. *$\forall \sigma \in \mathfrak{S}$,*

$$CSSA(\sigma) = (\Psi(\sigma), \mathcal{B}(\sigma^{-1})) \Longleftrightarrow SSA(\sigma^{-1}) = (\mathcal{B}(\sigma^{-1}), \Psi(\sigma))\,.$$

It is interesting to notice that these two correspondences look quite similar to the two Robinson-Schensted type correspondences on ribbons and quasi-ribbons, introduced in [**5**].

6. The Sylvester Schensted-Knuth correspondence

We first recall that there is an easy bijection between integer matrices and commutative biwords on A^* which consists to repeat m_{ij} times the biletter $\begin{pmatrix} \mathrm{i} \\ \mathrm{j} \end{pmatrix}$ for a matrix $M = (m_{ij})_{(i,j)\in[1,n]\times[1,m]}$ of dimensions $n \times m$. For example commutative biwords $\begin{pmatrix} 1111222333 \\ 1113123133 \end{pmatrix}$ and $\begin{pmatrix} 1112321233 \\ 1111123333 \end{pmatrix}$ (which are equal) have the same corresponding matrix which is :

$$\begin{pmatrix} 3 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}.$$

Definition 6.1. Let $\begin{pmatrix} u \\ v \end{pmatrix}$ be a commutative biword and $\begin{bmatrix} u' \\ v' \end{bmatrix}$, $\begin{bmatrix} u'' \\ v'' \end{bmatrix}$ be the two biwords associated with $\begin{bmatrix} u \\ v \end{bmatrix}$ as explained in Section 2 . We note κ_S the *Sylvester Schensted-Knuth correspondence* defined as follows:

$$\kappa_S \begin{pmatrix} u \\ v \end{pmatrix} := (\Psi(v'), \mathcal{B}(u''))\,.$$

This definition holds since for any biletter permutation of $\begin{bmatrix} u \\ v \end{bmatrix}$, v' and u'' remain the same. Moreover by Lemma 5.3 the following Proposition is straightforward.

Proposition 6.1. *For any commutative biword* $\begin{pmatrix} u \\ v \end{pmatrix}$ *,* $Shape(\Psi(v')) = Shape(\mathcal{B}(u''))$.

We notice that $CSSA$ is recovered by encoding a word w by $\begin{pmatrix} 1 & 2 & \cdots & m \\ w_1 & w_2 & \cdots & w_m \end{pmatrix}$.

In order to prove that κ_S is a bijection, we proceed as Lascoux, Leclerc and Thibon did in [**6**] for the usual Knuth correspondence [**4**].

Theorem 6.2. *The algorithm κ_S is a bijection.*

Proof – Using again arguments of the proof of Lemma 5.3 we obtain $\mathcal{B}(Std(u'')) = \mathcal{B}(Std(v')^{-1})$. Then, applying Theorem 5.4 we have that $\mathcal{B}(Std(v')^{-1}) = \psi(Std(v'))$ and from Corollary 5.5 we get:

$$\mathcal{B}(Std(u'')) = \psi(v')\,.$$

This means that $\mathcal{B}(u'')$ is the unique binary search tree of evaluation $ev(u'')$ such that $\mathcal{B}(Std(u'')) = \psi(v')$. ■

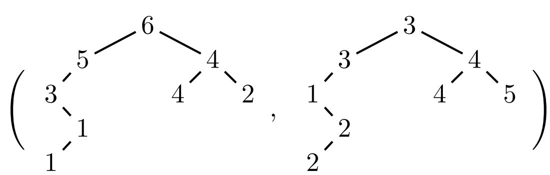
An easy remark is the following:

$$\kappa_S \begin{pmatrix} u \\ v \end{pmatrix} = (\Psi(v'), \mathcal{B}(u'')) \Longleftrightarrow \kappa_S \begin{pmatrix} v \\ u \end{pmatrix} = (\Psi(u''), \mathcal{B}(v'))\,,$$

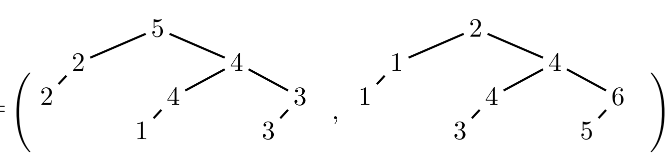
which generalizes results of Section 5. Nevertheless the symmetry of the usual Knuth correspondence is broken for P and Q symbols. At last, we give a full example of our construction.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 2\,1\,3\,3\,5\,4\,2\,4 \\ 1\,3\,6\,5\,2\,4\,1\,4 \end{bmatrix}, \quad \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} 1\,2\,2\,3\,3\,4\,4\,5 \\ 3\,1\,1\,5\,6\,4\,4\,2 \end{bmatrix} \quad \begin{bmatrix} u'' \\ v'' \end{bmatrix} = \begin{bmatrix} 2\,2\,5\,1\,4\,4\,3\,3 \\ 1\,1\,2\,3\,4\,4\,5\,6 \end{bmatrix},$$

$(\Psi(v'), \mathcal{B}(u''))$ =

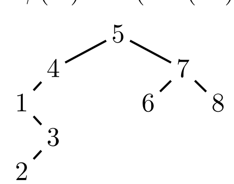


$(\Psi(u''), \mathcal{B}(v'))$ =



Since $Std(22514433) = 23816745$, we can check that $\psi(v') = \mathcal{B}(Std(u''))$:

$\psi(31156442) = \mathcal{B}(23816745) =$



.

This bijection leads to a Cauchy formula type for binary trees.

Proposition 6.2. *Let A and B be two non-commutative alphabets, such that A and B are commuting one with the other and words on alphabet B are quotiented by the sylvester congruence (see* [**2**]*). Then,*

$$\Big\langle \prod_{a\in A,\, b\in B}^{\rightarrow} \frac{1}{1-ab} \Big\rangle = \sum_{T\in BT} P_T(A)\otimes Q_T(B)\,,$$

where $(P_T)_{T\in BT}$ comes from Definition 7.1 and $(Q_T)_{T\in BT}$ is its dual basis introduced in [**2**].

Proof – We recall that the sylvester monoid introduced in [**2**] is the monoid such that two words having the same shape through algorithm $\mathcal{B}$ are equal. This proof needs the Free Cauchy identity mentionned in [**3**] and to be fully introduced in [**1**]:

$$\Big\langle \prod_{a\in A,\, b\in B}^{\rightarrow} \frac{1}{1-ab} \Big\rangle = \sum_{Std(v)=Std(u)^{-1}} u\otimes v\,. \tag{6.1}$$

Since v is an element of the sylvester monoid and by definition of $(Q_T)_{T\in BT}$ (see [**2**]), by Theorem 6.2 and by Proposition 6.1, right-hand side of Equation (6.1) can be rewritten as

$$\sum_{T\in BT}\Bigg(\sum_{Shape(\Psi(w))=T} w\Bigg)\otimes Q_T\,.$$

Using Definition 7.1 we obtain the desired equality. ∎

7. Back to dendriform structure

Definition 7.1. Let $T\in BT$. We define

$$P_T := \sum_{w;\ Shape(\Psi(w))=T} w\,. \tag{7.1}$$

As a special case we recover:

$$P_\bullet = \sum_{a\in A} a\,. \tag{7.2}$$

We now provide an algorithm that associates a plane binary tree to an element of the dendriform algebra $\mathfrak{D}$.

Definition 7.2. We consider

$$\Phi : BT \longrightarrow (\mathbb{K}[A], \leftharpoonup, \rightharpoondown)\,,$$

whose recursive definition is the following:

$$\left\{\begin{array}{lcll}
\Phi\Big(\begin{array}{ccc} & \bullet & \\ \diagup & & \diagdown \\ \alpha & & \beta \end{array}\Big) & = & (\Phi(\alpha)\rightharpoondown(P_\bullet\leftharpoonup\Phi(\beta)))\,, & (i)\\[2ex]
\Phi\Big(\begin{array}{cc} & \bullet \\ \diagup & \\ \alpha & \end{array}\Big) & = & (\Phi(\alpha)\rightharpoondown P_\bullet)\,, & (ii)\\[2ex]
\Phi\Big(\begin{array}{cc} \bullet & \\ & \diagdown \\ & \beta \end{array}\Big) & = & (P_\bullet\leftharpoonup\Phi(\beta))\,, & (iii)\\[2ex]
\Phi(\bullet) & = & P_\bullet\,, & (iv)
\end{array}\right. \tag{7.3}$$

where $\alpha,\beta\in BT$.

For example:

$$\Phi\left(\begin{array}{ccccc} & \bullet & & & \\ \diagup & & \diagdown & & \\ \bullet & & & \bullet & \\ & & \diagup & & \\ & \bullet & & & \\ & & \diagdown & & \\ & & & \bullet & \end{array}\right) = (P_\bullet\rightharpoondown(P_\bullet\leftharpoonup((P_\bullet\leftharpoonup P_\bullet)\rightharpoondown P_\bullet)))\,.$$

LEMMA 7.3. *Let $T \in BT$. Then, $P_T = \Phi(T)$.*

Proof – We proceed by induction. The initial case is immediate by (iv) of Definition 7.2. Assume that this property is satisfied for trees of size $n-1$. We consider a tree T of size n, with $n \geq 2$. Three kinds of trees are possible, according to wether their roots have only a non-empty right subtree or only a non-empty left subtree or finally have both left and right subtrees are non-empty. These cases correspond to $(i), (ii)$ and (iii) of (7.2).

(i) By Definition 7.1 this means that for all words appearing in the sum P_T, the $|\alpha|+1$ letter is greater or equal than its $|\alpha|$ first letters and greater to its $|\beta|$ last letters. This is the exact meaning of the right-hand side of (i) using (3.1), (3.2), remembering relation (ii) (associativity) of (3.4) and assuming inductive hypothesis.

(ii) By Definition 7.1 this means that for all words appearing in the sum P_T, their last letter is greater or equal than all others letters appearing in it. Then, by (3.2) and by induction hypothesis we have proved this case.

(iii) By Definition 7.1 this means that for all words appearing in the sum P_T, their first letter is greater than all others letters appearing in it. Hence by (3.1) and by induction hypothesis we have proved this case.

■

We now are able to provide the proof of Theorem 3.2:

Proof of Theorem 3.2 – Since the Loday-Ronco algebra of plane binary trees is the free dendriform algebra on one generator, we only have to check that the Hilbert serie of this subalgebra $\mathfrak{D}$ generated by $P_\bullet$ is counted by Catalan numbers. To this purpose we consider the familly of $(P_T)_{T\in BT}$. By Lemma 7.3 they are clearly elements of $\mathfrak{D}$. Moreover the intersection of any pairs of them is always empty, by construction. Then $(P_T)_{T\in BT}$ are linearly independent. ■

References

[1] G. Duchamp, F. Hivert, J.-C. Novelli, J.-Y. Thibon, Noncommutative symmetric functions VII (under construction).

[2] F. Hivert, J.-C. Novelli, J.-Y. Thibon, The algebra of binary search trees, Theoret. Computer Sci. 339 (2005), 129–165.

[3] F. Hivert, J.-C. Novelli, J.-Y. Thibon, Un analogue du monoïde plaxique pour les arbres binaires de recherche, C.R. Acad. Sci. Paris Sr I Math. 332 (2002) 577-580.

[4] D. Knuth, "Permutations, matrices, and generalized Young tableaux", Pacific J. Math. 34 (1970), 709-727.

[5] D. Krob and J.-Y. Thibon, Noncommutative symmetric functions IV: Quantum linear groups and Hecke algebras at q = 0, J. Alg. Comb. 6 (1997), 339–376.

[6] A. Lascoux, B. Leclerc, and J.-Y. Thibon, The Plactic Monoid, in: Algebraic Combinatorics on Words, M. Lothaire, Univ. Press Cambridge (2001).

[7] J.-L. Loday, Dialgebras and related operads, Springer Lecture Notes in Mathematics 1763 (2001), 7-66.

[8] J.-L. Loday, M. O. Ronco, Hopf algebra of the planar binary trees, Adv. Math. 139 (1998), no. 2, 293–309.

[9] J.-C Novelli and J.-Y. Thibon, Construction of dendriform trialgebras, preprint math.CO/0510218.

INSTITUT GASPARD MONGE, UNIVERSITÉ DE MARNE-LA-VALLÉE, 5 BOULEVARD DESCARTES, CHAMPS-SUR-MARNE, 77454 MARNE-LA-VALLÉE CEDEX 2, FRANCE

E-mail address: `rey@univ-mlv.fr`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Kazhdan-Lusztig immanants and products of matrix minors, II

Brendon Rhoades and Mark Skandera

ABSTRACT. We show that for each permutation w containing no decreasing subsequence of length k, the Kazhdan-Lusztig immanant $\mathrm{Imm}_w(x)$ vanishes on all matrices having k equal columns. We also construct new and simple inequalities satisfied by the minors of totally nonnegative matrices.

RÉSUMÉ. Nous démontrons que pour chaque permutation w qui ne contient aucune sous-suite décroissante de longeur k, l'immanant de Kazhdan-Lusztig $\mathrm{Imm}_w(x)$ s'annule sur toutes les matrices avec k colonnes identiques. Nous introduisons par ailleurs des inégalités simples et nouvelles satisfaites par les mineurs des matrices complètement non-negatives.

1. Introduction and Preliminaries

The Kazhdan-Lusztig basis $\{C'_w(q) \mid w \in S_n\}$ of the Hecke algebra $H_n(q)$, originally introduced in [**10**], has seen several applications in combinatorics and positivity. In [**14**] Rhoades and Skandera define the Kazhdan-Lusztig immanants via the Kazhdan-Lusztig basis and obtain various positivity results concerning linear combinations of products of matrix minors. Lam, Postnikov, and Pylyavskyy, in turn, use these results in [**11**] to resolve several conjectures in Schur positivity. In this paper, we further develop algebraic properties of the Kazhdan-Lusztig immanants and apply these immanants to obtain additional positivity results.

Fix $n \in \mathbb{N}$ and let $x = (x_{ij})_{1 \leq i,j \leq n}$ be a matrix of n^2 variables. For a pair of subsets $I, J \subseteq [n]$, with $|I| = |J|$, define the (I, J)-minor of x, denoted $\Delta_{I,J}(x)$, to be the determinant of the submatrix of x indexed by rows in I and columns in J. We adopt the convention that the empty minor $\Delta_{\emptyset,\emptyset}(x)$ is equal to 1. An $n \times n$ matrix A is said to be *totally nonnegative* (TNN) if every minor of A is a nonnegative real number. A polynomial $p(x)$ in n^2 variables is called totally nonnegative if whenever $A = (a_{i,j})_{1 \leq i,j \leq n}$ is a totally nonnegative matrix, $p(A) \underset{\mathrm{def}}{=} p(a_{1,1}, \ldots, a_{n,n})$ is a nonnegative real number. [**2**], [**3**], [**9**], [**13**], [**12**], [**19**] give a graph theoretic characterization of totally nonnegative matrices which is used by Rhoades and Skandera in [**15**] and [**14**] to construct several examples of totally nonnegative polynomials.

Let H denote the infinite array $(h_{j-i})_{i,j \geq 1}$, where h_i denotes the complete homogeneous symmetric function of degree i. (see, for example, [**18**]) Here we use the convention that $h_i = 0$ whenever $i < 0$. A polynomial $p(x)$ in n^2 variables is called *Schur nonnegative* (SNN) if whenever K is an $n \times n$ submatrix of H, the symmetric function $p(K)$ is a nonnegative linear combination of Schur functions. By the Jacobi identity, the determinant is a trivial example of a SNN polynomial.

For $i \in [n-1]$, let s_i denote the adjacent transposition in S_n which is written $(i, i+1)$ in cycle notation. For a fixed $w \in S_n$, call an expression $s_{i_1} \cdots s_{i_\ell}$ representing w *reduced* if ℓ is minimal. In this case, define the *length* of w, denoted $\ell(w)$, to be ℓ.

For q a formal indeterminate, define the *Hecke algebra* $H_n(q)$ to be the $\mathbb{C}[q^{1/2}, q^{-1/2}]$-algebra with generators $T_{s_1}, \ldots, T_{s_{n-1}}$ subject to the relations

$$\begin{aligned} T_{s_i}^2 &= (q-1)T_{s_i} + q, && \text{for } i = 1, \ldots, n-1, \\ T_{s_i}T_{s_j}T_{s_i} &= T_{s_j}T_{s_i}T_{s_j}, && \text{if } |i-j| = 1, \\ T_{s_i}T_{s_j} &= T_{s_j}T_{s_i}, && \text{if } |i-j| \geq 2. \end{aligned}$$

For $w \in S_n$, define the Hecke algebra element T_w by

$$T_w = T_{s_{i_1}} \cdots T_{s_{i_\ell}},$$

where $s_{i_1} \cdots s_{i_\ell}$ is any reduced expression for w. Specializing at $q = 1$, the map $T_{s_i} \mapsto s_i$ induces an isomorphism between $H_n(1)$ and the symmetric group algebra $\mathbb{C}[S_n]$.

The elements $\{C'_v(q) \,|\, v \in S_n\}$ of the Kazhdan-Lusztig basis of $H_n(q)$ have the form

$$C'_v(q) = \sum_{w \leq v} P_{w,v}(q) q^{-\ell(v)/2} T_w, \tag{1.1}$$

where

$$\{P_{w,v}(q) \,|\, w, v \in S_n\}$$

are polynomials in $\mathbb{N}[q]$ called the Kazhdan-Lusztig polynomials. We recall a couple of elementary properties of the Kazhdan-Lusztig polynomials.

Lemma 1.1. *For $w, v \in S_n$, $P_{w,v}(q) \equiv 0$ if and only if $w \not\leq v$, where $\leq$ is (strong) Bruhat ordering. Also, $P_{w,w}(q) \equiv 1$.*

A polynomial $p(x)$ in n^2 variables is called an *immanant* if it belongs to the $\mathbb{C}$-linear span of $\{x_{1,w(1)} \cdots x_{n,w(n)} \mid w \in S_n\}$. Following [**14**], for $w \in S_n$, define the *w-Kazhdan-Lusztig immanant* by

$$\mathrm{Imm}_w(x) \underset{\mathrm{def}}{=} \sum_{v \in S_n} (-1)^{\ell(w)-\ell(v)} P_{w_0 v, w_0 w}(1) x_{1,v(1)} \cdots x_{n,v(n)}, \tag{1.2}$$

where w_o denotes the long element of S_n, written $n(n-1)\ldots 1$ in one-line notation. Specializing at $w = 1$, we have that $\mathrm{Imm}_1(x) = det(x)$.

It follows from Lemma 1.1 that the expression $(-1)^{\ell(w)-\ell(v)} P_{w_0 v, w_0 w}(1)$ is nonzero if and only if $w \leq v$ in the Bruhat order and that $P_{w_0 w, w_0 w}(1) = 1$. Therefore, the set $\{\mathrm{Imm}_w(x) \mid w \in S_n\}$ forms a basis for the vector space of immanants. The Kazhdan-Lusztig immanants are both TNN and SNN and various examples of TNN and SNN polynomials can be constructed by studying the cone generated by the Kazhdan-Lusztig immanants [**14**]. Moreover, when w is 321-avoiding, the Kazhdan-Lusztig immanant $\mathrm{Imm}_w(x)$ is satisfies a natural generalization of Lindström's Lemma [**15**].

2. Main

For $1 \leq k \leq n$, let $\Gamma_{n,k}$ denote the subset of $\mathbb{C}[x_{1,1}, \ldots, x_{n,n}]$ consisting of all products of the form $\Delta_{I_1,J_1}(x) \cdots \Delta_{I_k,J_k}(x)$, where $I_1, J_1, \ldots, I_k, J_k \subseteq [n]$, $I_1 \uplus \cdots \uplus I_k = J_1 \uplus \cdots \uplus J_k = [n]$, and $|I_j| = |J_j|$ for all $j \in [k]$. Here $\uplus$ denotes disjoint union. Elements of $\Gamma_{n,k}$ are sometimes called complementary products of minors. In [**15**], Kazhdan-Lusztig immanants are used to find that the dimension of span($\Gamma_{n,2}$) is equal to the n^{th} Catalan number C_n. In this paper we shall relate the dimension of span($\Gamma_{n,k}$) to pattern avoidence in S_n for arbitrary k.

For $k \in \mathbb{N}$, let $S_{n,k}$ denote the set of permutations in S_n which do not have a decreasing subsequence of length $k+1$. For example, in one-line notation, $S_{3,2} = \{123, 213, 132, 312, 231\}$. Notice that $S_{n,k} = S_n$ for all $k \geq n$. We start by examining the image of $S_{n,k}$ under the Robinson-Schensted correspondence.

Let $\leq_{LR}$ be the preorder on S_n defined in [**10**] and let $s_{[1,k]}$ be the longest element in the subgroup of S_n generated by $s_1, \ldots, s_{k-1}$.

Lemma 2.1. *Suppose $v \notin S_{n,k-1}$. Then we have $v \leq_{LR} s_{[1,k]}$.*

Proof. Given any permutation w, define the pair of tableaux $(P'(w), Q'(w))$ to be the image of w under the Robinson-Schensted column insertion correspondence. Let $\lambda'(w)$ be the shape of these tableaux.

A well-known property of the Robinson-Schensted correspondence implies that $\lambda'(v) \geq \lambda'(s_{[1,k]})$ in the dominance order. This dominance relation in turn is known to be equivalent to the partial order on Kazhdan-Lusztig cells induced by the preorder $\leq_{LR}$. Thus in the preorder $\leq_{LR}$, every permutation in the cell of v precedes every permutation in the cell of $s_{[1,k]}$. (See [**1**], [**6**, Sec. 1], [**8**, Appendix].) ☐

Proposition 2.2. *Suppose $A \in \mathrm{Mat}_n(\mathbb{C})$ has k equal rows and let $v \in S_{n,k-1}$. Then, $\mathrm{Imm}_v(A) = 0$.*

This result generalizes Proposition 3.14 of Rhoades and Skandera [**15**], which together with [**14**] implies that Proposition 2.2 holds when $k = 2$.

Proof. Define the element $[A]$ of $\mathbb{C}[S_n]$ by

$$[A] = \sum_{w \in S_n} a_{1,w(1)} \cdots a_{n,w(n)} w.$$

Let $i_1 < \cdots < i_k$ be the indices of k rows in A which are equal and let U be the subgroup of S_n which fixes all indices not contained in the set $\{i_1, \ldots, i_k\}$. Then

$$\sum_{u \in U} u$$

factors as $wz_{[1,k]}w'$ for some elements w, w' of S_n. It follows that $[A]$ factors as

$$\begin{aligned}[A] &= \left(\sum_{u \in U} u \right) f(A) \\ &= (wz_{[1,k]}w')f(A)\end{aligned}$$

for some group algebra element $f(A)$.

Let I be the two-sided ideal of $\mathbb{C}[S_n]$ spanned by $\{C'_u(1) \,|\, u \leq_{LR} s_{[1,k]}\}$ and let $\theta : \mathbb{C}[S_n] \to \mathbb{C}[S_n]/I$ be the canonical homomorphism. Clearly we have $\theta([A]) = 0$.

On the other hand, we have

$$\begin{aligned}\theta([A]) &= \theta \left(\sum_{w \in S_n} \mathrm{Imm}_w(A) C'_w(1) \right) \\ &= \sum_{w \in S_n} \mathrm{Imm}_w(A) \theta(C'_w(1)).\end{aligned}$$

Since $\theta(C'_w(1)) = 0$ for all permutations $w \leq_{LR} s_{[1,k]}$, we have

$$0 = \sum_{w} \mathrm{Imm}_w(A) \theta(C'_w(1)),$$

where the sum is over all permutations $w \not\leq_{LR} s_{[1,k]}$, i.e., those permutations having no decreasing subsequence of length k. Since the elements $\theta(C'_w(1))$ in this sum are linearly independent, we must have $\mathrm{Imm}_w(A) = 0$ for each permutation w having no decreasing subsequence of length k. ☐

Proposition 2.3. *Suppose $\Delta_{I_1,J_1}(x) \cdots \Delta_{I_k,J_k}(x) \in \Gamma_{n,k}$. Then, there exist $d_w \in \mathbb{C}$ such that $\Delta_{I_1,J_1}(x) \cdots \Delta_{I_k,J_k}(x) = \sum_{w \in S_{n,k}} d_w \mathrm{Imm}_w(x)$.*

Proof. The Kazhdan-Lusztig immanants form a basis for the vector space of immanants, so we may write

$$\Delta_{I_1,J_1}(x) \cdots \Delta_{I_k,J_k}(x) = \sum_{w \in S_n} d_w \mathrm{Imm}_w(x), \tag{2.1}$$

for some $d_w \in \mathbb{C}$. If $k \geq n$ the claim is trivial, so we assume that $k < n$. We show that $d_w = 0$ whenever $w \notin S_{n,k}$.

Suppose that $C \in Mat_n(\mathbb{C})$ has $k+1$ equal rows. Then, by the pigeonhole principle, there exist two equal rows of C indexed by integers lying in one of $I_1, \ldots, I_k$. Hence, $\Delta_{I_1,J_1}(C) \cdots \Delta_{I_k,J_k}(C) = 0$.

Now let $B = (b_{ij}) \in Mat_n(\mathbb{C})$ be defined by $b_{ij} = 1$ for all i and j. By Proposition 2.2, since $k < n$ we have that $Imm_w(B) = 0$ for every $w \neq w_o$. Also, $Imm_{w_o}(B) = 1$. By the above paragraph, $\Delta_{I_1,J_1}(B) \cdots \Delta_{I_k,J_k}(B) = 0$. Therefore, applying both sides of (2.1) to B, we get that $d_{w_o} = 0$.

For $l \in \mathbb{N}$, define $T_{n,l}$ to be the set difference $S_{n,l} \smallsetminus S_{n,l-1}$. Suppose that $k < m < n$ and suppose that for all p satisfying $m < p \leq n$ we have that $d_w = 0$ for every $w \in T_{n,p}$. Give the elements of $T_{n,m}$ a total order which is an extension of their Bruhat ordering and write $T_{n,m} = \{w_1 < w_2 < \cdots < w_h\}$. Let $t \in [h]$ and suppose by induction that $d_w = 0$ for $w \in \{w_{t+1}, \ldots, w_h\}$. Since $w_t \in T_{n,m}$, there exist $i_1 < i_2 < \cdots < i_m$ such that $w_t(i_1) > w_t(i_2) > \cdots > w_t(i_m)$. Let $D \in Mat_n(\mathbb{C})$ be the matrix obtained by replacing the rows $i_1, \ldots, i_m$ in the permutation matrix for w_t by rows of 1's. By Proposition 2.2, $Imm_w(D) = 0$ for every $w \in S_{n,m-1}$. By (1.1) we also have that $\mathrm{Imm}_w(D) = 0$ for every $w \not\leq w_t$ in the Bruhat order. Since $k < m$, we have that $\Delta_{I_1,J_1}(D)\cdots\Delta_{I_k,J_k}(D)=0$. Thus, applying both sides of (2.1) to D, we get that $d_{w_t}=0$ and the Proposition follows by induction.

$\square$

If $k=2$ in Proposition 2.3, results in [**15**] and [**14**] imply that the d_w must be nonnegative. For k arbitrary, Skandera [**16**] has given an elementary proof that whenever w avoids the patterns 3412 and 4231, (i.e., when the Schubert variety Γ_w corresponding to w is smooth), the coefficient d_w is also nonnegative. Using deeper properties of the dual canonical basis of $\mathcal{O}SL_n\mathbb{C}$, it is possible to show that the coefficients d_w are nonnegative for general k and w.

Proposition 2.4. $\dim(\mathrm{span}_{\mathbb{C}}(\Gamma_{n,k}))=|S_{n,k}|$.

Specializing at $k=2$, $S_{n,2}$ is the set of 321-avoiding permutations, so we have that $|S_{n,2}|=C_n$, the n^{th} Catalan number. Thus, this result is a generalization of Proposition 4.7 of [**15**].

Proof. By Proposition 2.3 we have that $\dim(\mathrm{span}_{\mathbb{C}}(\Gamma_{n,k}))\leq|S_{n,k}|$.

For each collection of sets $I_1,J_1,\ldots,I_k,J_k\subseteq[n]$ with $[n]=I_1\uplus\cdots\uplus I_k=J_1\uplus\cdots\uplus J_k$ and $|I_j|=|J_j|$ for each $j\in k$, let $min(I_1,J_1,\ldots,I_k,J_k)$ denote the unique minimal permutation in the Bruhat order which maps I_i into J_i for each $i\in[k]$. For example, if we set $n=6$, $I_1=\{1,3,6\},I_2=\{2,4\},I_3=\{5\},J_1=\{3,4,6\},J_2=\{1,5\},J_3=\{2\}$, we have that $min(I_1,J_1,I_2,J_2,I_3,J_3)=314526$ in one-line notation.

For $\Delta_{I_1,J_1}(x)\cdots\Delta_{I_k,J_k}(x)\in\Gamma_{n,k}$ it is easy to see that there exist $d_w\in\mathbb{C}$ such that

$$\Delta_{I_1,J_1}(x)\cdots\Delta_{I_k,J_k}(x)=\sum_{w\geq min(I_1,J_1,\ldots,I_k,J_k)}d_w x_{1,w(1)}\cdots x_{n,w(n)},$$

where $d_{min(I_1,J_1,\ldots,I_k,J_k)}=1$. In light of this, it suffices to show that for every permutation $w\in S_{n,k}$, there exists a collection of sets $I_1,J_1,\ldots,I_k,J_k\subseteq[n]$ such that $[n]=I_1\uplus\cdots\uplus I_k=J_1\uplus\cdots\uplus J_k$ and $|I_j|=|J_j|$ for each $j\in k$ and $w=min(I_1,J_1,\ldots,I_k,J_k)$. For then, we have that $dim(span_{\mathbb{C}}(\Gamma_{n,k}))\geq|S_{n,k}|$.

Let $w\in S_{n,k}$. Define a partial order on the set $P=\{(i,w(i))|i\in[n]\}$ by setting $(i,w(i))<(j,w(j))$ if $i<j$ and $w(i)<w(j)$. Now $\{(i_1,w(i_1)),\ldots,(i_m,w(i_m))\}\subseteq P$ with $i_1<\cdots<i_m$ is an antichain in P if and only if $(w(i_1),\ldots,w(i_m))$ is an decreasing subsequence of w. Hence, width$(P)<k+1$ (see [**17**] for definitions). By Dilworth's Theorem, there exist k disjoint (possibly empty) chains $C_1,\ldots,C_k$ which partition P. Now, for each $j\in[k]$, write $C_j=\{(i_1,w(i_1)),\ldots,(i_{m_j},w(i_{m_j}))\}$, with $i_1<\cdots<i_{m_j}$. Since C_j is a chain in P, $(w(i_1),\ldots,w(i_{m_j}))$ is an increasing subsequence of w. Define $I_j=\{i_1,\ldots i_{m_j}\}$ and $J_j=\{w(i_1),\ldots w(i_{m_j})\}$. It is now easy to check that $w=min(I_1,J_1,\ldots,I_k,J_k)$ and we are done.

$\square$

The numbers $|S_{n,k}|$ were studied by Gessel [**7**] who found an expression involving Bessel functions for the generating function $\sum_{n\geq1}|S_{n,k}|t^n$. The authors do not know of a simple form of the polynomial $\sum_{k=1}^{n}|S_{n,k}|t^k$.

Corollary 2.5. *Suppose that $I_1\uplus I_2=J_1\uplus J_2=[n]$, $|I_1|=|J_1|=n_1$, $|I_2|=|J_2|=n_2$, $w_1\in S_{n_1,k_1}$, and $w_2\in S_{n_2,k_2}$. For $i=1,2$ let x_i be the submatrix of x with row set I_i and column set J_i. Then, there exist $d_v\in\mathbb{C}$ such that $\mathrm{Imm}_{w_1}(x_1)\mathrm{Imm}_{w_2}(x_2)=\sum_v d_v\mathrm{Imm}_v(x)$, where the sum is over v in S_{n,k_1+k_2}.*

Specializing at $w_1=w_2=1$, we have that the coefficients d_v in the Corollary are in fact nonnegative real numbers. (see [**15**], [**14**]) Again, one may use the properties of the dual canonical basis of $\mathcal{O}SL_n\mathbb{C}$ to show that $\{d_w\,|\,w\in S_n\}$ are nonnegative real numbers.

Proof. For $i=1,2$, by Propositions 2.3 and 2.4 there exist $p_{i,j}(x_i)\in\Gamma_{n_i,k_i}$ and $d_j\in\mathbb{C}$ such that $\mathrm{Imm}_{w_i}(x)=\sum_j d_jp_{i,j}(x_i)$. Since x_1 and x_2 are complementary submatrices of x, for any $p_1(x_1)\in\Gamma_{n_1,k_1}$ and $p_2(x_2)\in\Gamma_{n_2,k_2}$, the product $p_1(x_1)p_2(x_2)$ is contained in Γ_{n,k_1+k_2}. So, the product $\mathrm{Imm}_{w_1}(x_1)\mathrm{Imm}_{w_2}(x_2)$ is a linear combination of elements in Γ_{n,k_1+k_2}. The result now follows from Proposition 2.3.

$\square$

Taken together, Propositions 2.3 and 2.4 imply that for $w\in S_{n,k}$, there exist $p_i(x)\in\Gamma_{n,k}$ and $d_i\in\mathbb{C}$ such that $\mathrm{Imm}_w(x)=\sum_{i=1}^{m}d_ip_i(x)$. Results in [**15**] and [**14**] show that, for $k=2$, we may in fact assume that the $p_i(x)$ are contained in a subset of $\Gamma_{n,2}$ which is in a natural bijective correspondence with the set of Dyck paths of length $2n$. It would be interesting to see if an analogous result holds for general k.

We now investigate when polynomials in span($\Gamma_{n,k}$) are TNN or SNN. For any integers n and k satisfying $1 \leq k \leq n$, define the poset $P_{n,k}$ on $\Gamma_{n,k}$ by
$\Delta_{I_1,J_1}(x) \cdots \Delta_{I_k,J_k}(x) \leq \Delta_{I'_1,J'_1}(x) \cdots \Delta_{I'_k,J'_k}(x)$ if and only if the difference
$\Delta_{I'_1,J'_1}(x) \cdots \Delta_{I'_k,J'_k}(x) - \Delta_{I_1,J_1}(x) \cdots \Delta_{I_k,J_k}(x)$ is TNN. In [**15**] the authors develop necessary and sufficient combinatorial conditions for polynomials $p(x) \in$ span($\Gamma_{n,2}$) to be TNN. For all positive integers n, $P_{n,2}$ has a unique maximal element given by $\Delta_{I,I}(x)\Delta_{J,J}(x)$, where $I = \{1, 3, 5, \ldots\}$ and $J = \{2, 4, 6, \ldots\}$. Also, the determinant $\Delta_{[n],[n]}(x)\Delta_{\emptyset,\emptyset}(x)$ is always a minimal element of $P_{n,2}$. In [**14**] the authors show that the combinatorial tests in [**15**] constitute sufficient conditions for polynomial in span($\Gamma_{n,2}$) to be SNN. Therefore, whenever $\Delta_{I,J}(x)\Delta_{I',J'}(x) \leq \Delta_{K,L}(x)\Delta_{K',L'}(x)$ in $P_{n,2}$ we also have that $\Delta_{K,L}(x)\Delta_{K',L'}(x) - \Delta_{I,J}(x)\Delta_{I',J'}(x)$ is SNN. It is unknown whether the converse of the last sentence is true.

In [**4**], [**5**], and [**14**] the positivity properties of differences of the form $x_{1,w(1)} \cdots x_{n,w(n)} - x_{1,u(1)} \cdots x_{n,u(n)}$ for $w, u \in S_n$ are studied. The authors prove the following about the subposet $P_{n,n}$ consisting of products of n nonempty minors.

THEOREM 2.6. *Let $w, u \in S_n$. Then, the following statements are equivalent.*
1. $w \leq u$ in the Bruhat order.
2. The difference $x_{1,w(1)} \cdots x_{n,w(n)} - x_{1,u(1)} \cdots x_{n,u(n)}$ is TNN.
3. The difference $x_{1,w(1)} \cdots x_{n,w(n)} - x_{1,u(1)} \cdots x_{n,u(n)}$ is SNN.
4. Whenever the difference $x_{1,w(1)} \cdots x_{n,w(n)} - x_{1,u(1)} \cdots x_{n,u(n)}$ is applied to a Jacobi-Trudi matrix, the result is a nonnegative linear combination of monomial symmetric functions.
5. The difference $x_{1,w(1)} \cdots x_{n,w(n)} - x_{1,u(1)} \cdots x_{n,u(n)}$ is a nonnegative linear combination of Kazhdan-Lusztig immanants.

With the above results as motivation, we show that $P_{n,k}$ has a unique maximal element for arbitrary k and that certain comparable elements in $P_{n,k}$ have differences which are SNN as well as TNN.

LEMMA 2.7. *Let $(I_1, \ldots, I_p)$ and $(I'_1, \ldots, I'_p)$ be seqences of sets satisfying*
$I_1 \uplus \cdots \uplus I_p = I'_1 \uplus \cdots \uplus I'_p,$
$|I_i| = |I'_i| \qquad$ *for all i.*
Fix indices $k < \ell$ and define increasing sequences $(\alpha_1, \ldots, \alpha_p)$ and $(\alpha'_1, \ldots, \alpha'_p)$ by

$$I_k \cup I_\ell = \{\alpha_1, \ldots, \alpha_p\},$$
$$I'_k \cup I'_\ell = \{\alpha'_1, \ldots, \alpha'_p\}.$$

Define the sequences of sets $(J_1, \ldots, J_p)$ and $(J'_1, \ldots, J'_p)$ by

$$J_i = \begin{cases} \{\alpha_1, \alpha_3, \ldots, \} & \text{if } i = k, \\ \{\alpha_2, \alpha_4, \ldots, \} & \text{if } i = \ell, \\ I_i & \text{otherwise.} \end{cases}$$

$$J'_i = \begin{cases} \{\alpha'_1, \alpha'_3, \ldots, \} & \text{if } i = k, \\ \{\alpha'_2, \alpha'_4, \ldots, \} & \text{if } i = \ell, \\ I'_i & \text{otherwise.} \end{cases}$$

Then the immanant

$$\Delta_{J_1,J'_1}(x) \cdots \Delta_{J_p,J'_p}(x) - \Delta_{I_1,I'_1}(x) \cdots \Delta_{I_p,I'_p}(x)$$

is totally nonnegative and Schur nonnegative.

PROOF. This difference is

$$\frac{\Delta_{J_1,J'_1}(x) \cdots \Delta_{J_p,J'_p}(x)}{\Delta_{J_k,J'_k}(x)\Delta_{J_\ell,J'_\ell}(x)}(\Delta_{J_k,J'_k}(x)\Delta_{J_\ell,J'_\ell}(x) - \Delta_{I_k,I'_k}(x)\Delta_{I_\ell,I'_\ell}(x)),$$

which is totally nonnegative and Schur nonnegative by [**15**, Prop. 4.6] and [**14**, Thm. 5.2]. ☐

Our next result implies that the poset $P_{n,k}$ has a maximal element for any n and k.

THEOREM 2.8. *Let* $(I_1, \dots, I_p)$ *and* $(I'_1, \dots, I'_p)$ *be two sequences of sets satisfying*

$$I_1 \uplus \cdots \uplus I_p = I'_1 \uplus \cdots \uplus I'_p = [n],$$
$$|I_1| = |I'_i| \qquad \text{for all } i$$

and define sets $J_1, \dots, J_p$ *by*

$$J_i = \{i \in [n] \,|\, i \equiv j \mod p\}.$$

Then the immanant

$$\Delta_{J_1,J_1}(x) \cdots \Delta_{J_p,J_p}(x) - \Delta_{I_1,I'_1}(x) \cdots \Delta_{I_p,I'_p}(x)$$

is totally nonnegative and Schur nonnegative.

PROOF. Applying several iterations of Lemma 2.7 to the sets $I_1, \dots, I_p, I'_1, \dots I'_p$, we obtain the desired result. $\square$

COROLLARY 2.9. *Let* $k < \ell$ *and define the sequences of sets* $(I_1, \dots, I_k)$ *and* $(J_1, \dots, J_\ell)$ *by*

$$I_j = \{i \in [n] \,|\, i \equiv j \mod k\},$$
$$J_j = \{i \in [n] \,|\, i \equiv j \mod \ell\}.$$

Then the immanant

$$\Delta_{J_1,J_1}(x) \cdots \Delta_{J_p,J_p}(x) - \Delta_{I_1,I'_1}(x) \cdots \Delta_{I_p,I'_p}(x)$$

is totally nonnegative and Schur nonnegative.

Not much is known about the posets $P_{n,k}$ in general. Obviously we have that $P_{n,1} \subset P_{n,2} \subset \cdots \subset P_{n,n}$. By Theorem 2.6 $P_{n,n}$ contains a subposet isomorphic to (the dual of) the Bruhat order on S_n. Also, it is possible to show that any element of span($\Gamma_{3,3}$) is TNN or SNN if and only if it may be expressed as a nonnegative linear combination of Kazhdan-Lusztig immanants. In particular, this allows one to construct the poset $P_{3,3}$ and see that it coincides with the analogous poset constructed by considering SNN differences. It would be interesting to see what $P_{n,k}$ looks like in general.

3. Acknowledgements

The authors are grateful to Mark Haiman and Peter Trapa for helpful conversations.

References

[1] Dan Barbasch and David Vogan. Primitive ideals and orbital integrals in complex exceptional groups. *J. Algebra*, 80(2):350–382, 1983.

[2] Francesco Brenti. Combinatorics and total positivity. *J. Combin. Theory Ser. A*, 71(2):175–218, 1995.

[3] C. W. Cryer. Some properties of totally positive matrices. *Linear Algebra Appl.*, 15:1–25, 1976.

[4] B. Drake, S. Gerrish, and M. Skandera. Two new criteria for comparison in the Bruhat order. *Electron. J. Combin.*, 11(1), 2004. Note 6, 4 pp. (electronic).

[5] B. Drake, S. Gerrish, and M. Skandera. Monomial nonnegativity and the Bruhat order. Submitted, 2005.

[6] Jie Du. Canonical bases for irreducible representations of quantum GL_n. *Bull. London Math. Soc.*, 24(4):325–334, 1992.

[7] I. Gessel. Symmetric functions and p-recursiveness. *J. Combin. Theory Ser. A*, 58:257–285, 1990.

[8] M. Haiman. Hecke algebra characters and immanant conjectures. *J. Amer. Math. Soc.*, 6(3):569–595, 1993.

[9] S. Karlin and G. McGregor. Coincidence probabilities. *Pacific J. Math.*, 9:1141–1164, 1959.

[10] D. Kazhdan and G. Lusztig. Representations of Coxeter groups and Hecke algebras. *Inv. Math.*, 53:165–184, 1979.

[11] Thomas Lam, Alexander Postnikov, and Pavlo Pylyavskyy. Schur positivity conjectures: $2\frac{1}{2}$ are no more! Preprint math.CO/0502446 on ArXiv, 2005.

[12] B. Lindström. On the vector representations of induced matroids. *Bull. London Math. Soc.*, 5:85–90, 1973.

[13] C. Loewner. On totally positive matrices. *Math. Z.*, 63:338–340, 1955.

[14] B. Rhoades and M. Skandera. Kazhdan-Lusztig immanants and products of matrix minors. Submitted, 2005.

[15] B. Rhoades and M. Skandera. Temperley-Lieb immanants. *Ann. Comb.*, 2005. To appear.

[16] M. Skandera. On the nonnegativity properties of the dual canonical basis. 2005. In progress.

[17] R. Stanley. *Enumerative Combinatorics*, volume 1. Cambridge University Press, Cambridge, 1997.

[18] R. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, Cambridge, 1999.

[19] A. Whitney. A reduction theorem for totally positive matrices. *J. Anal. Math.*, 2:88–92, 1952.

DEPT. OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN, 55455

DEPARTMENT OF MATHEMATICS, HAVERFORD COLLEGE, 370 LANCASTER AVENUE, HAVERFORD, PA, 19041

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Some Expansions of the Dual Basis of Z_λ

Amanda Riehl

Abstract.

A zigzag or ribbon is a connected skew diagram that contains no 2×2 boxes. Given a composition $\beta = (\beta_1, \ldots \beta_k)$, we let Z_β denote the skew Schur function corresponding to the zigzag shape whose row lengths are $\beta_1, \ldots \beta_k$ reading from top to bottom. For each n, the set $\{Z_\lambda\}_{\lambda \vdash n}$ is a basis for Λ_n, the space of homogeneous symmetric functions of degree n. In this paper, we investigate some characteristics of the dual basis of $\{Z_\lambda\}_{\lambda \vdash n}$ relative to the Hall inner product which we denote by $\{DZ_\lambda\}_{\lambda \vdash n}$. We give a combinatorial interpretation for the coefficients in the expansion of DZ_λ in terms of the monomial symmetric functions $\{m_\mu\}_{\mu \vdash n}$ as a certain signed sum of paths in the partition lattice under refinement. We shall show that in many cases, we can give an explicit formulas for the coefficients $a_{\mu,\lambda} = DZ_\lambda |_{m_\mu}$. In addition, we give explicit formulas for the coefficients that arise in the expansion of DZ_λ in terms of Schur functions for several special cases. As an application, we obtain combinatorial interpretations for the coefficients in the expansion of Schur functions and general ribbon Schur functions in terms of ribbon Schur functions indexed by partitions.

Résumé. Un zigzag ou un ruban est un diagramme connexe oblique qui ne contient aucune boite 2×2. Soit une composition $\beta = (\beta_1, \ldots \beta_k)$, notons Z_β la fonction oblique de Schur correspondant la forme de zigzag dont les longueurs des lignes sont $\beta_1, \ldots \beta_k$ lu de haut en bas. Pour chaque n, l'ensemble $\{Z_\lambda\}_{\lambda \vdash n}$ est une base Λ_n, de l'éspace des fonctions symétriques homogène de degré n. Dans cet article, nous étudions certaines caractéristiques de la base duale de $\{Z_\lambda\}_{\lambda \vdash n}$ relativement au produit intérieur de Hall que nous dénotons par $\{DZ_\lambda\}_{\lambda \vdash n}$. Nous donnons une interprétation combinatoire des coefficients dans l'expansion de DZ_λ en termes des fonctions symétriques monômiales $\{m_\mu\}_{\mu \vdash n}$ comme une somme signée de chemin dans le treillis des partages (l'ordre est le raffinement). Nous montrerons que, dans beaucoup de cas, nous pouvons donner des formules explicites pour les coefficients $a_{\mu,\lambda} = DZ_\lambda |_{m_\mu}$. De plus, nous donnons dans plusieurs cas des formules explicites pour les coefficients dans l'expansion de DZ_λ en termes de fonctions de Schur. Comme application, nous obtenons des interprétations combinatoires pour les coefficients dans l'expansion des fonctions de Schur et des fonctions Schur de ruban en termes de fonctions de Schur ruban indexées par les partages.

1. Introduction

Zigzag (or ribbon) Schur functions are the skew Schur functions with a ribbon shape and indexed by compositions. A composition $\beta = (\beta_1, \ldots, \beta_k)$ of n, denoted $\beta \models n$, is a sequence of positive integers such that $\beta_1 + \beta_2 + \ldots + \beta_k = n$. We define a zigzag shape to be a connected skew shape that contains no 2 x 2 array of boxes. Given a composition $\beta = (\beta_1, \ldots \beta_k)$, we let Z_β denote the skew Schur function corresponding to the zigzag shape whose row lengths are $\beta_1, \ldots \beta_k$ reading from top to bottom. For example Figure 1 shows the zigzag shape corresponding to the composition $(2, 3, 1, 4)$. As pointed out in [**2**], zigzag Schur functions arise in many contexts. For example, the scalar product of any two zigzags gives the number of permutations σ such that σ and σ^{-1} have the associated pair of descent sets [**9**]. Zigzags can also be used to compute the number of permutations with a given descent set and cycle structure [**5**]. MacMahon [**8**] showed their coefficients in terms of the monomial symmetric functions count descents in permutations with repeated

2000 *Mathematics Subject Classification.* Primary 05E05; Secondary 05A17.

Key words and phrases. symmetric functions, zigzag Schur functions, dual basis.

This research was done under the supervision of Jeffrey Remmel as part of the author's thesis, and the author would like to thank him for his support and numerous helpful conversations.

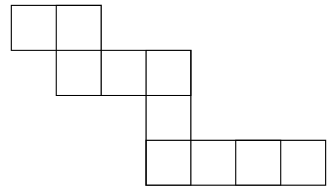


Figure 1. The ribbon shape corresponding to the composition $(2,3,1,4)$, so that $s_{(7,4,4,2)/(3,3,1)} = Z_{(2,3,1,4)}$.

elements. They also show up as sl_n-characters of the irreducible components of the Yangian representation in level 1 modules of $\hat{sl}_n$[**6**].

The zigzag Schur functions corresponding to partitions of n form a basis of Λ_n, the space of homogeneous symmetric functions of degree n, and therefore they have a dual basis relative to the Hall inner product which we denote by $\{DZ_\lambda\}_{\lambda\vdash n}$. We shall call DZ_λ the dual zigzag symmetric function corresponding to λ. The basis $\{DZ_\lambda\}_{\lambda\vdash n}$ has not been extensively studied. Let $\{m_\lambda\}_{\lambda\vdash n}$ denote the set of monomial symmetric functions, $\{h_\lambda\}_{\lambda\vdash n}$ denote the set of homogeneous symmetric functions, and $\{s_\lambda\}_{\lambda\vdash n}$ denote the set of Schur functions. The main result of this paper is to give a combinatorial interpretation to coefficients that arise in the expansion of DZ_λ in terms of the monomial symmetric functions. That is, we shall give a combinatorial interpretation to $a_{\mu,\lambda}$ where

$$DZ_\lambda = \sum_\mu a_{\mu,\lambda} m_\mu. \tag{1.1}$$

Our main result will show that $a_{\mu,\lambda}$ is a signed sum over the weights of certain paths in the lattice of partitions under refinement. In general such a signed sum is complicated, but we will show that in many special cases, we can explicitly evaluate this sum. For example, we will show that $a_{\mu,(n)} = 1$ for all μ so that

$$DZ_{(n)} = \sum_\mu m_\mu = s_{(n)}$$

where $s_{(n)}$ is the Schur function associated to the partition with only one part.

Once we have found our combinatorial interpretation for $a_{\mu,\lambda}$, we can obtain combinatorial interpretations for the expansion of DZ_λ in terms of any other basis by using the combinatorial interpretations of the transition matrices between bases of symmetric functions found in [**1**]. In particular, we shall use this method to find explicit values for $b_{\mu,\lambda}$ where

$$DZ_\lambda = \sum_\mu b_{\mu,\lambda} s_\mu \tag{1.2}$$

for certain special cases.

We now give brief explanations of the concepts to state our main result. There is a natural correspondence between a composition β of n and subsets of $[n-1]$. That is, given a composition $\beta = (\beta_1, \ldots, \beta_k)$ of n, we define a subset of $[n-1]$ by

$$Set(\beta) = \{\beta_1, \beta_1+\beta_2, \beta_1+\beta_2+\beta_3, \ldots, \beta_1+\beta_2+\ldots+\beta_{k-1}\}. \tag{1.3}$$

We also let $\lambda(\beta)$ denote the partition that arises from β by arranging its parts in decreasing order and $\ell(\beta)$ denote the number of parts of β. For example, if $\beta = (2,3,1,2)$, then $Set(\beta) = \{2,5,6\}$ and $\lambda(\beta) = (3,2,2,1)$. Given a subset $S = \{a_1 < a_2 < \cdots < a_r\} \subseteq [n-1]$, we define a composition of n by

$$\beta_n(S) = (a_1, a_2 - a_1, \ldots, a_r - a_{r-1}, n - a_r). \tag{1.4}$$

For example, if $S = \{2,4,8\}$, then $\beta_{10}(S) = (2,2,4,2)$. We also define $shape_n(S) = \lambda(\beta_n(S))$. Given two compositions β and γ, we say that β is a refinement of γ, denoted $\beta \leq_r \gamma$, if by adding together adjacent components of β, we can obtain γ. For two partitions μ and λ with $\mu \leq_r \lambda$, we define $Path(\mu,\lambda)$ to be the set of all $P = (\mu_0, \mu_1, \ldots, \mu_k)$, such that $\mu = \mu_0 <_r \mu_1 <_r \ldots <_r \mu_k = \lambda$. If $P = (\mu_0, \mu_1, \ldots, \mu_k)$ is such a path, we let $\ell(P) = k$ denote the length of P. Finally, μ and λ are partitions of n, then we define

$$[\mu \to \lambda] = |\{S \subseteq Set(\mu) : shape_n(S) = \lambda\}|$$

For example, if $\mu = (2,2,2,1)$ and $\lambda = (4,2,1)$, then $[\mu \to \lambda] = 2$, since $Set(\mu) = \{2,4,6\}$ and $\lambda(\beta_7(\{2,6\})) = \lambda(\beta_7(\{4,6\})) = (4,2,1)$.

This given, our main result is to give a combinatorial interpretation of for the coefficients $a_{\mu,\lambda}$ that arise in (1.1).

Theorem 1.1. *If λ and μ are partitions of n, then*

$$a_{\mu,\lambda} = (-1)^{l(\mu)-l(\lambda)} \sum_{P \in Path(\mu,\lambda)} [P](-1)^{l(P)}$$

where $P = (\mu_0, \mu_1, \ldots, \mu_k)$, $\mu = \mu_0 <_r \mu_1 <_r \ldots <_r \mu_k = \lambda$ and $[P] = [\mu_0 \to \mu_1][\mu_1 \to \mu_2] \ldots [\mu_{k-1} \to \mu_k]$.

As one application of our main result, we can give a combinatorial interpretation of the expansion of Z_α in terms of Z_λ's, where α is a composition of n, and λ is a partition of n. It is known, see [**4**], that

$$Z_\alpha = \sum_{T \subseteq Set(\alpha)} (-1)^{|Set(\alpha)-T|} h_{\lambda(\beta(T))}.$$

Thus if $Z_\alpha = \sum_{\mu \vdash n} f_{\mu,\alpha} Z_\mu$, then

$$f_{\mu,\alpha} = \langle Z_\alpha, DZ_\mu \rangle = \sum_{T \subseteq Set(\alpha)} (-1)^{|Set(\alpha)-T|} a_{\lambda(\beta(T)),\mu}. \tag{1.5}$$

In principle, (1.5) gives rise to a combinatorial algorithm to compute the coefficients $f_{\mu,\alpha}$. However, such an algorithm is not necessarily the most efficient way to compute these coefficients.

The outline of this paper is as follows. In Section 2, we shall review the necessary background for symmetric functions and the combinatorial interpretation of the entries of the transition matrices between various bases of symmetric functions that we shall need. In particular, we shall use the Jacobi-Trudi identity to give a combinatorial interpretation of the coefficients $Z_\lambda |_{h_\mu}$. In Section 3, we outline the proof of our main theorem and give some examples of the computations involved in computing the coefficients $a_{\mu,\lambda}$. In Section 4, we give closed forms for several of the coefficients, independent of the size of the composition. In Section 5, we give the expansion of several dual zigzags in terms of Schur functions which are independent of the size of the partition. In Section 6, we give a brief explanation of two applications of our main result.

2. Background Information

We say that $\lambda = (\lambda_1 \geq \cdots \geq \lambda_k)$ is a partition of n, written $\lambda \vdash n$ if $\lambda_1 + \ldots + \lambda_k = n = |\lambda|$. We $\ell(\lambda)$ denote the number of parts of λ. We let F_λ denote the Ferrers diagram of λ. If $\mu = (\mu_1, \ldots, \mu_m)$ is a partition where $m \leq k$ and $\lambda_i \geq \mu_i$ for all $i \leq m$, we let $F_{\lambda/\mu}$ denote the skew shape that results by removing the cells of F_μ from F_λ.

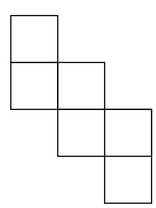


Figure 2. The skew Ferrers diagram of $(3,3,2,1)/(2,1)$.

A *column-strict tableau* T of shape λ is any filling of F_λ with natural numbers such that entries in each row are weakly increasing from left to right, and entries in each column are strictly increasing from bottom to top. We define the content of T to be $c(T) = (\alpha_1, \alpha_2, \ldots,)$ where α_i is the number of times that i occurs in T. If λ is a partition denoted by $\lambda = (\lambda_1, \ldots, \lambda_l) = (1^{m_1}, 2^{m_2}, \ldots, n^{m_n})$, where m_i is the number of parts of λ equal to i, then we define $z_\lambda = 1^{m_1} 2^{m_2} \cdots n^{m_n} m_1! m_2! \ldots m_n!$.

There are six standard bases of the space of homogeneous symmetric functions of degree n, $\Lambda_n(x)$, which are generally notated as: $\{m_\lambda\}_{\lambda \vdash n}$ (the monomial symmetric functions), $\{h_\lambda\}_{\lambda \vdash n}$ (the complete homogeneous symmetric function), $\{e_\lambda\}_{\lambda \vdash n}$ (the elementary symmetric functions), $\{p_\lambda\}_{\lambda \dashv n}$ (the power sum

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

FIGURE 3. A column strict tableau of shape $(3, 2, 2, 1)$ and content $(2, 2, 1, 2, 1)$.

symmetric functions), $\{f_\lambda\}_{\lambda \vdash n}$ (the forgotten symmetric functions) and $\{s_\lambda\}_{\lambda \vdash n}$ (the Schur functions), where λ is a partition of n.

The Hall inner product is a standard scalar product on the space of homogeneous symmetric functions $\Lambda_n(x)$, which is defined by:

$$\langle m_\lambda, h_\mu \rangle = \delta_{\lambda,\mu}$$

where

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

Under this scalar product, $\{s_\lambda\}_{\lambda \vdash n}$ and $\{p_\lambda/\sqrt{z_\lambda}\}_{\lambda \vdash n}$ are known to be self-dual, and $\{e_\lambda\}_{\lambda \vdash n}$ and $\{f_\lambda\}_{\lambda \vdash n}$ are dual [**1**].

When given two bases of $\Lambda_n(x)$, $\{a_\lambda\}_{\lambda \vdash n}$ and $\{b_\lambda\}_{\lambda \vdash n}$, we first fix some ordering of the partitions of n, e.g. the lexicographic order, and then we may think of the bases as row vectors, $\langle a_\lambda \rangle_{\lambda \vdash n}$ and $\langle b_\lambda \rangle_{\lambda \vdash n}$. We can define the transition matrix $M(a, b)$ that transforms the basis $\langle a_\lambda \rangle_{\lambda \vdash n}$ into the basis $\langle b_\lambda \rangle_{\lambda \vdash n}$ by

$$\langle b_\lambda \rangle_{\lambda \vdash n} = \langle a_\lambda \rangle_{\lambda \vdash n} M(a, b).$$

The (λ, μ) entry of $M(a, b)$ is given by the equation

$$b_\lambda = \sum_{\mu \vdash n} a_\mu M(a, b)_{\mu,\lambda}.$$

The main goal of this paper is to find a combinatorial interpretation of the entries of $M(m, DZ)$. That is, we want find a combinatorial interpretation for the $a_{\mu,\lambda}$ where

$$DZ_\lambda = \sum_\mu a_{\mu,\lambda} m_\mu.$$

In addition, we shall also be interested in finding a combinatorial interpretation for the entries of $M(s, DZ)$. That is, we want to find a combinatorial interpretation for $b_{\mu,\lambda}$ where

$$DZ_\lambda = \sum_\mu b_{\mu,\lambda} s_\mu.$$

We now give examples of the expansion of $\{DZ_\lambda\}_{\lambda\vdash n}$ when $n = 6$. We first give the expansion of DZ_λ in terms of the monomial symmetric functions, when $\lambda \vdash 6$.

$$\begin{aligned}
DZ_{(6)} &= m_6 + m_{5,1} + m_{4,2} + m_{4,1,1} + m_{3,3} + m_{3,2,1} + m_{3,1,1,1} \\
&\quad + m_{2,2,2} + m_{2,2,1,1} + m_{2,1,1,1,1} + m_{1,1,1,1,1,1} \\
DZ_{(5,1)} &= m_{5,1} + m_{4,1,1} + m_{3,2,1} + 2m_{3,1,1,1} + m_{2,2,1,1} + m_{2,1,1,1,1} - 2m_{1,1,1,1,1,1} \\
DZ_{(4,2)} &= m_{4,2} + m_{4,1,1} + 2m_{2,2,2} + m_{2,2,1,1} + 2m_{2,1,1,1,1} + 7m_{1,1,1,1,1,1} \\
DZ_{(4,1,1)} &= m_{4,1,1} + m_{3,1,1,1} + m_{2,2,1,1} + 3m_{2,1,1,1,1} + 8m_{1,1,1,1,1,1} \\
DZ_{(3,3)} &= m_{3,3} + m_{3,2,1} + m_{3,1,1,1} + m_{2,2,1,1} + m_{2,1,1,1,1} \\
DZ_{(3,2,1)} &= m_{3,2,1} + 2m_{3,1,1,1} + m_{2,2,1,1} + m_{2,1,1,1,1} - 3m_{1,1,1,1,1,1} \\
DZ_{(3,1,1,1)} &= m_{3,1,1,1} + m_{2,1,1,1,1} + m_{1,1,1,1,1,1} \\
DZ_{(2,2,2)} &= m_{2,2,2} + m_{2,2,1,1} + 2m_{2,1,1,1,1} + 5m_{1,1,1,1,1,1} \\
DZ_{(2,2,1,1)} &= m_{2,2,1,1} + 3m_{2,1,1,1,1} + 9m_{1,1,1,1,1,1} \\
DZ_{(2,1,1,1,1)} &= m_{2,1,1,1,1} + 5m_{1,1,1,1,1,1} \\
DZ_{(1,1,1,1,1,1)} &= m_{1,1,1,1,1,1}.
\end{aligned}$$

We note that we can get an indirect combinatorial interpretation of the coefficients $b_{\mu,\gamma}$ by using the combinatorial interpretation of the entries of the transition matrix $M(s,m)$ given in [**3**]. That is,

$$M(s,m)_{\lambda\mu} = K^{-1}_{\mu,\lambda},$$

where $||K^{-1}_{\mu,\lambda}||$ is the inverse Kostka matrix which will be described below. Thus

$$DZ_\lambda = \sum_{\mu \le_r \lambda} a_{\mu,\lambda} \sum_\gamma s_\gamma K^{-1}_{\mu,\gamma} = \sum_\gamma s_\gamma \sum_{\mu\le_r\lambda} a_{\mu,\lambda} K^{-1}_{\mu,\gamma}. \tag{2.1}$$

Hence

$$b_{\mu,\gamma} = \sum_{\mu\le_r\lambda} a_{\mu,\lambda} K^{-1}_{\mu,\gamma}. \tag{2.2}$$

The expansion of DZ_λ in terms of the Schur functions, when $\lambda \vdash 6$, is given below.

$$\begin{aligned}
DZ_{(6)} &= s_6 \\
DZ_{(5,1)} &= s_{5,1} - s_{4,2} + s_{3,2,1} - s_{2,2,2} - s_{2,2,1,1} \\
DZ_{(4,2)} &= s_{4,2} - s_{3,3} - s_{3,2,1} + 2s_{2,2,2} + s_{2,2,1,1} \\
DZ_{(4,1,1)} &= s_{4,1,1} - s_{3,2,1} + s_{2,2,2} + s_{2,2,1,1} \\
DZ_{(3,3)} &= s_{3,3} - s_{2,2,2} \\
DZ_{(3,2,1)} &= s_{3,2,1} - 2s_{2,2,2} - s_{2,2,1,1}.
\end{aligned}
\qquad
\begin{aligned}
DZ_{(3,1,1,1)} &= s_{3,1,1,1} - s_{2,2,1,1} \\
DZ_{(2,2,2)} &= s_{2,2,2} \\
DZ_{(2,2,1,1)} &= s_{2,2,1,1} \\
DZ_{(2,1,1,1,1)} &= s_{2,1,1,1,1} \\
DZ_{(1,1,1,1,1,1)} &= s_{1,1,1,1,1,1}
\end{aligned}$$

Next we shall describe the combinatorial interpretation of the coefficients that arise in expanding a skew Schur function in terms of the homogeneous symmetric functions. In particular, we will need to use the expansion of skew-Schur functions in terms of h_λ. To do so, we introduce rim hooks, special rim hooks and special rim hook tabloids. More detail is given in [**3**] where they are used to give a combinatorial interpretation of the inverse Kostka matrix.

For a partition λ, consider the Ferrers diagram F_λ. A rim hook of λ is a sequence of cells, h, along the northeast boundary of F_λ such that any two consecutive cells in h share an edge and if we remove h from F_λ, we are left with the Ferrers diagram of another partition. More generally, h is a rim hook of a skew shape λ/μ if h is a rim hook of λ which does not intersect μ.

A rim hook tableau of shape λ/ν and type μ, T, is a sequence of partitions

$$T = (\nu = \lambda^{(0)} \subset \lambda^{(1)} \subset \cdots \lambda^{(k)} = \lambda),$$

such that for each $1 \le i \le k, \lambda^{(i)}/\lambda^{(i-1)}$ is a rim hook of $\lambda^{(i)}$ of size μ_i. We define the sign of a rim hook $h_i = \lambda^{(i)}/\lambda^{(i-1)}$ to be

$$sgn(h_i) = (-1)^{r(h_i)-1},$$

where $r(h_i)$ is the number of rows that h_i occupies. The sign of a rim hook tableau T is

$$sgn(T) = \Pi_{i=1}^{k} sgn(h_i).$$

Given two partitions $\lambda^{(i-1)} \subset \lambda^{(i)}$, we say that $\lambda^{(i)}/\lambda^{(i-1)}$ is a special rim hook if $\lambda^{(i)}/\lambda^{(i-1)}$ is a rim hook of $\lambda^{(i)}$ and $\lambda^{(i)}/\lambda^{(i-1)}$ contains a cell from the first column of λ. A special rim hook tabloid (SRHT) T of shape λ/μ is a sequence of partitions

$$T = (\mu = \lambda^{(0)} \subset \lambda^{(1)} \subset \cdots \lambda^{(k)} = \lambda),$$

such that for each $1 \le i \le k, \lambda^{(i)}/\lambda^{(i-1)}$ is a special rim hook of $\lambda^{(i)}$. We have a partition determined by the integers $|\lambda^{(i)}/\lambda^{(i-1)}|$ which is the type of the special rim hook tabloid T. Notice that we have used the word *tabloid* instead of *tableau* in order to highlight there is no implicit order in the size of each successive special rim hook, unlike rim hook tableau.

The sign of a special rim hook, $h_i = \lambda^{(i)}/\lambda^{(i-1)}$, and the sign of a special rim hook tabloid T, are defined as we did for rim hooks and rim hook tableaux. We show an example of a special rim hook tabloid of type $(6, 5, 4, 2)$ and shape $(5, 4, 4, 3, 1)$ in Fig 4. For $|\lambda/\nu| = |\mu|$, Eğecioğlu and Remmel [**3**] show that

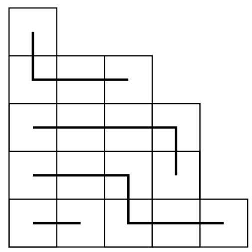


Figure 4. A special rim hook tabloid of shape (5,4,4,3,1) and type (6,5,4,2).

$$s_{\lambda/\nu} = \sum_{\mu} K^{-1}_{\mu,\lambda/\nu} h_\mu \tag{2.3}$$

where

$$K^{-1}_{\mu,\lambda/\nu} = \sum_{T \text{ is a SRHT of shape } \lambda/\nu \text{ and type } \mu} sgn(T).$$

Hence we obtain a combinatorial description of

$$M(s, m)_{\lambda,\mu} = K^{-1}_{\mu,\lambda}.$$

Recall that we defined a composition β of n, denoted $\beta \models n$, as a list of positive integers $(\beta_1, \beta_2, \ldots, \beta_k)$ such that $\beta_1 + \beta_2 + \ldots + \beta_k = n$. We call β_i a *component* of β, and we say that β has length $l(\beta) = k$ and size $|\beta| = n$. From this definition, we can see that β is a partition if each of its components are weakly decreasing. For any composition β, we define the partition determined by β, $\lambda(\beta)$, which we obtain by reordering the components of β in weakly decreasing order, e.g. $\lambda(2, 8, 9, 4) = (9, 8, 4, 2)$. Notice that two compositions β, γ can determine the same partition, e.g. if $\beta = (2, 8, 9, 4)$ and $\gamma = (2, 9, 8, 4)$, then $\lambda(2, 8, 9, 4) = (9, 8, 4, 2) = \lambda(2, 9, 8, 4)$.

There is a natural correspondence between a composition $\beta \models n$ and a subset $Set(\beta) \subseteq [n-1] = \{1, 2, \ldots, n-1\}$ where

$$Set(\beta) = \{\beta_1, \beta_1 + \beta_2, \beta_1 + \beta_2 + \beta_3, \ldots, \beta_1 + \beta_2 + \ldots + \beta_{k-1}\}.$$

We can also reverse this process so that for any subset $S = \{j_1, j_2, \ldots, j_{k-1}\} \subseteq [n-1]$, we can find the composition $\beta_n(S) \models n$ where

$$\beta_n(S) = (j_1, j_2 - j_1, \ldots, n - j_{k-1}).$$

For example, the composition $\beta = (2, 9, 8, 4)$ has $Set(\beta) = \{2, 11, 19\} \subseteq [22]$. We also define $shape_n(S) = \lambda(\beta_n(S))$. For example if $S = \{2, 5, 6, 10\}$ and $n = 11$, then $\beta_{11}(S) = (2, 3, 1, 4, 1)$, and $shape_{11}(S) = (4, 3, 2, 1, 1)$.

Given two partitions λ and μ of n, we say that λ is a refinement of μ, written $\lambda \le_r \mu$, if λ can be created from μ by splitting some of the parts of μ into pieces. For example, $(4, 2, 1, 1, 1, 1) \le_r (5, 3, 2)$ since we can split 5 into $4+1$ and 3 into $1+1+1$ to obtain λ. The cover relations in the lattice of partitions of n under refinement arise by starting with a partition λ and combining two of the parts of λ to get μ. Similarly, given two compositions β and γ, we say that β is a refinement of γ, denoted $\beta \leq_r \gamma$, if by adding together adjacent components of β, we can obtain γ. For example, $421131 \leq_r 4314$, meaning $\gamma = 421131$ is a refinement of $\beta = 4314$. If we only add together a single pair of adjacent components of a partition β to get γ, then we will say that γ covers β.

The refinement ordering restricted to the set of partitions forms a lattice which we call the lattice of partitions under refinement, or more briefly, the refinement lattice. For two partitions μ and λ, with $\mu \leq_r \lambda$ we define $Path(\mu, \lambda)$ to be the set of all $P = (\mu_0, \mu_1, \ldots, \mu_k)$, such that $\mu = \mu_0 <_r \mu_1 <_r \ldots <_r \mu_k = \lambda$. We define the length of P, $l(P) = k$.

Given two partitions of λ and μ of n such that $\mu \leq_r \lambda$, we define

$$[\mu \to \lambda] = |\{S \subseteq Set(\mu) : shape_n(S) = \lambda\}|.$$

As an example, let's calculate $[(2, 1^4) \to (4, 2)]$. Note that $Set(2, 1^4) = \{2, 3, 4, 5\}$. We want to find $|\{S \subseteq \{2, 3, 4, 5\} : shape_6(S) = (4, 2)|$. The only two subsets of $\{2, 3, 4, 5\}$ that have the appropriate shape are $\{2\}$ and $\{4\}$, so $[(2, 1^4) \to (4, 2)] = 2$.

3. A sketch of the proof of Theorem 1.1

Before proceeding with the proof of Theorem 1.1, we shall demonstrate how it can be used to calculate $a_{\mu,\lambda}$ in the case where $\mu = (1^6)$ and $\lambda = (3, 2, 1)$. Since our theorem says we sum over all paths in the refinement lattice, we give the relevant portion of the refinement lattice in Fig. 5. First we give several

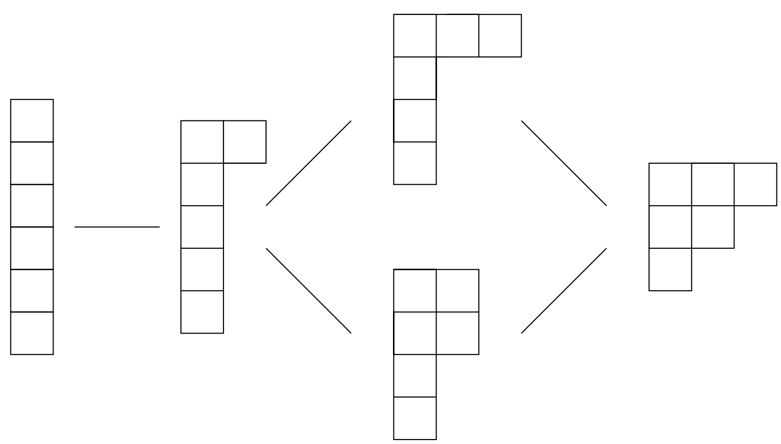


FIGURE 5. The refinement lattice from (1,1,1,1,1,1) to (3,2,1).

examples of how to calculate $[\alpha \to \beta]$. Recall that $Set(\lambda) = \{\lambda_1, \lambda_1 + \lambda_2, \ldots, \lambda_1 + \cdots + \lambda_{k-1}\}$. We first calculate $[(1^6) \to (2, 1^4)]$, which is equal to $|\{S \subset Set(1^6) : shape_6(S) = (2, 1^4)\}|$. $Set(1^6) = \{1, 2, 3, 4, 5\}$, and the subsets $\{2, 3, 4, 5\}, \{1, 3, 4, 5\}, \{1, 2, 4, 5\}, \{1, 2, 3, 5\}$, and $\{1, 2, 3, 4\}$ all have shape equal to $(2, 1^4)$. Therefore $[(1^6) \to (2, 1^4)] = 5$. Similarly $[(1^6) \to (3, 2, 1)] = 6$ since $\{3, 4\}, \{3, 5\}, \{2, 5\}, \{2, 3\}, \{1, 3\}, \{1, 4\}$ are the only subsets T of $Set(1^6) = \{1, 2, 3, 4, 5\}$ such that $shape_6(T) = (3, 2, 1)$. Finally we calculate $[(2, 1^4) \to (3, 1^3)]$. In this case, $Set(2, 1^4) = \{2, 3, 4, 5\}$ and the only subset T of $Set(2, 1^4)$ such that $shape_6(T) = (3, 1^3)$ is $\{3, 4, 5\}$. Thus $[(2, 1^4) \to (3, 1^3)] = 1$.

From these three examples we see that a considerable amount of work goes into calculating $[\alpha \to \beta]$ for every possibility in our refinement lattice. In Table 1, we give the values needed to calculate $[\alpha \to \beta]$ for all pairs in the refinement lattice from (1^6) to $(3, 2, 1)$.

Once we have calculated those values, we can easily calculate the weights of each possible path in our refinement lattice. These paths and weights are listed in Table 2. The length of the path will be used in our calculation of $a_{\mu,\lambda}$.

$[1^6 \to 2, 1^4] = 5$	$[2, 1^4 \to 3, 1^3] = 1$	$[3, 1^3 \to 3, 2, 1] = 2$
$[1^6 \to 3, 1^3] = 4$	$[2, 1^4 \to 2^2, 1^1] = 3$	$[2^2, 1^2 \to 3, 2, 1] = 1$
$[1^6 \to 2^2, 1^2] = 6$	$[2, 1^4 \to 3, 2, 1] = 4$	
$[1^6 \to 3, 2, 1] = 6$		

TABLE 1. Values for $[\alpha \to \beta]$ for pairs in the refinement lattice from (1^6) to $(3,2,1)$.

Possible Paths	Length of Path	Weight of Path
$[(1^6) \to (3,2,1)]$	1	6
$[(1^6) \to (3,1^3)][(3,1^3) \to (3,2,1)]$	2	8
$[(1^6) \to (2^2,1^2)][(2^2,1^2) \to (3,2,1)]$	2	6
$[(1^6) \to (2,1^4)][(2,1^4) \to (3,2,1)]$	2	20
$[(1^6) \to (2,1^4)][(2,1^4) \to (3,1^3)][(3,1^3 \to (3,2,1)]$	3	10
$[(1^6) \to (2,1^4)][(2,1^4) \to (2^2,1^2)][(2^2,1^2 \to (3,2,1)]$	3	15

TABLE 2. The weight of each possible path in the refinement lattice from (1^6) to $(3,2,1)$.

Finally, we combine this information:

$$\begin{aligned} a_{(1^6),(3,2,1)} &= (-1)^{6-3} \sum_{P \in Path((1^6),(3,2,1))} -1^{l(P)}[P] \\ &= -1^3(-1^1(6) + -1^2(8+6+20) + -1^3(10+15)) \\ &= -(-6+34-25) \\ &= -3. \end{aligned}$$

We should note that although this first example required many calculations, we have now done almost all of the work for several other coefficients for $n = 6$ since our the set of paths that we considered also arise in the computation of $a_{\alpha,\beta}$ for other pairs of partitions. In addition, we will see later that the same calculations allow us to evaluate an infinite number of coefficients $a_{\alpha,\beta}$ where α and β are partitions of $n > 6$.

Outline of proof of Theorem 1.1:

We start by expanding the zigzag Schur functions in terms of the homogeneous symmetric functions $\{h_\lambda\}_{\lambda \vdash n}$ derived from the Jacobi-Trudi by Egecioglu and Remmel [**3**],

$$s_{\lambda/\mu} = \det(h_{\lambda_i - \mu_j - i + j}) = \sum_\nu K^{-1}_{\nu,\lambda/\mu} h_\mu$$

where $h_0 = 1$ and $h_k = 0$ if $k < 0$. Applying it specifically to zigzag Schur functions and using compositions as subscripts, we can show that for any $\alpha \models n$,

$$Z_\alpha = (-1)^{l(\alpha)} \sum_{\beta \leq_r \alpha} (-1)^{l(\beta)} h_{\lambda(\beta)}.$$

Alternatively,

$$Z_\alpha = h_{\lambda(\beta(\alpha))} + \sum_{T \subset Set(\alpha)} (-1)^{|Set(\alpha) - T|} h_{\lambda(\beta(\alpha))}. \tag{3.1}$$

The result in 3.1 is well-known and can be proved by inclusion-exclusion [**4**]. Recall that $[\mu \to \lambda] = |\{S \subseteq Set(\mu) : shape_n(S) = \lambda\}|$. So

$$Z_\lambda = h_\lambda + \sum_{\lambda \leq_r \alpha} (-1)^{l(\lambda) - l(\alpha)} [\lambda \to \alpha] h_\alpha.$$

Since $\{Z_\lambda\}_{\lambda \vdash n}$ and $\{DZ_\lambda\}_{\lambda \vdash n}$ are dual bases, it follows that

$$\sum_\gamma Z_\gamma(x) DZ_\gamma(y) = \sum_\gamma h_\gamma(x) m_\gamma(y)$$

or, equivalently,

$$\sum_{\gamma} Z_\gamma(x) DZ_\gamma(y)|_{h_\lambda(x)m_\mu(y)} = \delta_{\lambda,\mu}.$$

Given our expansion of $Z_\lambda(x)$ in terms of $h_\lambda(x)$'s and the fact that $\langle h_\lambda(x), m_\mu(x)\rangle = \delta_{\lambda,\mu}$, we can then show that

$$\sum_{\gamma} Z_\gamma(x) DZ_\gamma(y)|_{h_\lambda(x)} = \sum_{\alpha \leq_r \lambda} (-1)^{l(\alpha)-l(\lambda)}[\alpha \to \lambda] m_\alpha(y)$$

and

$$\begin{aligned}
\sum_{\gamma} Z_\gamma(x) DZ_\gamma(y)|_{h_\lambda(x)m_\mu(y)} &= \sum_{\mu \leq_r \alpha \leq_r \lambda} (-1)^{l(\alpha)-l(\lambda)}[\alpha \to \lambda] a_{\mu,\alpha} \\
&= \sum_{\mu \leq_r \alpha \leq_r \lambda} \sum_{P \in Path(\mu,\alpha)} [P][\alpha \to \lambda] \\
&= \sum_{Q \in Paths(\mu,\lambda)} sgn(Q)[Q]
\end{aligned}$$

Thus we need only show that $\sum_{Q\in Path(\mu,\lambda)} sgn(Q)[Q] = \delta_{\lambda,\mu}$. This can be done by defining a weight preserving involution on the set of paths in the lattice of partitions under refinement but we do not have the space to give the argument in this paper.

4. Special Cases of the $a_{\mu,\lambda}$'s

We saw in our example calculating $a_{(1^6),(3,2,1)}$ how difficult and time-consuming it can be to find these coefficients. However, in a number of special cases, we can actually compute a closed form for the sum $a_{\mu,\lambda} = (-1)^{l(\mu)-l(\lambda)} \sum_{P\in Path(\mu,\lambda)} [P](-1)^{l(P)}$. For example, if $\mu <_r \lambda$ is a cover relation in the refinement lattice, then there is only one path and the formula for the coefficient $a_{\mu,\lambda}$ consists of a single term. In fact, we can prove the following.

1. If λ and μ are a cover relation in the refinement lattice, then $a_{\mu,\lambda} = [\mu \to \lambda]$.
2. Similarly, we can show that $a_{\mu,\mu} = 1$ for all μ.
3. For any μ such that $\mu \vdash n$, $a_{\mu,(n)} = 1$, so that we find $DZ_{(n)} = \sum_\mu m_\mu = s_{(n)}$.

We outline a proof of 3 by induction on the length of the refinement.

$$\begin{aligned}
a_{\mu,(n)} &= (-1)^{l(\mu)-1} \sum_{P\in Path(\mu,(n))} (-1)^{l(P)}[P] \\
&= (-1)^{l(\mu)-1} \sum_{\mu <_r \alpha <_r (n)} (-1)[\mu \to \alpha] \sum_{P \in Path(\alpha,(n))} (-1)^{l(P)}[P] \\
&\quad + (-1)^{l(\mu)-1}(-1)[\mu \to (n)]
\end{aligned}$$

Our inductive assumption that $a_{\alpha,(n)} = 1$ gives that $\sum_{P\in Path(\alpha,(n))} (-1)^{l(P)}[P] = (-1)^{l(\alpha)-1}$. Thus Note that

$$a_{\mu,(n)} = (-1)^{l(\mu)-1} \Big(\sum_{\mu <_r \alpha <_r (n)} (-1)[\mu \to \alpha](-1)^{l(\alpha)-1} \Big) + (-1)^{l(\mu)-1}(-1)[\mu \to (n)].$$

But if we think about the definition of $[\mu \to \alpha]$, now we are summing over all possibilities of ways to remove at least one element from $Set(\mu)$ so

$$\begin{aligned}
a_{\mu,(n)} &= (-1)^{l(\mu)-1} \sum_{\emptyset \subsetneq S \subseteq Set(\mu)} (-1)^{|Set(\mu)|-|S|} \\
&= (-1)^{l(\mu)-1} \Big(\Big(\sum_{\emptyset \subseteq S \subseteq Set(\mu)} (-1)^{|Set(\mu)|-|S|} \Big) - (-1)^{|Set(\mu)|} \Big)
\end{aligned}$$

But $\sum_{S\subseteq Set(\mu)} (-1)^{|S|} = 0$. So

$$a_{\mu,(n)} = (-1)^{l(\mu)}(0 - (-1)^{|Set(\mu)|}) = (-1)^{l(\mu)}((-1)^{|Set(\mu)|+1})$$

But $|Set(\mu)| + 1 = l(\mu)$, so $a_{\mu,(n)} = 1$.

Other results can be found using careful examination of the lattice of refinement. The proofs of some of the below items are very straightforward. For example, the proof of item 4 is plain because the relevant portion of the refinement lattice contains only two shapes. Moreover, $Set(1^k) = \{1, 2, \ldots, k-1\}$ and when we remove any element from $Set(1^k)$, one ends up with a set that has shape $(2, 1^{k-2})$. Since there are $k-1$ ways to remove one element from $Set(1^k)$, it follows that $a_{(1^k),(2,1^{k-2})} = k-1$. The proofs of other items are more involved.

Results with $\mu = (1^k)$ and $\lambda = (b, 1^{k-b})$ for $b = 1, 2, \ldots, 7$:

4. $a_{(1^k),(2,1^{k-2})} = k-1$
5. $a_{(1^k),(3,1^{k-3})} = 1$
6. $a_{(1^k),(4,1^{k-4})} = \binom{k-1}{2} - 2$
7. $a_{(1^k),(5,1^{k-5})} = -\frac{1}{2}(k-1)(k-4) + 3$
8. $a_{(1^k),(6,1^{k-6})} = \frac{1}{6}(k^3 - 3k^2 - 16k - 6)$
9. $a_{(1^k),(7,1^{k-7})} = -\frac{1}{3}(k)(k+1)(k-7) + 1$

Here are some other results which are useful for the computation of the coefficients $b_{\mu,\lambda}$ of (**??**):

10. $a_{(1^k),(3^2,1^{k-6})} = 0$
11. $a_{(1^k),(3,2,1^{k-5})} = -\frac{1}{2}k(k-5)$
12. $a_{(2,1^{k-2}),(4,1^{k-4})} = k-3$
13. $a_{(2,1^{k-2}),(3,2,1^{k-5})} = 1$

Theorem 4.1. *If $d \neq 1$,*

$$a_{(2^c,1^b),(2^{c+d},1^{b-2d})} = \frac{b(b-1)\cdots(b-d+2)}{d!}(b-2d+1)$$

Note that if $d = 1$, the product on the right is not defined, so that Theorem 4.1 would not make sense. However the case where $d = 1$ and $c = 0$ is a special case of one our previous formulas.

Finding the value of one coefficient also tells us the value of an infinite number of other coefficients. Let $\mu = (\mu_1, \ldots, \mu_j)$. That is, define $k\mu$ to be the partition obtained when each part of μ is multiplied by k so that $k\mu = (k\mu_1, \ldots, k\mu_j)$. Then we can prove the following result.

Theorem 4.2. *For all $k \in \mathbb{N}$,*

$$a_{\mu,\lambda} = a_{k\mu,k\lambda}.$$

In particular, if we apply Theorem 4.2 to Theorem 4.1, we obtain infinite number of cases where we have explicit formulas for $a_{\mu,\lambda}$. The proof of Theorem 4.2 follows from an obvious bijection between paths in the refinement lattice of (μ, λ) to paths in the refinement lattice of $(k\mu, k\lambda)$.

Here is another result of the same sort.

Theorem 4.3. *Let $\mu = (\mu_1, \ldots, \mu_s)$ and $\lambda = (\lambda_1, \ldots, \lambda_t)$. Then for any j such that $1 \leq j < min(\mu_s, \lambda_t)$,*

$$a_{\mu,\lambda} = a_{(\mu_1,\ldots,\mu_s,j),(\lambda_1,\ldots,\lambda_t,j)}.$$

The proof of Theorem 4.3 follows from examining the compositions and noticing that we must always have the last element of the composition in our subsets S in order for $shape_n(S)$ to match $(\lambda_1, \ldots, \lambda_t, k)$. This theorem works in "both directions", so to speak. Knowledge of the coefficients $a_{\mu,\lambda}$ where $\mu \vdash n$ and $\lambda \vdash n$ both with smallest part larger than 1 allows us to compute values of $a_{\alpha,\beta}$ for certain partitions α and β of size larger than n. Conversely, knowledge of coefficients $a_{\mu,\lambda}$ where μ and λ have identical unique smallest part allows us to compute values of $a_{\alpha,\beta}$ where α and β are partitions of size smaller than n by removing that smallest part from both μ and λ.

Thus the combination of Theorem 4.2 and Theorem 4.3 enables us to calculate the value $a_{\alpha,\beta}$ for infinitely many α and β starting with a single value of $a_{\mu,\lambda}$. That is, starting with $a_{\mu,\lambda}$, we can first multiply each part by k, then add smaller parts on the end, and so on.

5. Special Cases of the $b_{\mu,\lambda}$'s

Our method of expansion in terms of Schur functions in section 2 is useful not only in calculating particular expansions, but can also be used to make general statements independent of the size of λ.

T	$(-1)^{\lvert Set(\alpha)-T\rvert}$	$\lambda(\beta(T))$
$\emptyset$	-1	(10)
$\{2\}$	1	(8,2)
$\{4\}$	1	(6,4)
$\{8\}$	1	(8,2)
$\{2,4\}$	-1	(6,2,2)
$\{2,8\}$	-1	(6,2,2)
$\{4,8\}$	-1	(4,4,2)
$\{2,4,8\}$	1	(4,2,2,2)

TABLE 3. Values for $(-1)^{\lvert Set(\alpha)-T\rvert}$ and $\lambda(\beta(T))$ for each possible $T \subseteq Set(2,2,4,2)$.

μ	$(4,2,2,2)$	$(4,4,2)$	$(6,2,2)$	$(6,4)$	$(8,2)$	(10)
$a_{(4,2,2,2),\mu}$	1	2	1	1	2	1
$-a_{(4,4,2),\mu}$	0	-1	0	-1	-1	-1
$-2a_{(6,2,2),\mu}$	0	0	-2	-2	-2	-2
$a_{(6,4),\mu}$	0	0	0	1	0	1
$2a_{(8,2),\mu}$	0	0	0	0	2	2
$-a_{(10),\mu}$	0	0	0	0	0	-1
Sum for each μ	1	1	-1	-1	1	0

TABLE 4. Values for $a_{\gamma,\mu}$ used to compute $Z_{(2,2,4,2)} = \sum_{\mu\vdash n} f_{\mu,(2,2,4,2)} Z_\mu$.

We can use the fact that $b_{\mu,\lambda}$ can be expressed as $a_{\mu,\lambda}$ to prove further results, in particular that

1. $DZ_{(n)} = s_{(n)}$
2. $DZ_{(1^n)} = s_{1^n}$
3. $DZ_{(2^k,1^{n-2k})} = s_{(2^k,1^{n-2k})}\ \forall\ k$
4. $DZ_{(3^k,1^{n-3k})} = s_{(3,1^{n-3})} - s_{(2^2,1^{n-4})}\ \forall\ k$
5. $DZ_{(3,2,1^{n-5})} = s_{(3,2,1^{n-5})} - 2s_{(2^3,1^{n-6})} - s_{(2^2,1^{n-4})}$
6. $DZ_{(4,1^{n-4})} = s_{(4,1^{n-4})} - s_{(3,2,1^{n-5})} + s_{(2^2,1^{n-4})} + s_{(2^3,1^{n-6})}$

The proof of 1 was given above. The proofs of the others involve using the combinatorial interpretation of the coefficients that arise in (2.1) and defining some appropriate involutions to simplify the sum.

6. Applications of Our Main Result

As noted in the introduction, one application of our main result is to give a combinatorial interpretation of the expansion of Z_α in terms of Z_λ's, where α is a composition of n and λ is a partition of n. We noted that if $Z_\alpha = \sum_{\mu\vdash n} f_{\mu,\alpha} Z_\mu$, then

$$f_{\mu,\alpha} = \langle Z_\alpha, DZ_\mu\rangle = \sum_{T\subseteq Set(\alpha)} (-1)^{\lvert Set(\alpha)-T\rvert} a_{\lambda(\beta(T)),\mu}.$$

We now present an example of this fact; we will expand $Z_{(2,2,4,2)}$ as a sum of Z_λ's indexed by partitions of 10.

Table 3 tells us that

$$f_{\mu,(2,2,4,2)} = a_{(4,2,2,2),\mu} - a_{(4,4,2),\mu} - 2a_{(6,2,2),\mu} + a_{(6,4),\mu} + 2a_{(8,2),\mu} - a_{(10),\mu}.$$

Then Table 4 gives that $Z_{(2,2,4,2)} = Z_{(4,2,2,2)} + Z_{(4,4,2)} - Z_{(6,2,2)} - Z_{(6,4)} + Z_{(8,2)}$.

As another application of our results is that we can give a combinatorial interpretation of the coefficients that arise in the expansion of a Schur function s_γ in terms of the Z_λ's where $,\gamma,\lambda \vdash n$. That is, we can give a combinatorial interpretation of $e_{\mu,\gamma}$ where $s_\gamma = \sum_{\mu\vdash n} e_{\mu,\gamma} Z_\mu$.

Note that by 2.3, $s_\gamma = \sum_\mu K^{-1}_{\mu,\gamma} h_\mu$, so that

λ	$(3,2,1)$	$(4,1,1)$	$(3,3)$	$(4,2)$	$(5,1)$	(6)
$a_{(3,2,1),\lambda}$	1	0	1	0	1	1
$-a_{(4,1,1),\lambda}$	0	-1	0	-1	-1	-1
$-a_{(3,3),\lambda}$	0	0	-1	0	0	-1
$a_{(5,1),\lambda}$	0	0	0	0	1	1
Sum for each λ	1	-1	0	-1	1	0

Table 5. Values for $a_{\gamma,\lambda}$ used to compute $s_{(3,2,1)} = \sum_{\mu\vdash n} e_{\mu,(3,2,1)} Z_\mu$.

$$
\begin{aligned}
e_{\lambda,\gamma} &= \langle s_\gamma, DZ_\lambda \rangle \\
&= \langle \sum_{\mu} K^{-1}_{\mu,\gamma} h_\mu, \sum_{\beta \leq_r \lambda} a_{\beta,\lambda} m_\beta \rangle \\
&= \sum_{\beta \leq_r \lambda} K^{-1}_{\beta,\gamma} a_{\beta,\lambda}.
\end{aligned}
$$

We now present an example by expanding $s_{(3,2,1)}$ as a sum of ribbon Schur functions indexed by partitions. We can easily see that $s_{(3,2,1)} = h_1h_2h_3 - h_1h_1h_4 - h_3h_3 + h_1h_5$ by writing down all the special rim hook tabloids of shape $(3,2,1)$. Then

$$\langle s_{(3,2,1)}, DZ_\lambda \rangle = a_{(3,2,1),\lambda} - a_{(4,1,1),\lambda} - a_{(3,3),\lambda} + a_{(5,1),\lambda)}.$$

In Table 5, we present the relevant values of $a_{\mu,\lambda}$.

Thus

$$s_{(3,2,1)} = Z_{(3,2,1)} - Z_{(4,1,1)} - Z_{(4,2)} + Z_{(5,1)}.$$

This may not be the most efficient algorithm in all cases, for example another approach is to use a result of Lascoux and Pragacz [**7**] which gives the expansion of a Schur function as a product of ribbon Schur functions using a determinantal formula. Any product ribbon Schur functions can be simplified to a sum of ribbon Schur functions. However the ribbon Schur functions that result from such an expansion are just arbitrary Z_α where α is a composition. Thus one would need to expand $Z_\alpha = \sum_{\lambda\vdash n} f_{\lambda,\alpha} Z_\lambda$, where α is a composition of n and λ is a partition of n, as we did above. In special cases, such as when γ is a double hook, this method may be more efficient. However this method does not give a combinatorial interpretation of the coefficients of the Z_λ's that arise in the expansion.

7. Conclusions and Further Research

In this paper we have given combinatorial interpretations of the coefficients in the expansion of DZ_λ in terms of the monomial symmetric functions. We also found more indirect combinatorial interpretations of the expansion DZ_λ in terms of the Schur functions by using the inverse Kostka matrix. Moreover, we have given explicit formulas for such coefficients in many special cases.

There are many unanswered questions in this area. Of particular interest is what happens when we apply the ω transformation to DZ_λ. That is, recall the $\omega : \Lambda_n \to \Lambda_n$ is defined by the fact for all $\lambda \vdash n$, $\omega(h_\lambda) = e_\lambda$. Then the question is: can we give a combinatorial interpretation of $\omega(DZ_\lambda)$ in terms of $\{Z_\lambda\}_{\lambda\vdash n}$ or $\{DZ_\lambda\}_{\lambda\vdash n}$? We can clearly give a combinatorial interpretations of $\omega(DZ_\lambda)$ in terms of $\{f_\lambda\}_{\lambda\vdash n}$, since we can already expand DZ in terms of $\{m_\lambda\}_{\lambda\vdash n}$ and $\omega(m_\lambda) = f_\lambda$.

We also examined the coefficients in the expansion in terms of the power and elementary symmetric functions. Again the coefficients that arise in such expansions are not all positive. Thus another unanswered question is to find good combinatorial interpretations for the coefficients in the expansion of DZ_λ in terms of the other standard bases for the space of symmetric functions.

References

[1] D. Beck, J. Remmel, and T. Whitehead, The combinatorics of transition matrices between the bases of the symmetric functions and the B_n analogues, *Discrete Mathematics* **153** 1996, 3–27.

[2] L. Billera, H. Thomas, and S. van Willigenburg, Decomposable compositions, symmetric quasisymmetric functions and equality of ribbon Schur functions, preprint 2005, to appear in *Adv. Math.*.

[3] Ö. Eğecioğlu and J. Remmel, A combinatorial interpretation of the inverse Kostka matrix, *Linear Multilinear Algebra* **26** (1990), 59–84.

[4] I.M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V.S. Retakh and J.-Y. Thibon, Noncommutative symmetric functions, *Adv. in Math.*, **112** (1995), 218–348.

[5] I. Gessel and C. Reutenauer, Counting permutations with given cycle structure and descent set, *J. Combin. Theory Ser. A* **64** (1993), 189–215.

[6] A.N. Kirillov, A. Kuniba and T. Nakanishi, Skew diagram method in spectral decomposition of integrable lattice models, *Comm. Math. Phys.* **185** (1997), 441–465.

[7] A. Lascoux and P. Pragacz, Ribbon Schur functions, *European J. Combin.* **9** (1988), no. 6, 561–574.

[8] P.A. MacMahon, *Combinatory Analysis*, Cambridge University Press, 1917 (Vol. I), 1918 (Vol. II), reprint, Chelsea, New York, USA, 1960.

[9] R. Stanley, *Enumerative Combinatorics, Vol. 2*, Cambridge Studies in Advanced Mathematics, Vol. 62, Cambridge University Press, Cambridge, UK, 1999.

Department of Mathematics, University of California San Diego, La Jolla CA 92093-0112. USA
E-mail address: `ariehl@math.ucsd.edu`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A Labelling of the Faces in the Shi Arrangement

Felipe Rincón

Abstract. Let $\mathcal{F}_n$ be the face poset of the n-dimensional Shi arrangement, and let $\mathcal{P}_n$ be the poset of parking functions of length n with the order defined by $(a_1, a_2, \ldots, a_n) \leq (b_1, b_2, \ldots, b_n)$ if $a_i \leq b_i$ for all i. Pak and Stanley constructed a labelling of the regions in $\mathcal{F}_n$ by elements of $\mathcal{P}_n$. We extend this in a natural way to a labelling of all faces in $\mathcal{F}_n$ by closed intervals of $\mathcal{P}_n$, and explore some interesting and unexpected properties of this bijection. We give some results that contribute to characterize the intervals that appear as labels and consequently to a better comprehension of $\mathcal{F}_n$.

Résumé. Soit $\mathcal{F}_n$ l'ensemble partiellement ordonné des faces de l'arrangement de Shi en dimension n, et soit $\mathcal{P}_n$ l'ensemble partiellement ordonné des fonctions de parking de longueur n dont l'ordre est défini par $(a_1, a_2, \ldots, a_n) \leq (b_1, b_2, \ldots, b_n)$ si $a_i \leq b_i$ pout tout i. Pak et Stanley ont construit un étiquetage des regions de $\mathcal{F}_n$ avec des éléments de $\mathcal{P}_n$. On généralise cette étude de manière naturelle à un étiquetage à toutes les faces de $\mathcal{F}_n$ en utilisant des intervalles fermés de $\mathcal{P}_n$ et on éxamine quelques curieuses et inattendues propriétés de cette bijection. On donne des resultats qui contribuent à caractériser les intervalles qui apparaissent comme étiquèttes et ainsi une meilleure compréhension de $\mathcal{F}_n$.

1. Preliminaries

1.1. The Shi arrangement. A n-dimensional (real) hyperplane arrangement is a finite collection of affine hyperplanes in $\mathbb{R}^n$. Any hyperplane arrangement $\mathcal{A}$ cuts $\mathbb{R}^n$ into open regions that are polyhedra (called the regions of $\mathcal{A}$), so they have faces. More specifically, faces of $\mathcal{A}$ are nonempty intersections between the closure of a region and some or none hyperplanes in $\mathcal{A}$. The poset consisting of all these faces ordered by inclusion is called the face poset of $\mathcal{A}$.

The n-dimensional Shi arrangement $\mathcal{S}_n$ consists of the $n(n-1)$ hyperplanes

$$\mathcal{S}_n : \quad x_i - x_j = 0, 1 \qquad \text{for } 1 \leq i < j \leq n.$$

Let $\mathcal{F}_n$ be the face poset of $\mathcal{S}_n$, and let $\mathcal{R}_n$ be the set of n-dimensional faces in $\mathcal{F}_n$. Then $\mathcal{R}_n$ is the set of closures of the regions of $\mathcal{S}_n$. However we will identify the regions of $\mathcal{S}_n$ with their closure, so we will make no distinction between the elements of $\mathcal{R}_n$ and the regions of $\mathcal{S}_n$. This arrangement was first considered by Shi [**4**], who showed that $|\mathcal{R}_n| = (n+1)^{n-1}$.

Faces of any hyperplane arrangement $\mathcal{A}$ can be described by specifying for every $H \in \mathcal{A}$, which side of H contains the face. That is, for any $H \in \mathcal{A}$ define H^+ and H^- as the two closed halfspaces determined by H (the choice of which one is H^+ is arbitrary), and let $H^0 = H$. Then the elements in the face poset of $\mathcal{A}$ are precisely the nonempty intersections of the form

$$F = \bigcap_{H \in \mathcal{A}} H^{\sigma_H}$$

where $\sigma_H \in \{+, -, 0\}$. So every face F is encoded by its sign sequence $(\sigma_H)_{H \in \mathcal{A}}$, where $\sigma_H \neq 0$ if and only if $F \subseteq H^{\sigma_H}$ and $F \nsubseteq H$.

2000 *Mathematics Subject Classification.* Primary 52C35; Secondary 06A07.

Key words and phrases. hyperplane arrangements, Shi arrangement, parking functions, face poset.

These results will be part of the author's undergraduate thesis in mathematics in Universidad de los Andes under the direction of Federico Ardila. The author is very grateful for all his help and suggestions.

For the Shi arrangement it is useful to represent this sequence as a matrix. We will assume as convention that for $i < j$ if $H : x_i = x_j$ then $H^- : x_i \geq x_j$, and if $H : x_i = x_j + 1$ then $H^- : x_i \leq x_j + 1$. First define $\mathcal{M}_n$ as the set of all $n \times n$ matrices whose entries belong to $\{+, -, 0\}$. Then for any $F \in \mathcal{F}_n$ consider its sign sequence $(\sigma_H)_{H \in \mathcal{S}_n}$, and define its *associated matrix* $M_F \in \mathcal{M}_n$ as follows:

$$(M_F)_{i,j} = \begin{cases} \sigma_H & \text{if } j < i, \text{ where } H : x_j = x_i \\ \sigma_H & \text{if } i < j, \text{ where } H : x_i = x_j + 1 \\ 0 & \text{if } i = j. \end{cases}$$

For example, the matrix associated to the region defined by $x_n \leq x_{n-1} \leq \ldots \leq x_1 \leq x_n + 1$ has all entries equal to $-$, except for diagonal ones which are 0. In general, if $F \in \mathcal{F}_n$ then F is a region if and only if all non-diagonal entries of M_F are different from zero. And if $F, G \in \mathcal{F}_n$ then $F \subseteq G$ if and only if M_G has the same entries as M_F except for some non-diagonal zero entries of M_F which become $-$ or $+$ in M_G.

However, there is another way of representing a face that will be very useful for us. For notation simplicity, if n is a positive integer let $[n] = \{1, 2, \ldots, n\}$. Now, if $F \in \mathcal{F}_n$, we will say a function $X : [n] \to \mathbb{R}$ is an *interval representation* of F if the point $\big(X(1), X(2), \ldots, X(n)\big) \in \mathbb{R}^n$ belongs to F and not to any other face properly contained in F. We will denote by $\mathcal{X}_n$ the set of all functions from $[n]$ to $\mathbb{R}$. Two interval representations $X, X' \in \mathcal{X}_n$ will be called *equivalent* if they represent the same face. We can imagine these interval representations as ways in which n numbered intervals of length 1 can be placed on the real line: any $X \in \mathcal{X}_n$ can be thought as the collection of the n intervals $[X(i), X(i) + 1]$ for $i \in [n]$. Interval $[X(i), X(i) + 1]$ will be refered as the i-th interval of X. So the face represented by X is determined only by the relative position of the endpoints of the intervals of X.

1.2. Parking functions. A parking function of length n is a sequence $P = (P_1, P_2, \ldots, P_n) \in [n]^n$ such that if $Q_1 \leq Q_2 \leq \ldots \leq Q_n$ is the increasing rearrangement of the terms of P, then $Q_i \leq i$. Parking functions were first considered by Konheim and Weiss [**3**] under a slightly different definition, but equivalent to ours. Let $\mathcal{P}_n$ be the poset of the parking functions of length n with the order defined by $(P_1, P_2, \ldots, P_n) \leq (Q_1, Q_2, \ldots, Q_n)$ if $P_i \leq Q_i$ for all $i \in [n]$.

Pak and Stanley constructed a bijection between $\mathcal{R}_n$ and the parking functions of length n as follows [**5**]: Let $R_0 \in \mathcal{R}_n$ be the region defined by $x_n \leq x_{n-1} \leq \ldots \leq x_1 \leq x_n + 1$, and define its label $\lambda(R_0) = (1, 1, \ldots, 1) \in \mathbb{Z}^n$. Suppose that $R, R' \in \mathcal{R}_n$, R is labelled and R' is unlabelled, R and R' are only separated by the hyperplane $H : x_i = x_j$ $(i < j)$, and R_0 and R are on the same side of H; then define $\lambda(R') = \lambda(R) + e_i$ ($e_i \in \mathbb{Z}^n$ is the i-th vector of the canonical basis). If under the same hypothesis R and R' are only separated by the hyperplane $H : x_i = x_j + 1$ $(i < j)$ and R_0 and R are on the same side of H; then define $\lambda(R') = \lambda(R) + e_j$.

Figure 1 shows the projection of the arrangement $\mathcal{S}_3$ on the plane defined by $x + y + z = 0$, and the labelling of the regions in a simplified notation.

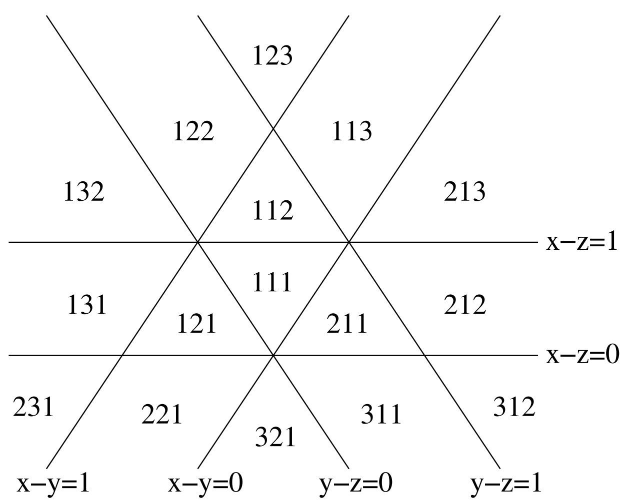


FIGURE 1. Arrangement $\mathcal{S}_3$ and the labelling λ

Notice that in our convention, if $R \in \mathcal{R}_n$ then we have that $\lambda(R) = (a_1 + 1, a_2 + 1, \dots, a_n + 1)$ where a_i is the number of $+$ entries in the i-th column of M_R. Stanley showed that this labelling was in fact a bijection between $\mathcal{R}_n$ and the elements of $\mathcal{P}_n$, that is, he showed that all these labels were parking functions, each appearing once.

2. The labelling of $\mathcal{F}_n$

We will now extend this labelling to all faces in $\mathcal{F}_n$. First we prove a lemma that allows us to define the labelling.

Lemma 2.1. *Let $F \in \mathcal{F}_n$. Then there exist two unique regions $F^-, F^+ \in \mathcal{R}_n$ such that $F \subseteq F^-$, $F \subseteq F^+$ and for any region $R \in \mathcal{R}_n$, if $F \subseteq R$ then $\lambda(F^-) \leq \lambda(R) \leq \lambda(F^+)$ in $\mathcal{P}_n$. Moreover, $F^- \cap F^+ = F$.*

Proof. Consider an interval representation $X \in \mathcal{X}_n$ of F. Clearly the lemma is true if $F \in \mathcal{R}_n$, that is, if there are no equalities in X of the form $X(i) = X(j)$ or $X(i) = X(j) + 1$ with $i < j$, because in this case $F^- = F^+ = F$. In other case, let r be the maximum $X(i)$ for which there exists a $j > i$ such that $X(i) = X(j)$ or $X(i) = X(j) + 1$. Take k as the maximum i such that $X(i) = r$. Define a new interval representation $X' \in \mathcal{X}_n$ by

$$X'(i) = \begin{cases} X(i) & \text{if } i \neq k \\ X(i) + \epsilon & \text{if } i = k \end{cases}$$

where ϵ is a sufficiently small positive real number so that for all j, if $X(k) < X(j)$ then $X(k) + \epsilon < X(j)$, and if $X(k) < X(j) + 1$ then $X(k) + \epsilon < X(j) + 1$. So X' is the same interval representation as X, but its k-th interval is moved a little bit to the right. Let $F' \in \mathcal{F}_n$ be the face represented by X'.

By the definition of X' it is clear that inequalities in X remain unchanged in X', and also equalities that do not involve $X(k)$. That is, if $X(i) < X(j)$ then $X'(i) < X'(j)$, if $X(i) < X(j) + 1$ with $i < j$ then $X'(i) < X'(j) + 1$, and if $X(i) > X(j) + 1$ with $i < j$ then $X'(i) > X'(j) + 1$. Also if $X(i) = X(j)$ and $i, j \neq k$ then $X'(i) = X'(j)$, and if $X(i) = X(j) + 1$ with $i < j$ and $i, j \neq k$ then $X'(i) = X'(j) + 1$. Notice as well that there are no equalities in X of the form $X(i) = X(k) + 1$ with $i < k$ because it imply be a contradiction with the maximality of r, neither equalities of the form $X(k) = X(i)$ with $k < i$ because they contradict the choice of k. So all equalities in X involving $X(k)$ must be of the form $X(k) = X(i) + 1$ with $k < i$, or $X(i) = X(k)$ with $i < k$. In the first case we have that $X'(k) > X(i) + 1 = X'(i) + 1$, so $(M_{F'})_{k,i} = +$. In the second case $X'(i) = X(i) < X'(k)$, so $(M_{F'})_{k,i} = +$. All this shows that $M_{F'}$ has the same entries as M_F except for the non-diagonal zero entries in the k-th row and k-th column of M_F, which become $+$ in $M_{F'}$.

If we repeat this construction starting with the face F' we obtain a face F'', satisfying that $M_{F''}$ has the same entries as $M_{F'}$ except for some non-diagonal zero entries in $M_{F'}$ that become $+$ in $M_{F''}$. And continuing with this process we finally get a face F^+, such that M_{F^+} is the same matrix as M_F but replacing all its non-diagonal zero entries by $+$.

Consider now the same construction, but defining X' by moving the k-th interval of X a little bit to the left. The non-diagonal zero entries in the k-th row and k-th column of M_F become now $-$ in $M_{F'}$, so repeating the process we finally get a face F^- such that M_{F^-} is the same matrix as M_F but replacing all its non-diagonal zero entries by $-$.

By this description of their associated matrices, it is easy to see that $F^+ \cap F^- = F$. Now, let $R \in \mathcal{R}_n$ be any region containing F. Remember that M_R must be the same matrix as M_F, but changing the non-diagonal zero entries in M_F by $-$ or $+$. Then for every $i \in [n]$ the number of $+$ entries in the i-th column of M_R must be at least the number of $+$ entries in the i-th column of M_{F^-}, and at most the number of $+$ entries in the i-th column of M_{F^+}. Hence $\lambda(R) \in \mathcal{P}_n$ must satisfy the relation $\big(\lambda(F^-)\big)_i \leq \big(\lambda(R)\big)_i \leq \big(\lambda(F^+)\big)_i$ for all i, that is, $\lambda(F^-) \leq \lambda(R) \leq \lambda(F^+)$ in $\mathcal{P}_n$. This property implies easily the uniqueness of F^- and F^+, so the proof is complete. ☐

This lemma is interesting by itself, as the following result shows.

Corollary 2.2. *Let $R_1, R_2, \dots, R_k \in \mathcal{R}_n$, and define $P^i = (P_1^i, P_2^i, \dots, P_n^i) = \lambda(R_i)$ for $1 \leq i \leq k$. If*

$$Q = \left(\max_i P_1^i, \max_i P_2^i, \dots, \max_i P_n^i\right)$$

is not a parking function then $\bigcap_{i=1}^k R_i = \emptyset$.

PROOF. If $F = \bigcap_{i=1}^{k} R_i \neq \emptyset$ then $F \in \mathcal{F}_n$. Hence by the Lemma we have that $P^i \leq \lambda(F^+)$ for all i, but this implies that Q is a parking function. $\square$

We now define the labelling of the faces in $\mathcal{F}_n$. Denote by $\mathrm{Int}(\mathcal{P}_n)$ the set of all closed intervals of $\mathcal{P}_n$.

DEFINITION 2.3. The labelling $\lambda : \mathcal{F}_n \to \mathrm{Int}(\mathcal{P}_n)$ is defined by $\lambda(F) = \big[\lambda(F^-), \lambda(F^+)\big]$.

We will use λ also for this labelling because it can be considered as an extension of the labelling we had for regions (by identifying $\lambda(R)$ with $\big\{\lambda(R)\big\}$).

Notice that different faces have different labels, because $F^- \cap F^+ = F$ for all $F \in \mathcal{F}_n$. Unfortunately, not all closed intervals of $\mathcal{P}_n$ appear as labels of some face.

3. Properties of the labelling

Clearly the main property of this labelling is stated in the following surprising theorem.

THEOREM 3.1. *Let $F \in \mathcal{F}_n$. Then $\lambda(F) = \big\{\lambda(R) \mid R \in \mathcal{R}_n$ and $F \subseteq R\big\}$.*

PROOF. Let $I(F) = \big\{\lambda(R) \mid R \in \mathcal{R}_n \text{ and } F \subseteq R\big\}$. Lemma 2.1 tells us that $I(F) \subseteq \lambda(F)$. Now, notice that

$$\Big|\lambda(F)\Big| = \prod_{i=1}^{n} \Big(\big(\lambda(F^+)\big)_i - \big(\lambda(F^-)\big)_i + 1\Big)$$

because $P = (P_1, P_2, \ldots, P_n)$ is a parking function in $\lambda(F)$ if and only if $\big(\lambda(F^-)\big)_i \leq P_i \leq \big(\lambda(F^+)\big)_i$ for all i. Now let $X \in \mathcal{X}_n$ be an interval representation of F, and define $A(F,i) = \big\{j \in [n] \mid j > i \text{ and } X(i) = X(j)\big\}$ and $B(F,i) = \big\{j \in [n] \mid j < i \text{ and } X(j) = X(i)+1\big\}$. Then $c(F,i) = |A_i| + |B_i|$ is the number of non-diagonal zero entries in the i-th column of M_F. So $c(F,i)$ is the difference between the number of $+$ entries in the i-th column of M_{F^+} and the number of $+$ entries in the i-th column of M_{F^-}. Hence $c(F,i) = \big(\lambda(F^+)\big)_i - \big(\lambda(F^-)\big)_i$, and

$$\Big|\lambda(F)\Big| = \prod_{i=1}^{n} \big(c(F,i) + 1\big).$$

We will then prove that $\big|I(F)\big| \geq \prod_{i=1}^{n} \big(c(F,i)+1\big)$, which is equivalent to the equality between $I(F)$ and $\lambda(F)$ by a cardinality argument. Notice that $\big|I(F)\big|$ is the number of regions that contain F as a face. Then $\big|I(F)\big|$ is the number of ways (up to equivalence) in which the intervals of X can be moved a little bit, changing all equalities in X of the form $X(i) = X(j)$ or $X(i) = X(j)+1$ $(1 \leq i < j \leq n)$ to inequalities. So we will prove there are at least $\prod_{i=1}^{n} \big(c(F,i)+1\big)$ different ways of doing this.

The proof is by induction on n. If $n = 2$ there are 5 faces in $\mathcal{F}_2$, and it is easy to check that for each one of them the equality holds. Now assume the assertion is true for $n-1$. Consider $F \in \mathcal{F}_n$ and let $X \in \mathcal{X}_n$ be an interval representation of F. Let r be the minimum $X(i)$, and let k be the minimum i such that $X(i) = r$. By the choice of k there is no i such that $i < k$ and $X(i) = X(k)$, or $i > k$ and $X(k) = X(i)+1$. That is, for all $i \neq k$ we have that $k \notin A(F,i)$ and $k \notin B(F,i)$. Then, ignoring the k-th interval, by induction hypothesis there are at least $\prod_{i \neq k} \big(c(F,i)+1\big)$ different ways of moving (as explained before) all intervals of X except the k-th interval. Consider one of these ways in which these intervals can be moved, and for $i \neq k$ let $X'(i)$ be the new position of the i-th interval. We can assume without loss of generality that the intervals were moved very little, so that there exists an open interval U around $X(k)+1$ such that $X'(i)+1 \in U$ if and only if $X(k)+1 = X(i)+1$, and $X'(i) \in U$ if and only if $X(i) = X(k)+1$. Then the $c(F,k)$ points of $\big\{X'(i)+1 \mid i \in A(F,k)\big\} \cup \big\{X'(i) \mid i \in B(F,k)\big\}$ separate the interval U in $c(F,k)+1$ disjoint open intervals $U_0, U_1, \ldots, U_{c(F,k)}$. For every j such that $0 \leq j \leq c(F,k)$ let z_j be some point inside interval U_j, and define $Y_j \in \mathcal{X}_n$ as follows:

$$Y_j(i) = \begin{cases} X'(i) & i \neq k \\ z_j & \text{if } i = k. \end{cases}$$

So Y_j is an interval representation obtained by moving all intervals of X a little bit (as explained before). Because U was chosen sufficiently small, Y_j represents a region in $\mathcal{R}_n$ that contains F. Moreover, if $i \neq j$ then Y_i and Y_j represent different regions, because $Y_i(k) \in U_i$ and $Y_j(k) \in U_j$. So we have proved that for every way of moving all intervals of X except the k-th interval there are at least $c(F,k)+1$ different regions in $\mathcal{R}_n$ that contain F. Hence

$$\big|I(F)\big| \geq \big(c(F,k)+1\big)\prod_{i\neq k}\big(c(F,i)+1\big) = \prod_{i=1}^{n}\big(c(F,i)+1\big)$$

as we wanted, so the proof is complete. □

This Theorem can also be stated as follows.

Corollary 3.2. *Let $R_1, R_2, \ldots, R_k \in \mathcal{R}_n$ such that $F \subseteq \bigcap_{i=1}^{k} R_i \neq \emptyset$, and define $P^i = (P^i_1, P^i_2, \ldots, P^i_n) = \lambda(R_i)$ for $1 \leq i \leq k$. If $R \in \mathcal{R}_n$ is such that*

$$\Big(\min_i P^i_1, \min_i P^i_2, \ldots, \min_i P^i_n\Big) \leq \lambda(R) \leq \Big(\max_i P^i_1, \max_i P^i_2, \ldots, \max_i P^i_n\Big)$$

then $F \subseteq R$.

Another important consequence is stated in the next corollary.

Corollary 3.3. *Let $F, G \in \mathcal{F}_n$. Then $F \subseteq G$ if and only if $\lambda(F) \supseteq \lambda(G)$.*

So if we define $\mathcal{I}_n$ as the poset of all intervals of $\mathcal{P}_n$ appearing as labels, ordered by reverse inclusion, then λ is an isomorphism between $\mathcal{F}_n$ and $\mathcal{I}_n$. This means that the characterization of all intervals in $\mathcal{I}_n$ will give us a complete combinatorial description of $\mathcal{F}_n$. We already know that all intervals of $\mathcal{P}_n$ consisting of exactly one element appear in $\mathcal{I}_n$ as labels of some region.

Now, every $F \in \mathcal{F}_n$ has a dimension, which determines the rank of F in the poset $\mathcal{F}_n$. To see how this dimension is represented in $\mathcal{I}_n$ we need the following definition.

Let $X \in \mathcal{X}_n$. A *chain* of X is a k-tuple $(a_1, a_2, \ldots, a_k) \in [n]^k$ constructed as follows:

- Choose a_1 so that there is no $i < a_1$ such that $X(i) = X(a_1)+1$, neither $i > a_1$ such that $X(a_1) = X(i)$.
- Once a_j has been chosen, if there exists some $i < a_j$ such that $X(i) = X(a_j)$ then $a_{j+1} = \max\big\{i < a_j \mid X(i) = X(a_j)\big\}$. If this i does not exist, but there exists some $l > a_j$ such that $X(a_j) = X(l)+1$, then $a_{j+1} = \max\big\{l > a_j \mid X(a_j) = X(l)+1\big\}$.
- The chain ends when there are no such i nor l as in the last step.

X can have several different chains, but the definition implies that all of them must be disjoint, and every $i \in [n]$ must belong to some chain of X. It is easy to see that chains represent sets of intervals that are binded one to another in X. That is, if we move a little bit the j-th interval to obtain a new interval representation $X' \in \mathcal{X}_n$, then for all i in the same chain as j we must also move the i-th interval in the same way if we want X' to represent the same face as X. Hence, the number of chains of X is the dimension of the face represented by X.

Proposition 3.1. *Let $F \in \mathcal{F}_n$, and $\lambda(F) = [P,Q]$. Then* $\dim(F) = \big|\{i \in [n] \mid P_i = Q_i\}\big|$.

Proof. Let $X \in \mathcal{X}_n$ be an interval representation of F. Remember the definitions of $A(F,i)$, $B(F,i)$ and $c(F,i)$ given in the proof of Theorem 3.1. Notice that if $H = (a_1, a_2, \ldots, a_k) \in [n]^k$ is a chain of X then $c(F,a_j) = 0$ if and only if $j = 1$, because $a_j \in A(F,a_{j+1}) \cup B(F,a_{j+1})$. So the number of chains of X is equal to the number of $i \in [n]$ such that $c(F,i) = 0$. But we had seen that $c(F,i) = Q_i - P_i$, so the proof is complete. □

Continuing with the same ideas we can prove the following proposition.

Proposition 3.2. *Let $F \in \mathcal{F}_n$, and $\lambda(F) = [P,Q]$. Then*

$$\{Q_1 - P_1, Q_2 - P_2, \ldots, Q_n - P_n\} = \{0, 1, 2, \ldots, m\}$$

for some $m \in \mathbb{N}$.

Proof. Let $X \in \mathcal{X}_n$ be an interval representation of F. Notice that if $H = (a_1, a_2, \ldots, a_k) \in [n]^k$ is a chain of X then $A(F,a_{j+1}) \cup B(F,a_{j+1}) \subseteq A(F,a_j) \cup B(F,a_j) \cup \{a_j\}$for all j, so $c(F,a_{j+1}) \leq c(F,a_j)+1$. Then $Q_{a_{j+1}} - P_{a_{j+1}} \leq Q_{a_j} - P_{a_j} + 1$ for all j. Remembering that $Q_{a_1} - P_{a_1} = 0$ we have that $\{Q_{a_1} - P_{a_1}, Q_{a_2} - P_{a_2}, \ldots, Q_{a_k} - P_{a_k}\} = \{0, 1, \ldots, m_H\}$ for some $m_H \in \mathbb{N}$. Therefore, by taking the union over all chains of X, the proof is finished. □

Last proposition restricts a lot the possible intervals appearing as labels, and makes a step toward the characterization of the elements of $\mathcal{I}_n$.

We now characterize the possible sizes of the intervals that appear as labels of faces of a fixed dimension.

PROPOSITION 3.3. *The set* $\{|\lambda(F)| \mid F \in \mathcal{F}_n$ *and* $\dim(F) = k\}$ *is the set of all positive numbers* d *such that* $d = 2^{a_1}3^{a_2}\dots(m+1)^{a_m}$ *for some* $m \in \mathbb{N}$*, where* $a_i > 0$ *for all* $i \le m$*, and* $a_1 + a_2 + \ldots + a_m = n - k$.

PROOF. Let $F \in \mathcal{F}_n$ be a face such that $\dim(F) = k$, and let $\lambda(F) = [P, Q]$. Define $a_i = \big|\{j \mid Q_j - P_j = i\}\big|$. Proposition 3.2 tells us there exists $m \in \mathbb{N}$ such that $a_i > 0$ if and only if $i \le m$. Then

$$\big|\lambda(F)\big| = \big|[P,Q]\big| = \prod_{i=1}^{n}(Q_i - P_i + 1) = 2^{a_1}3^{a_2}\dots(m+1)^{a_m}.$$

It is clear that $a_0 + a_1 + \ldots + a_m = n$, so by Proposition 3.1 we have that $a_1 + a_2 + \ldots + a_m = n - k$.

On the other hand, if we take $a_0, a_1, \ldots, a_m$ such that $a_i > 0$ for all $i \le m$ and $a_0 + a_1 + \ldots + a_m = n$ then it is easy to construct an interval representation X of a face $F \in \mathcal{F}_n$ satisfying $a_i = \big|\{j \mid c(F,j) = i\}\big|$. Therefore, remembering that if $\lambda(F) = [P,Q]$ then $c(F,j) = Q_j - P_j$, the proposition follows. □

Remember that if $F \in \mathcal{F}_n$ then $\big|\lambda(F)\big| = \big|\{R \in \mathcal{R}_n \mid F \subseteq R\}\big|$, so this proposition is also giving some geometrical information about the Shi arrangement.

Finally, we characterize the intervals appearing as labels of 1-dimensional faces.

PROPOSITION 3.4. *Let* $I = [P,Q]$ *be an interval of* $\mathcal{P}_n$*. Then* I *is the label of a* 1*-dimensional face if and only if the following statements hold:*

- Q *is a permutation of* $[n]$.
- P *is determined by* Q *in the following way. Denote* $(a_1, a_2, \ldots, a_n) = (Q^{-1}(1), Q^{-1}(2), \ldots, Q^{-1}(n))$*, and let* $0 = i_0 < i_1 < i_2 < \ldots < i_k = n$ *be the numbers such that*
$$\{i_1, i_2, \ldots, i_{k-1}\} = \{j \in [n] \mid a_j < a_{j+1}\}.$$
Then for all $r \in [n]$*, if* j *is such that* $i_j < r \le i_{j+1}$ *we have that*
$$P_{a_r} = i_{j-1} + \big|\{l \in [n] \mid i_{j-1} < l \le i_j \text{ and } a_l > a_r\}\big| + 1,$$
where $i_{-1} = 0$.

PROOF. To see that the conditions are necessary, let $F \in \mathcal{F}_n$ be a 1-dimensional face such that $\lambda(F) = [P,Q]$, and let $X \in \mathcal{X}_n$ be an interval representation of F. Then X consists only of one chain $H = (b_1, b_2, \ldots, b_n)$. Remember that $Q_i - 1$ is the number of non-diagonal + or 0 entries in the i-th column of M_F, that is,

$$Q_i = \big|\{j \in [n] \mid j > i \text{ and } X(j) \ge X(i)\}\big| + \big|\{j \in [n] \mid j < i \text{ and } X(j) \ge X(i) + 1\}\big| + 1.$$

But all intervals of X are on the same chain, so we have that for all i

$$\{j \in [n] \mid j > b_i \text{ and } X(j) \ge X(b_i)\} \cup \{j \in [n] \mid j < b_i \text{ and } X(j) \ge X(b_i) + 1\} = \{b_1, b_2, \ldots, b_{i-1}\},$$

hence $Q_{b_i} = i$. This shows that Q is a permutation of $[n]$, and that $a_i = b_i$ for all i.

Notice that the numbers $i_0, i_1, \ldots, i_k$ satisfy that for all m, $i_j < m \le i_{j+1}$ if and only if $X(a_m) = X(a_1) - j$. Then $i_j = \big|\{l \in [n] \mid X(l) > X(a_1) - j\}\big|$. Remember also that $P_i - 1$ is the number of + entries in the i-th column of M_F, that is,

$$P_i = \big|\{j \in [n] \mid j > i \text{ and } X(j) > X(i)\}\big| + \big|\{j \in [n] \mid j < i \text{ and } X(j) > X(i) + 1\}\big| + 1.$$

Let $r \in [n]$ and j such that $i_j < r \le i_{j+1}$, so $X(a_r) = X(a_1) - j$. Therefore, because X consists only of the chain H,

$$\begin{aligned}
P_{a_r} &= \Big|\big\{l \mid l > a_r \text{ and } X(l) > X(a_r)\big\}\Big| + \Big|\big\{l \mid l < a_r \text{ and } X(l) > X(a_r) + 1\big\}\Big| + 1 \\
&= \Big|\big\{l \mid l > a_r \text{ and } X(l) = X(a_r) + 1\big\}\Big| + \Big|\big\{l \mid X(l) > X(a_r) + 1\big\}\Big| + 1 \\
&= \Big|\big\{l \mid l > a_r \text{ and } X(l) = X(a_1) - (j-1)\big\}\Big| + \Big|\big\{l \mid X(l) > X(a_1) - (j-1)\big\}\Big| + 1 \\
&= \Big|\big\{m \mid a_m > a_r \text{ and } i_{j-1} < m \le i_j\big\}\Big| + i_{j-1} + 1,
\end{aligned}$$

as we wanted.

On the other hand, it is easy to see that if $[P,Q]$ is an interval of $\mathcal{P}_n$ satisfying the previous conditions, then it appears as the label of a 1-dimensional face. In fact, the function $X \in \mathcal{X}_n$ defined by

$$X\left(Q^{-1}(i)\right) = -\left|\left\{l \in [n] \mid l < i \text{ and } Q^{-1}(l) < Q^{-1}(l+1)\right\}\right|$$

represents a 1-dimensional face F such that $\lambda(F) = [P,Q]$. □

This characterization has an interesting corollary.

Corollary 3.4. *Each region $R \in \mathcal{R}_n$ such that $\lambda(R)$ is a permutation of $[n]$ contains a unique 1-dimensional face $F \in \mathcal{F}_n$. Moreover, each 1-dimensional face $F \in \mathcal{F}_n$ is contained in a unique region $R \in \mathcal{R}_n$ such that $\lambda(R)$ is a permutation of $[n]$.*

Proof. Suppose R is a region such that $\lambda(R)$ is a permutation of $[n]$. By the last characterization we know that there exists a unique $P \in \mathcal{P}_n$ such that $[P,Q]$ is the label of a 1-dimensional face F. Theorem 3.1 implies that $F \subseteq R$. Moreover, if F' is a one dimensional face contained in R then $Q \in \lambda(F')$, and because Q is a maximal element of $\mathcal{P}_n$ we have that $\lambda(F') = [P', Q]$ for some $P' \in \mathcal{P}_n$. Therefore $P = P'$ and $F = F'$, so the face F is unique.

Now, if F is a 1-dimensional face then by the characterization $\lambda(F) = [P,Q]$, with Q a permutation of $[n]$. By Theorem 3.1, if R is the region such that $\lambda(R) = Q$ then $F \subseteq R$. Moreover, if R' is a region that contains F then $Q' = \lambda(R') \in [P,Q]$. Therefore, if Q' is a permutation of $[n]$ then $Q' = Q$, because Q' is a maximal element of $\mathcal{P}_n$. So $R = R'$, proving that the region R is unique. □

Corollary 3.5. *The number of 1-dimensional faces of $\mathcal{S}_n$ is $n!$.*

This is a particular example of a general result first stated by Athanasiadis [**1**]. However, this bijective proof allows a better comprehension of the geometrical organization of these faces.

4. Perspectives

After developing these results, it seems clear that there are still many aspects to understand about this labelling. We are now working on three main problems. In first place, we are trying to achieve a total and simple characterization of the intervals of $\mathcal{I}_n$. This would give a complete combinatorial description of the poset $\mathcal{F}_n$, thus a better comprehension of the geometry of the Shi arrangement. We are also trying to generalize to higher dimensions the way in which 1-dimensional faces were counted, obtaining this way a similar result to the one given by Athanasiadis [**1**]. Finally, we want to apply all these results to the theory of random walks on hyperplane arrangements, as defined by Brown and Diaconis in [**2**].

References

[1] C.A. Athanasiadis, Characteristic polynomials of subspace arrangements and finite fields, *Adv. Math.* 122 (1996), 193-233.

[2] K. S. Brown and P. Diaconis, Random walks and hyperplane arrangements, *The Annals of Probability* 26 (1998), no. 4, 1813-1854.

[3] A. G. Konheim and B. Weiss, An occupancy discipline and applications, *SIAM J. Applied Math.* 14 (1966), 1266-1274.

[4] J.-Y. Shi, The Kazhdan-Lusztig Cells in Certain Affine Weyl Groups, *Lecture Notes in Mathematics*, no. 1179, Springer, Berlin/Heidelberg/New York, 1986.

[5] R. Stanley, Hyperplane arrangements, parking functions and tree inversions, *Mathematical Essays in Honor of Gian-Carlo Rota*, Birkhäuser, Boston/Basel/Berlin, 1998, 359-375.

Departamento de Matemáticas, Universidad de los Andes, Bogotá, Colombia

E-mail address: `ed-rinco@uniandes.edu.co`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

"Elliptic" enumeration of nonintersecting lattice paths

Michael Schlosser

Abstract. We enumerate lattice paths in $\mathbb{Z}^2$ consisting of unit vertical and horizontal steps in the positive direction using elliptic weights, composed of appropriately chosen products of theta functions. The "elliptic" generating function of paths from a given starting point to a given end point evaluates, by virtue of Riemann's addition formula for theta functions and induction, to an elliptic generalization of the binomial coefficient. Convolution gives an identity equivalent to Frenkel and Turaev's ${}_{10}V_9$ summation. (This appears to be the first combinatorial proof of the latter, and at the same time of some important degenerate cases including Jackson's ${}_8\phi_7$ and Dougall's ${}_7F_6$ summation.) We then turn to nonintersecting lattice paths in $\mathbb{Z}^2$ where, using the Lindström–Gessel–Viennot theory combined with an elliptic determinant evaluation by Warnaar, we compute the elliptic generating function of selected families of paths with given starting points and end points. Here convolution gives a multivariate extension of the ${}_{10}V_9$ summation which turns out to be a special case of an identity originally conjectured by Warnaar, later proved by Rosengren. We conclude with discussing some future perspectives.

Résumé. On énumère les chemins dans le réseau $\mathbb{Z}^2$, dont chaque pas unitaire est vertical ou horizontal dans le sens positif, par rapport à un poids elliptique, qui est produit choisis de façon apropriée de fonctions théta. L'évaluation de la fonction génératrice "elliptique" des chemins d'un point de départ donné à un point d'arrivée donné, à l'aide de la formule d'addition de Riemann pour fonctions theta et récurrence, donne lieu à une généralization du coefficient binomial. La formule de convolution donne une identitée équivalente à la formule sommatoire ${}_{10}V_9$ de Frenkel and Turaev. (Il semble que c'est la première preuve combinatoire de la dernière, et en même temps de certains cas importants dégénérés comprenant les formules sommatoires ${}_8\phi_7$ de Jackson et ${}_7F_6$ de Dougall.) On tourne ensuite vers les chemins non intersectant dans $\mathbb{Z}^2$. En utilisant la théorie de Lindström–Gessel–Viennot couplée avec l'évaluation d'un déterminant elliptique de Warnaar, on calcule la fonction génératrice elliptique de certaines familles choisies de chemins avec les points de départs et d'arrivées donnés. Dans ce cas la formule de convolution donne une extension multivariée de la formule sommatoire ${}_{10}V_9$ qui s'est avéré un cas particulier d'une identité originalement conjecturée par Warnaar, et puis démontrée par Rosengren. On conclut avec quelques discussions sur la perspective d'avenir.

1. Preliminaries

1.1. Lattice paths in $\mathbb{Z}^2$. We consider lattice paths in the plane integer lattice $\mathbb{Z}^2$ consisting of unit horizontal and vertical steps in the positive direction. Given points u and v, we denote the set of all lattice paths from u to v by $\mathcal{P}(u \to v)$. If $\mathbf{u} = (u_1, \dots, u_r)$ and $\mathbf{v} = (v_1, \dots, v_r)$ are vectors of points, we denote the set of all r-tuples $(P_1, \dots, P_r)$ of paths where P_i runs from u_i to v_i, $i = 1, \dots, r$, by $\mathcal{P}(\mathbf{u} \to \mathbf{v})$. A set of paths is *nonintersecting* if no two paths have a point in common. The set of all nonintersecting paths from $\mathbf{u}$ to $\mathbf{v}$ is denoted $\mathcal{P}_+(\mathbf{u} \to \mathbf{v})$. Let w be a function which assigns to each horizontal edge e in $\mathbb{Z}^2$ a *weight* $w(e)$. The weight $w(P)$ of a path P is defined to be the product of the weights of all its horizontal steps. The weight $w(\mathbf{P})$ of an r-tuple $\mathbf{P} = (P_1, \dots, P_r)$ of paths is defined to be the product $\prod_{i=1}^r w(P_i)$ of the weights of all

2000 *Mathematics Subject Classification.* Primary 05A15; Secondary 05A17, 05A19, 05E10, 11B65, 33D15, 33E20.

Key words and phrases. nonintersecting lattice paths, elliptic weights, elliptic hypergeometric series, Frenkel and Turaev's ${}_{10}V_9$ summation, elliptic determinant evaluations.

Partly supported by FWF Austrian Science Fund grants P17563-N13 and S9607 (the second is part of the Austrian National Research Network "Analytic Combinatorics and Probabilistic Number Theory").

the paths in the r-tuple. For any weight function w defined on a set M, we write $w(\mathcal{M};w) := \sum_{x\in\mathcal{M}} w(x)$ for the generating function of the set M with respect to the weight w.

For $\mathbf{u} = (u_1,\dots,u_r)$ and a permutation $\sigma \in \mathcal{S}_r$ we denote $\mathbf{u}_\sigma = (u_{\sigma(1)},\dots,u_{\sigma(r)})$. We say that $\mathbf{u}$ is compatible to $\mathbf{v}$ if no families $(P_1,\dots,P_r)$ of nonintersecting paths from $\mathbf{u}_\sigma$ to $\mathbf{v}$ exist unless $\sigma = \epsilon$, the identity permutation.

We need the following theorem which is a special case (sufficient for the purposes of the present exposition) of the Lindström–Gessel–Viennot theorem of nonintersecting lattice paths (cf. [**12**] and [**10**]).

Theorem 1.1. *Let $\mathbf{u},\mathbf{v} \in (\mathbb{Z}^2)^r$. If $\mathbf{u}$ is compatible to $\mathbf{v}$, then*

$$w(\mathcal{P}_+(\mathbf{u}\to\mathbf{v});w) = \det_{1\le i,j\le r} w(\mathcal{P}(u_j\to v_i)). \tag{1.1}$$

1.2. Ordinary, basic and elliptic hypergeometric series. For the following material, we refer to Gasper and Rahman's texts [**8**]. For motivation, we first define (ordinary) hypergeometric series and basic hypergeometric series, and only then elliptic hypergeometric series, although we will mainly be interested in the latter type of series (being the most general of the three).

For any (complex) parameter a and nonnegative integer k, the *shifted factorial* is defined as

$$(a)_k := a(a+1)\cdots(a+k-1).$$

(This definition can also be extended to the case where k is a negative integer.) It is convenient to use the compact notation

$$(a_1,\dots,a_m)_k := (a_1)_k\cdots(a_m)_k,$$

for products of shifted factorials.

We call a series $\sum c_k$ a *hypergeometric series* if $g(k) = c_{k+1}/c_k$ is a rational function of k. Without loss of generality, we may assume that

$$\frac{c_{k+1}}{c_k} = \frac{(a_1+k)(a_2+k)\dots(a_r+k)}{(1+k)(b_1+k)\dots(b_s+k)}z.$$

The general form of a hypergeometric series is thus

$${}_rF_s\left[\begin{matrix}a_1,\dots,a_r\\ b_1,\dots,b_s\end{matrix};z\right] := \sum_{k=0}^\infty \frac{(a_1,a_2,\dots,a_r)_k}{(1,b_1\dots,b_s)_k}z^k,$$

where $a_1,\dots,a_r$ are the upper parameters, $b_1,\dots,b_s$ the lower parameters, and z is the argument of the series. Several important summation theorems for hypergeometric series include the binomial theorem, the Chu–Vandermonde summation, the Gauß summation, the Pfaff–Saalschütz summation and Dougall's very-well-poised ${}_7F_6$ summation, to name a few.

Now consider q to be a complex parameter, called the "base", usually with $0<|q|<1$. For a nonnegative integer k, the *q-shifted factorial* is defined as

$$(a;q)_k := (1-a)(1-aq)\cdots(1-aq^{k-1}).$$

(This definition can also be extended to the case where k is a negative integer.) It is convenient to use the compact notation

$$(a_1,\dots,a_m;q)_k := (a_1;q)_k\cdots(a_m;q)_k,$$

for products of q-shifted factorials. Note that

$$\lim_{q\to1^-}\frac{(q^a;q)_k}{(1-q)^k} = (a)_k.$$

In this sense the q-shifted factorials generalize the (ordinary) shifted factorials.

We call a series $\sum c_k$ a *q-hypergeometric* or *basic hypergeometric series* if $g(k) = c_{k+1}/c_k$ is a rational function of q^k. Without loss of generality, we may assume that

$$\frac{c_{k+1}}{c_k} = \frac{(1-a_1q^k)(1-a_2q^k)\dots(1-a_rq^k)}{(1-q^k)(1-b_1q^k)\dots(1-b_sq^k)}(-q^k)^{1+s-r}z.$$

The general form of a basic hypergeometric series is thus

$${}_r\phi_s\left[\begin{matrix}a_1,\dots,a_r\\ b_1,\dots,b_s\end{matrix};q,z\right] := \sum_{k=0}^\infty \frac{(a_1,a_2,\dots,a_r;q)_k}{(q,b_1\dots,b_s;q)_k}\left((-1)^kq^{\binom{k}{2}}\right)^{1+s-r}z^k,$$

where $a_1,\dots,a_r$ are the upper parameters, $b_1,\dots,b_s$ the lower parameters, q is the base, and z is the argument of the series. Several important summation theorems for basic hypergeometric series include the q-binomial theorem, the q-Chu–Vandermonde summation, the q-Gauß summation, the q-Pfaff–Saalschütz summation and Jackson's very-well-poised ${}_8\phi_7$ summation, to name a few.

For the elliptic case, define a modified Jacobi theta function with argument x and nome p by

$$\theta(x;p)=(x,p/x;p)_\infty=(x;p)_\infty(p/x;p)_\infty\,,\qquad \theta(x_1,\dots,x_m;p)=\theta(x_1;p)\dots\theta(x_m;p), \tag{1.2}$$

where $x,x_1,\dots,x_m\neq 0,\ |p|<1$, and $(x;p)_\infty=\prod_{k=0}^\infty(1-xp^k)$. We note the following useful properties of theta functions:

$$\theta(x;p)=-x\,\theta(1/x;p),\qquad\qquad \theta(px;p)=-\frac{1}{x}\,\theta(x;p), \tag{1.3}$$

and Riemann's *addition formula*

$$\theta(xy,x/y,uv,u/v;p)-\theta(xv,x/v,uy,u/y;p)=\frac{u}{y}\,\theta(yv,y/v,xu,x/u;p) \tag{1.4}$$

(cf. [**24**, p. 451, Example 5]).

Further, define a *theta shifted factorial* analogue of the q-shifted factorial by

$$(a;q,p)_n=\begin{cases}\prod_{k=0}^{n-1}\theta(aq^k;p), & n=1,2,\dots\,,\\ 1, & \mathrm{n}=0,\\ 1/\prod_{k=0}^{-n-1}\theta(aq^{n+k};p), & n=-1,-2,\dots,\end{cases} \tag{1.5}$$

and let

$$(a_1,a_2,\dots,a_m;q,p)_n=(a_1;q,p)_n\dots(a_m;q,p)_n,$$

where $a,a_1,\dots,a_m\neq 0$. Notice that $\theta(x;0)=1-x$ and, hence, $(a;q,0)_n=(a;q)_n$ is a *q-shifted factorial* in base q. The parameters q and p in $(a;q,p)_n$ are called the *base* and *nome*, respectively, and $(a;q,p)_n$ is called the *q,p-shifted factorial.* Observe that

$$(pa;q,p)_n=(-1)^na^{-n}q^{-\binom{n}{2}}\,(a;q,p)_n, \tag{1.6}$$

which follows from (1.3). A list of other useful identities for manipulating the q,p-shifted factorials is given in [**8**, Sec. 11.2].

We call a series $\sum c_k$ an *elliptic hypergeometric series* if $g(k)=c_{k+1}/c_k$ is an elliptic function of k with k considered as a complex variable; i.e., the function $g(x)$ is a doubly periodic meromorphic function of the complex variable x. Without loss of generality, by the theory of theta functions, we may assume that

$$g(x)=\frac{\theta(a_1q^x,a_2q^x,\dots,a_{s+1}q^x;p)}{\theta(q^{1+x},b_1q^x,\dots,b_sq^x;p)}\,z,$$

where the *elliptic balancing condition*, namely

$$a_1a_2\cdots a_{s+1}=qb_1b_2\cdots b_s,$$

holds. If we write $q=e^{2\pi i\sigma}$, $p=e^{2\pi i\tau}$, with complex σ, τ, then $g(x)$ is indeed periodic in x with periods σ^{-1} and $\tau\sigma^{-1}$.

The general form of an elliptic hypergeometric series is thus

$${}_{s+1}E_s\left[\begin{matrix}a_1,\dots,a_{s+1}\\ b_1,\dots,b_s\end{matrix};q,p;z\right]:=\sum_{k=0}^\infty\frac{(a_1,a_2,\dots,a_{s+1};q,p)_k}{(q,b_1\dots,b_s;q,p)_k}z^k,$$

provided $a_1a_2\cdots a_{s+1}=qb_1b_2\cdots b_s$. Here $a_1,\dots,a_r$ are the upper parameters, $b_1,\dots,b_s$ the lower parameters, q is the base, p the nome, and z is the argument of the series. For convergence reasons, one usually requires $a_{s+1}=q^{-n}$ (n being a nonnegative integer), so that the sum is in fact finite.

Very-well-poised elliptic hypergeometric series are defined as

$$\begin{aligned}{}_{s+1}V_s(a_1;a_6,\dots,a_{s+1};q,p;z)&:={}_{s+1}E_s\left[\begin{matrix}a_1,\ qa_1^{\frac12},-qa_1^{\frac12},qa_1^{\frac12}/p^{\frac12},-qa_1^{\frac12}p^{\frac12},a_6,\dots,a_{s+1}\\ a_1^{\frac12},-a_1^{\frac12},a_1^{\frac12}p^{\frac12},-a_1^{\frac12}/p^{\frac12},a_1q/a_6,\dots,a_1q/a_{s+1}\end{matrix};q,p;-z\right]\\ &=\sum_{k=0}^\infty\frac{\theta(a_1q^{2k};p)}{\theta(a_1;p)}\frac{(a_1,a_6,\dots,a_{s+1};q,p)_k}{(q,a_1q/a_6,\dots,a_1q/a_{s+1};q,p)_k}(qz)^k,\end{aligned} \tag{1.7}$$

where

$$q^2a_6^2a_7^2\cdots a_{s+1}^2=(a_1q)^{s-5}.$$

It is convenient to abbreviate

$${}_{s+1}V_s(a_1;a_6,\dots,a_{s+1};q,p):={}_{s+1}V_s(a_1;a_6,\dots,a_{s+1};q,p;1).$$

Note that in (1.7) we have used

$$\frac{\theta(aq^{2k};p)}{\theta(a;p)}=\frac{(qa^{\frac12},-qa^{\frac12},qa^{\frac12}/p^{\frac12},-qa^{\frac12}p^{\frac12};q,p)_k}{(a^{\frac12},-a^{\frac12},a^{\frac12}p^{\frac12},-a^{\frac12}/p^{\frac12};q,p)_k}(-q)^{-k},$$

which shows that in the elliptic case the number of pairs of numerator and denominator paramters involved in the construction of the *very-well-poised term* is *four* (whereas in the basic case this number is *two*, in the ordinary case only *one*).

The above definitions for ${}_{s+1}E_s$ and ${}_{s+1}V_s$ series are due to Spiridonov [**20**], see [**8**, Ch. 11].

In their study of elliptic $6j$ symbols (which are elliptic solutions of the Yang–Baxter equation found by Baxter [**2**] and Date et al. [**6**]), Frenkel and Turaev [**7**] came across the following ${}_{12}V_{11}$ transformation:

$$\begin{aligned}(1.8)\quad &{}_{12}V_{11}(a;b,c,d,e,f,\lambda aq^{n+1}/ef,q^{-n};q,p)\\ &=\frac{(aq,aq/ef,\lambda q/e,\lambda q/f;q,p)_n}{(aq/e,aq/f,\lambda q/ef,\lambda q;q,p)_n}\,{}_{12}V_{11}(\lambda;\lambda b/a,\lambda c/a,\lambda d/a,e,f,\lambda aq^{n+1}/ef,q^{-n};q,p),\end{aligned}$$

where $\lambda=a^2q/bcd$. This is an extension of Bailey's very-well-poised ${}_{10}\phi_9$ transformation [**8**, Eq. (2.9.1)], to which it reduces when $p=0$.

The ${}_{12}V_{11}$ transformation in (1.8) appeared as a consequence of the tetrahedral symmetry of the elliptic $6j$ symbols. Frenkel and Turaev's transformation contains as a special case the following summation formula,

$$\text{(1.9)}\qquad {}_{10}V_9(a;b,c,d,e,q^{-n};q,p)=\frac{(aq,aq/bc,aq/bd,aq/cd;q,p)_n}{(aq/b,aq/c,aq/d,aq/bcd;q,p)_n},$$

where $a^2q^{n+1}=bcde$, see also (2.14). The ${}_{10}V_9$ summation is an elliptic analogue of Jackson's ${}_8\phi_7$ summation formula [**8**, Eq. (2.6.2)] (or of Dougall's ${}_7F_6$ summation formula [**8**, Eq. (2.1.6)]). A striking feature of elliptic hypergeometric series is that already the simplest identities involve many parameters. The fundamental identity at the "bottom" of the hierarchy of identities for elliptic hypergeometric series is the ${}_{10}V_9$ summation. When keeping the nome p arbitrary (while $|p|<1$) there is no way to specialize (for the sake of obtaining lower order identities) any of the free parameters of an elliptic hypergeometric series in form of a limit tending to zero or infinity, due to the issue of convergence. For the same reason, elliptic hypergeometric series are only well-defined as complex functions if they are terminating (i.e., the sums are finite). See Gasper and Rahman's texts [**8**, Ch. 11] for more details.

2. Elliptic enumeration of lattice paths

The identity responsible for q-calculus to "work" is the simple factorization

$$\text{(2.1)}\qquad q^k-q^{k+1}=(1-q)q^k.$$

This (almost embarrassingly simple) identity underlies not only q-integration (cf. [**1**, Eq. (2.12)]), but also the recursion(s) for the q-binomial coefficient (see (2.8) at the end of this section). As q-binomial coefficients can be combinatorially interpreted as generating functions of lattice paths in $\mathbb{Z}^2$ (from a given starting point to a given end point), one may wonder whether any suitable generalization of (2.1) would give rise to a corresponding extension of q-binomial coefficients with meaningful combinatorial interpretation. Indeed, by using the much more general identity (1.4), rather than (2.1), as the underlying three term relation, we obtain such an extension. In particular, we shall be considering *elliptic binomial coefficients*, resulting from the enumeration of lattice paths with respect to *elliptic weights*. The expressions and series occurring in our study belong to the world of elliptic hypergeometric series, which we just introduced in the previous section.

The most important ingredient for this analysis to work out is the particular "clever" choice of weight function in (2.2). This choice was made, on one hand, by matching the general indefinite sum (2.9) with the known indefinite sum in (2.11), such that induction can be applied (with appeal to the three term relation (1.4), actually a special case of (2.11)). One the other hand, factorization of the elliptic binomial coefficient $w(\mathcal{P}((l,k)\to(n,m)))$ was sought in general, in particular also when $(l,k)\neq(0,0)$. Once the right choice of weight function is made, everything becomes easy and a matter of pure verification. Nevertheless, at the conceptual level things remain interesting. For instance, the elliptic binomial coefficient $w(\mathcal{P}((l,k) \to (n,m)))$ indeed depends on l,k,n,m (besides other parameters), and is not a mere multiple of $w(\mathcal{P}((0,0) \to (n-l,m-k)))$, contrary to the basic ("q") or classical case.

Let a,b,q,p be arbitrary (complex) parameters with $a,b,q \neq 0$ and $|p|<1$. We define the ("standard") *elliptic weight function* on horizontal edges $(n-1,m) \to (n,m)$ of $\mathbb{Z}^2$ as follows.

$$w(n,m) = w(n,m;a,b;q,p) := \frac{\theta(aq^{n+2m}, bq^{2n}, bq^{2n-1}, aq^{1-n}/b, aq^{-n}/b;p)}{\theta(aq^n, bq^{2n+m}, bq^{2n+m-1}, aq^{1+m-n}/b, aq^{m-n}/b;p)}q^m. \tag{2.2}$$

Our terminology is perfectly justified as the weight function defined in (2.2) is indeed *elliptic* (i.e., doubly periodic meromorphic), even independently in each $\log_q a$, $\log_q b$, n and m (viewed as complex parameters). If we write $q=e^{2\pi i\sigma}$, $p=e^{2\pi i\tau}$, $a=q^\alpha$ and $b=q^\beta$ with complex σ, τ, α and β, then the weight $w(n,m)$ is clearly periodic in α with period σ^{-1}. A simple calculation involving (1.6) further shows that $w(n,m)$ is also periodic in α with period $\tau\sigma^{-1}$ (the latter means that $w(n,m)$ is invariant with respect to $a \mapsto pa$). The same applies to $w(n,m)$ viewed as a function in β (or n or m) with the same two periods σ^{-1} and $\tau\sigma^{-1}$. Spiridonov [**20**] calls expressions such as (2.2) where all free parameters have equal periods of double periodicity *totally elliptic*. In this respect we can also refer to (2.2) as a totally elliptic weight.

For $p=0$ (2.2) reduces to

$$w(n,m;a,b;q,0) = \frac{(1-aq^{n+2m})(1-bq^{2n})(1-bq^{2n-1})(1-aq^{1-n}/b)(1-aq^{-n}/b)}{(1-aq^n)(1-bq^{2n+m})(1-bq^{2n+m-1})(1-aq^{1+m-n}/b)(1-aq^{m-n}/b)}q^m. \tag{2.3}$$

If we further let $a \to 0$ and then $b \to 0$ (in this order; or take $b \to 0$ and then $a \to \infty$) this reduces to the standard q-weight q^m (counting the height of, or the area below, the horizontal edge $(n-1,m) \to (n,m)$).

By an *elliptic generating function* we mean, of course, a generating function with respect to an elliptic weight function (and in particular, we shall always take the weight defined in (2.2) unless stated otherwise). It is clear that an elliptic generating function is elliptic as a function in its free parameters.

The particular choice of our elliptic weight in (2.2) is justified by the following nice result.

THEOREM 2.1. *Let l,k,n,m be four integers with $n-l+m-k \geq 0$. The elliptic generating function of paths running from (l,k) to (n,m) is*

$$\begin{aligned} w(\mathcal{P}((l,k) \to (n,m))) &= \frac{(q^{1+n-l}, aq^{1+n+2k}, bq^{1+n+k+l}, aq^{1+k-n}/b;q,p)_{m-k}}{(q, aq^{1+l+2k}, bq^{1+2n+k}, aq^{1+k-l}/b;q,p)_{m-k}} \\ &\quad \times \frac{(aq^{1+l+2k}, aq^{1-n}/b, aq^{-n}/b;q,p)_{n-l}}{(aq^{1+l}, aq^{1+k-n}/b, aq^{k-n}/b;q,p)_{n-l}} \frac{(bq^{1+2l};q,p)_{2n-2l}}{(bq^{1+k+2l};q,p)_{2n-2l}} q^{(n-l)k}. \end{aligned} \tag{2.4}$$

PROOF. First, if $k>m$ (there is no path in this case), the expression in (2.4) vanishes due to the factor $(q;q,p)_{m-k}^{-1}$. On the other hand, if $m \geq k$ but $l>n$ (again there is no path) the expression vanishes due to the factor $(q^{1+n-l};q,p)_{m-k}$ since $n-l+m-k \geq 0$. We may therefore assume, besides $n-l+m-k \geq 0$, that $n \geq l$ and $m \geq k$. The statement is now readily proved by induction on $n-l+m-k$. For $n=l$ one has $w(\mathcal{P}((l,k) \to (l,m))) = 1$ as desired. For $m=k$ one readily verifies $w(\mathcal{P}((l,k) \to (n,k))) = \prod_{i=l+1}^n w(i,k)$. (In both cases there is just one path.) Next assume $n>l$ and $m>k$. We are done if we can verify the recursion

$$w(\mathcal{P}((l,k) \to (n,m))) = w(\mathcal{P}((l,k) \to (n,m-1))) + w(\mathcal{P}((l,k) \to (n-1,m)))\, w(n,m). \tag{2.5}$$

(The final step of a path is either vertical or horizontal.) However, this reduces to the addition formula (1.4). $\square$

Aside from the recursion (2.5), we also (automatically) have

$$w(\mathcal{P}((l,k) \to (n,m))) = w(\mathcal{P}((l,k+1) \to (n,m))) + w(l+1,k)\, w(\mathcal{P}((l+1,k) \to (n,m))). \tag{2.6}$$

(The first step of a path is either vertical or horizontal.) In the limit $p \to 0$, $a \to 0$, $b \to 0$ (in this order), the recursions (2.5) and (2.6) reduce to

$$\begin{bmatrix} n-l+m-k \\ n-l \end{bmatrix}_q q^{(n-l)k} = \begin{bmatrix} n-l+m-k-1 \\ n-l \end{bmatrix}_q q^{(n-l)k} + \begin{bmatrix} n-l+m-k-1 \\ n-l-1 \end{bmatrix}_q q^{(n-l-1)k+m}$$

and

$$\begin{bmatrix} n-l+m-k \\ n-l \end{bmatrix}_q q^{(n-l)k} = \begin{bmatrix} n-l+m-k-1 \\ n-l \end{bmatrix}_q q^{(n-l)(k+1)} + \begin{bmatrix} n-l+m-k-1 \\ n-l-1 \end{bmatrix}_q q^{(n-l-1)k+k},$$

respectively, where

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{(q;q)_n}{(q;q)_k (q;q)_{n-k}} \tag{2.7}$$

is the *q-binomial coefficient*, defined for nonnegative integers n, k with $n \geq k$. This pair of recursions is of course equivalent to the well-known pair

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q q^{n-k}, \qquad \begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q q^k + \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q. \tag{2.8}$$

We may therefore refer to the factored expression in (2.4) as an *elliptic binomial coefficient* (which should not be confused with the much simpler definition given in [**8**, Eq. (11.2.61)] which is a straightforward theta shifted factorial extension of (2.7) but actually *not* elliptic). In fact, it is not difficult to see that the expression in (2.4) is totally elliptic, i.e. elliptic in each $\log_q a$, $\log_q b$, l, k, n and m (viewed as complex parameters) which again fully justifies the notion "elliptic".

2.1. Immediate consequences. Let us consider the elliptic generating function of lattice paths in $\mathbb{Z}^2$ from $(0,0)$ to (n,m). (In what follows, there is in fact no loss of generality in choosing the starting point to be the origin.) We may distinguish paths by the height of their last step. This gives the simple identity

$$w(\mathcal{P}((0,0) \to (n,m))) = \sum_{k=0}^{m} w(\mathcal{P}((0,0) \to (n-1,k)))\, w(n,k). \tag{2.9}$$

In explicit terms, this is

$$\frac{(q^{1+n}, aq^{1+n}, bq^{1+n}, aq^{1-n}/b; q,p)_m}{(q, aq, bq^{1+2n}, aq/b; q,p)_m} \\ = \sum_{k=0}^{m} \frac{(q^n, aq^n, bq^n, aq^{2-n}/b; q,p)_k}{(q, aq, bq^{2n-1}, aq/b; q,p)_k} \frac{\theta(aq^{n+2k}, bq^{2n}, bq^{2n-1}, aq^{1-n}/b, aq^{-n}/b; p)}{\theta(aq^n, bq^{2n+k}, bq^{2n+k-1}, aq^{1+k-n}/b, aq^{k-n}/b; p)} q^k,$$

which, after simplifying the summand, is

$$\frac{(q^{1+n}, aq^{1+n}, bq^{1+n}, aq^{1-n}/b; q,p)_m}{(q, aq, bq^{1+2n}, aq/b; q,p)_m} = \sum_{k=0}^{m} \frac{\theta(aq^{n+2k}; p)(aq^n, q^n, bq^n, aq^{-n}/b; q,p)_k}{\theta(aq^n; p)(q, aq, aq/b, bq^{1+2n}; q,p)_k} q^k. \tag{2.10}$$

By analytic continuation to replace q^n by an arbitrary complex parameter ((2.10) is true for all $n \geq 0$, etc.; see Warnaar [**23**, Proof of Thms. 4.7–4.9] for a typical application of the identity theorem in the elliptic case) and substitution of variables, one gets the indefinite summation

$$\frac{(aq, bq, cq, aq/bc; q,p)_m}{(q, aq/b, aq/c, bcq; q,p)_m} = \sum_{k=0}^{m} \frac{\theta(aq^{2k}; p)(a, b, c, a/bc; q,p)_k}{\theta(a; p)(q, aq/b, aq/c, bcq; q,p)_k} q^k \tag{2.11}$$

(cf. [**8**, Eq. (11.4.10)]).

More generally, for a fixed l, $1 \leq l \leq n$, we may distinguish paths running from $(0,0)$ to (n,m) by the height k they have when they first reach a point on the vertical line $x = l$ (right after the horizontal step $(l-1,k) \to (l,k)$). This refined enumeration reads, in terms of elliptic generating functions,

$$w(\mathcal{P}((0,0) \to (n,m))) = \sum_{k=0}^{m} w(\mathcal{P}((0,0) \to (l-1,k)))\, w(l,k)\, w(\mathcal{P}((l,k) \to (n,m))). \tag{2.12}$$

Explicitly, this is (after some simplifictions)

$$\frac{(q^{1+n}, aq^{1+l}, bq^{1+n}, aq^{1-l}/b; q,p)_m}{(q^{1+n-l}, aq, bq^{1+n+l}, aq/b; q,p)_m} = \sum_{k=0}^{m} \frac{\theta(aq^{l+2k}; p)(aq^l, bq^l, q^l, aq^{-n}/b, aq^{1+n+m}, q^{-m}; q,p)_k}{\theta(aq^l; p)(q, aq/b, aq, bq^{1+n+l}, q^{l-n-m}, aq^{1+l+m}; q,p)_k} q^k, \tag{2.13}$$

which after analytic continuation (first to replace q^n, then q^l, by complex parameters) and substitution of variables becomes

$$\frac{(aq,aq/bc,aq/bd,aq/cd;q,p)_m}{(aq/b,aq/c,aq/d,aq/bcd;q,p)_m}=\sum_{k=0}^{m}\frac{\theta(aq^{2k};p)(a,b,c,d,a^2q^{1+m}/bcd,q^{-m};q,p)_k}{\theta(a;p)(q,aq/b,aq/c,aq/d,bcdq^{-m}/a,aq^{1+m};q,p)_k}q^k, \tag{2.14}$$

The result is *Frenkel and Turaev's* ${}_{10}V_9$ *summation* ([**7**]; cf. [**8**, Eq. (11.4.1)]), the elliptic extension of Jackson's very-well-poised balanced ${}_8\phi_7$ summation (cf. [**8**, Eq. (2.6.2)]), the latter of which is a q-analogue of Dougall's ${}_7F_6$ summation theorem.

We briefly sketch two other ways how to obtain the ${}_{10}V_9$ sum from Theorem 2.1 by convolution (and analytic continuation). For a fixed k, $1\le k\le m$, we may distinguish paths running from $(0,0)$ to (n,m) by the abscissa l they have when they first reach a point on the horizontal line $y=k$ (right after the vertical step $(l,k-1)\to(l,k)$). This refined enumeration reads, in terms of elliptic generating functions,

$$w(\mathcal{P}((0,0)\to(n,m)))=\sum_{l=0}^{m}w(\mathcal{P}((0,0)\to(l,k-1)))\,w(\mathcal{P}((l,k)\to(n,m))). \tag{2.15}$$

On the other hand, we may also fix an antidiagonal running through $(k,0)$ and $(0,k)$, $0<k<n+m$. We can then distinguish paths running from $(0,0)$ to (n,m) by where they cut the antidiagonal. This refined enumeration reads, in terms of elliptic generating functions,

$$w(\mathcal{P}((0,0)\to(n,m)))=\sum_{l=0}^{\min(k,n)}w(\mathcal{P}((0,0)\to(l,k-l)))\,w(\mathcal{P}((l,k-l)\to(n,m))). \tag{2.16}$$

The last two identities both constitute, when written out explicitly using Theorem 2.1, variants of Frenkel and Turaev's ${}_{10}V_9$ summation (like (2.12)) both of which can be extended to (2.14) by analytic continuation.

2.2. Determinant evaluations and elliptic generating functions for nonintersecting lattice paths. For obtaining explicit results the following determinant evaluation from [**23**, Cor. 5.4] is crucial.

Lemma 2.2 (Warnaar). *Let A, B, C, and $X_1,\dots,X_r$ be indeterminate. Then there holds*

$$\det_{1\le i,j\le r}\left(\frac{(AX_i,AC/X_i;q,p)_{r-j}}{(BX_i,BC/X_i;q,p)_{r-j}}\right)=A^{\binom{r}{2}}q^{\binom{r}{3}}\prod_{1\le i<j\le r}X_j\,\theta(X_i/X_j,C/X_iX_j;p)\prod_{i=1}^{r}\frac{(B/A,ABCq^{2r-2i};q,p)_{i-1}}{(BX_i,BC/X_i;q,p)_{r-1}}.$$

As a consequence of Theorem 1.1 and Lemma 2.2, we have the following explicit formulae which generalize Theorem 2.1:

Proposition 2.1. *(a) Let l,k,n, $m_1,\dots,m_r$ be integers such that $m_1\ge m_2\ge\cdots\ge m_r$ and $n-l+m_i-k\ge0$ for all $i=1,\dots,r$. Then the elliptic generating function for nonintersecting lattice paths with starting points $(l+i,k-i)$ and end points (n,m_i), $i=1,\dots,r$, is*

$$\begin{aligned}\det_{1\le i,j,\le r}&\big(w(\mathcal{P}((l+j,k-j)\to(n,m_i)))\big)\\
&=q^{3\binom{r+1}{3}+\binom{r+2}{3}+r(n-l)k-(n-l)\binom{r+1}{2}-r^2k+\sum_{i=1}^{r}(i-1)m_i}\prod_{1\le i<j\le r}\theta(q^{m_i-m_j},aq^{1+n+m_i+m_j};p)\\
&\quad\times\prod_{i=1}^{r}\frac{(q^{1+n-l-i};q,p)_{m_i-k+i}(aq^{1+n+2k-r-i};q,p)_{m_i-k+i}(aq^{1+l+2k-i};q,p)_{n-l-r}}{(q;q,p)_{m_i-k+r}(aq^{1+l+2k-i};q,p)_{m_i-k+i}(aq^{1+l+i};q,p)_{n-l-i}}\\
&\quad\times\prod_{i=1}^{r}\frac{(bq^{2+n+k+l-i};q,p)_{m_i-k+i}(bq^{1+2l+2i};q,p)_{2n-2l-2i}}{(bq^{1+2n+k-i};q,p)_{m_i-k+i}(bq^{1+2l+k+i};q,p)_{2n-2l-2i}}\\
&\quad\times\prod_{i=1}^{r}\frac{(aq^{1+k-n-i}/b;q,p)_{m_i-k+i}(aq^{1-n}/b,aq^{-n}/b;q,p)_{n-l-i}}{(aq^{k-l-i}/b;q,p)_{m_i-k+i}(aq^{1+k-n-i}/b,aq^{k-n-i}/b;q,p)_{n-l-i}}.\end{aligned} \tag{2.17}$$

(b) Let l,k,m, $n_1,\dots,n_r$, be integers such that $n_1\le n_2\le\cdots\le n_r$ and $n_i-l+m-k\ge0$ for all $i=1,\dots,r$. Then the elliptic generating function for nonintersecting lattice paths with starting points $(l+i,k-i)$ and end points (n_i,m), $i=1,\dots,r$, is

$$\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(n_i,m)))\big)=q^{2\binom{r+1}{3}+(l-k+1)\binom{r+1}{2}-rlk+\sum_{i=1}^r(k-i)n_i}$$
$$\times\prod_{i=1}^r\frac{(q^{n_i-l};q,p)_{m-k+1}(aq^{n_i+2k-i};q,p)_{m-k+1-r+i}(aq^{l+2k};q,p)_{n_i-l-i}}{(q;q,p)_{m-k+i}(aq^{l+2k};q,p)_{m-k+1-r+i}(aq^{1+l+i};q,p)_{n_i-l-i}}$$
$$\times\prod_{1\le i<j\le r}\theta(q^{n_j-n_i},bq^{1+m+n_i+n_j};p)\prod_{i=1}^r\frac{(bq^{1+n_i+k+l};q,p)_{m-k+1}(bq^{1+2l+2i};q,p)_{2n_i-2l-2i}}{(bq^{1+2n_i+k-i};q,p)_{m-k+i}(bq^{1+2l+k+i};q,p)_{2n_i-2l-2i}}$$
$$\times\prod_{i=1}^r\frac{(aq^{k-n_i}/b;q,p)_{m-k+1}(aq^{1-n_i}/b,aq^{-n_i}/b;q,p)_{n_i-l-i}}{(aq^{k-l-i}/b;q,p)_{m-k+1}(aq^{k-n_i}/b;q,p)_{n_i-l+1-2i}(aq^{k-r-n_i}/b;q,p)_{n_i-l+r-2i}}. \tag{2.18}$$

(c) Let $l,k,m,\ n_1,\dots,n_r$ be integers such that $n_1\le n_2\le\cdots\le n_r$ and $m-l-k\ge 0$. Then the elliptic generating function for nonintersecting lattice paths with starting points $(l+i,k-i)$ and end points $(n_i,m-n_i)$, $i=1,\dots,r$, is

$$\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(n_i,m-n_i)))\big)=q^{2\binom{r+1}{3}+(l-k+1)\binom{r+1}{2}-rlk+\sum_{i=1}^r(k-i)n_i}$$
$$\times\prod_{i=1}^r\frac{(q^{n_i-l};q,p)_{m-n_i-k+i}(aq^{n_i+2k-i};q,p)_{m-n_i-k+1}(aq^{l+2k};q,p)_{n_i-l-i}}{(q;q,p)_{m-n_i-k+r}(aq^{l+2k};q,p)_{m-n_i-k+1}(aq^{1+l+i};q,p)_{n_i-l-i}}$$
$$\times\prod_{1\le i<j\le r}\theta(q^{n_j-n_i},aq^{m-n_i-n_j}/b;p)\prod_{i=1}^r\frac{(bq^{1+n_i+k+l};q,p)_{m-n_i-k+i}(bq^{1+2l+2i};q,p)_{2n_i-2l-2i}}{(bq^{1+2n_i+k-i};q,p)_{m-n_i-k+i}(bq^{1+2l+k+i};q,p)_{2n_i-2l-2i}}$$
$$\times\prod_{i=1}^r\frac{(aq^{1+k-n_i-i}/b;q,p)_{m-n_i-k+i}(aq^{1-n_i}/b,aq^{-n_i}/b;q,p)_{n_i-l-i}}{(aq^{k-l-i}/b;q,p)_{m-n_i-k+i}(aq^{1+k-n_i-i}/b;q,p)_{n_i-l-i}(aq^{k-r-n_i}/b;q,p)_{n_i-l+r-2i}}. \tag{2.19}$$

(d) Let $l,n,m,\ k_1,\dots,k_r$ be integers such that $k_1\ge k_2\ge\cdots\ge k_r$ and $n-l+m-k_i\ge 0$ for all $i=1,\dots,r$. Then the elliptic generating function for nonintersecting lattice paths with starting points (l,k_i) and end points $(n+i,m-i)$, $i=1,\dots,r$, is

$$\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l,k_j)\to(n+i,m-i)))\big)=q^{\sum_{i=1}^r(n-l+i)k_i}\prod_{1\le i<j\le r}\theta(q^{k_i-k_j},aq^{l+k_i+k_j};p)$$
$$\times\prod_{i=1}^r\frac{(q^{1+n+i-l};q,p)_{m-k_i-i}(aq^{1+n+2k_i};q,p)_{m-k_i}(aq^{1+l+2k_i};q,p)_{n-l}}{(q;q,p)_{m-k_i-1}(aq^{1+l+2k_i};q,p)_{m-k_i-1}(aq^{1+l};q,p)_{n-l+i}}$$
$$\times\prod_{i=1}^r\frac{(bq^{1+n+k_i+l+r};q,p)_{m-k_i-r-1+i}(bq^{1+2l};q,p)_{2n-2l+2i}}{(bq^{1+2n+k_i};q,p)_{m-k_i+i}(bq^{1+2l+k_i};q,p)_{2n-2l}}$$
$$\times\prod_{i=1}^r\frac{(aq^{1+k_i-n}/b;q,p)_{m-k_i-i-1}(aq^{1-n-i}/b,aq^{-n-i}/b;q,p)_{n-l+i}}{(aq^{1+k_i-l}/b;q,p)_{m-k-i}(aq^{1+k_i-n}/b;q,p)_{n-l}(aq^{k_i-r-n}/b;q,p)_{n-l+r}}. \tag{2.20}$$

(e) Let $k,n,m,\ l_1,\dots,l_r$ be integers such that $l_1\le l_2\le\cdots\le l_r$ and $n-l_i+m-k\ge 0$ for all $i=1,\dots,r$. Then the elliptic generating function for nonintersecting lattice paths with starting points (l_i,k) and end points $(n+i,m-i)$, $i=1,\dots,r$, is

$$\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l_j,k)\to(n+i,m-i)))\big)=q^{(n+r+k)\binom{r}{2}+(n+1)rk-\sum_{i=1}^r(k+i-1)l_i}$$
$$\times\prod_{i=1}^r\frac{(q^{1+n+r-l_i};q,p)_{m-k-r}(aq^{1+n+2k+i};q,p)_{m-k-1}(aq^{1+l_i+2k};q,p)_{n+i-l_i}}{(q;q,p)_{m-k-i}(aq^{1+l_i+2k};q,p)_{m-k-1}(aq^{1+l_i};q,p)_{n+i-l_i}}$$
$$\times\prod_{1\le i<j\le r}\theta(q^{l_j-l_i},bq^{k+l_i+l_j};p)\prod_{i=1}^r\frac{(bq^{1+n+k+r+l_i};q,p)_{m-k-r}(bq^{1+2l_i};q,p)_{2n+2i-2l_i}}{(bq^{1+2n+k+2_i};q,p)_{m-k-i}(bq^{1+k+2l_i};q,p)_{2n+2i-2l_i}}$$
$$\times\prod_{i=1}^r\frac{(aq^{1+k-n-i}/b;q,p)_{m-k-1}(aq^{1-n-i}/b,aq^{-n-i}/b;q,p)_{n+i-l_i}}{(aq^{1+k-l_i}/b;q,p)_{m-k-1}(aq^{1+k-n-i}/b,aq^{k-n-i}/b;q,p)_{n+i-l_i}}. \tag{2.21}$$

(f) Let $k,n,m,\ l_1,\dots,l_r$ be integers such that $l_1\le l_2\le\cdots\le l_r$ and $n+m-k\ge 0$. Then the elliptic generating function for nonintersecting lattice paths with starting points $(l_i,k-l_i)$ and end points $(n+i,m-i)$, $i=1,\dots,r$, is

$$\begin{aligned}(2.22)\qquad &\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l_j,k-l_j)\to(n+i,m-i)))\big)=q^{k\binom{r+1}{2}+rnk-\sum_{i=1}^r(n+k+i-l_i)l_i}\\ &\times\prod_{i=1}^r\frac{(q^{1+n+r-l_i};q,p)_{m-k-r+l_i+i-1}(aq^{1+n+2k-2l_i};q,p)_{m-k+l_i}(aq^{1+2k-l_i};q,p)_{n-l_i}}{(q;q,p)_{m-k+l_i-1}(aq^{1+2k-l_i};q,p)_{m-k+l_i-i}(aq^{1+l_i};q,p)_{n+i-l_i}}\\ &\times\prod_{1\le i<j\le r}\theta(q^{l_j-l_i},aq^{k-l_i-l_j}/b;p)\prod_{i=1}^r\frac{(bq^{1+n+k+i};q,p)_{m-k+l_i-i}(bq^{1+2l_i};q,p)_{2n+2i-2l_i}}{(bq^{1+2n+k-l_i};q,p)_{m-k+l_i+i}(bq^{1+k+l_i};q,p)_{2n-2l_i}}\\ &\times\prod_{i=1}^r\frac{(aq^{1+k-n-l_i}/b;q,p)_{m-k+l_i-i-1}(aq^{1-n-i}/b,aq^{-n-i}/b;q,p)_{n+i-l_i}}{(aq^{1+k-2l_i}/b;q,p)_{m-k+l_i-1}(aq^{1+k-n-l_i}/b;q,p)_{n-l_i}(aq^{k-n-r-l_i}/b;q,p)_{n+r-l_i}}.\end{aligned}$$

Remark 2.3. In Proposition 2.1 we are considering generating functions for families of nonintersecting lattice paths where the set of starting points or end points are consecutive points on an antidiagonal parallel to $x+y=c$, for an integer c, such as $(l+i,c-l-i)$. What happens if, say, the starting points are instead considered to be consecutive points on a *horizontal* (resp. *vertical*) line, such as $(l+i,k)$ (resp. $(l,k-i)$), $i=1,\dots,r$? The answer is that the computation of the generating function is then readily reduced to the previous case where the starting points are consecutive points on an antidiagonal, namely $(l+i,k+r-i)$ (resp. $(l+i-1,k-i)$), $i=1,\dots,r$. (We thank Christian Krattenthaler for reminding us of this simple fact; during the preparations of this paper, we had namely computed these other determinants separately and were originally planning to include them explicitly in the above list). In fact, it is easy to see that in this case the second rightmost (resp. second highest) path must start with a vertical (resp. horizontal) step, the third rightmost (resp. third highest) path with two vertical (resp. horizontal) steps, and the leftmost (resp. lowest) path with $r-1$ vertical (resp. horizontal) steps. Explicitly, we have

$$(2.23)\qquad \det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k)\to(n_i,m_i)))\big)=\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k+r-j)\to(n_i,m_i)))\big),$$

and

(2.24)

$$\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l,k-j)\to(n_i,m_i)))\big)=\prod_{1\le i<j\le r}w(l+i,k-j)\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j-1,k-j)\to(n_i,m_i)))\big).$$

An analogous fact holds if one considers the *end* points instead of the starting points to be consecutive on a horizontal (resp. vertical) line.

3. Identities for multiple elliptic hypergeometric series

It is straightforward to extend the convolution formulae in (2.12), (2.15), and (2.16), to the multivariate setting using the interpretation of nonintersecting lattice paths. We have the following identities:

Proposition 3.1. *Let l,k,n,m be integers such that $n-l+m-k\ge 0$.*
(a) Fix an integer ν such that $l+r+1\le\nu\le n+1$. Then we have

$$\begin{aligned}(3.1)\qquad &\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(n+i,m-i)))\big)\\ &=\sum_{\substack{t_1>t_2>\cdots>t_r\\ t_1\le m-1,t_r\ge k-r}}\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(\nu-1,t_i)))\big)\prod_{s=1}^r w(\nu,t_s)\det_{1\le i,j,\le r}\big(w(\mathcal{P}((\nu,t_j)\to(n+i,m-i)))\big).\end{aligned}$$

(b) Fix an integer ν such that $k\le\nu\le m-r$. Then we have

$$\begin{aligned}(3.2)\qquad &\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(n+i,m-i)))\big)\\ &=\sum_{\substack{t_1<t_2<\cdots<t_r\\ t_1\ge l+1,t_r\le n+r}}\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(t_i,\nu-1)))\big)\det_{1\le i,j,\le r}\big(w(\mathcal{P}((t_j,\nu)\to(n+i,m-i)))\big).\end{aligned}$$

(c) Fix an integer ν such that $l+k\le\nu\le n+m$. Then we have

$$\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(n+i,m-i)))\big)\tag{3.3}$$
$$=\sum_{\substack{t_1<t_2<\dots<t_r\\ t_1\ge l+1,t_r\le n+r}}\det_{1\le i,j,\le r}\big(w(\mathcal{P}((l+j,k-j)\to(t_i,\nu-t_i)))\big)\det_{1\le i,j,\le r}\big(w(\mathcal{P}((t_j,\nu-t_j)\to(n+i,m-i)))\big).$$

We could also have formulated more general versions of convolutions where the respective starting and/or end points of the total paths are not consecutive on antidiagonals (in the above cases these points are $(l+i,k-i)$ and $(n+i,m-i)$, $i=1,\dots,r$). However, the advantage of our specific choice is that all the determinants involved in Proposition 3.1 factor into closed form, by virtue of the determinant evaluations in Proposition 2.1. We thus obtain, writing out the identities (3.1), (3.2), and (3.3) explicitly, summations which are particularly attractive since both the summands and the product sides are completely factored. Each of the above three cases leads, after suitable substitution of variables, simplification, and analytic continuation, to the same result. It is a special case of a multivariate ${}_{10}V_9$ summation formula conjectured by Warnaar (let $x=q$ in [**23**, Cor. 6.2]) which has subsequently been proved by Rosengren [**16**].

Theorem 3.1 (A multivariate extension of Frenkel and Turaev's ${}_{10}V_9$ summation formula). *Let a,b,c,d be indeterminates, let m be a nonnegative integer, and $r\ge 1$. Then we have*

$$\sum_{0\le k_1<k_2<\dots<k_r\le m} q^{\sum_{i=1}^r(2i-1)\lambda_i}\prod_{1\le i<j\le r}\theta(q^{k_i-k_j},aq^{k_i+k_j};p)^2\tag{3.4}$$
$$\times\prod_{i=1}^r\frac{\theta(aq^{2k_i};p)(a,b,c,d,a^2q^{3-2r+m}/bcd,q^{-m};q,p)_{k_i}}{\theta(a;p)(q,aq/b,aq/c,aq/d,bcdq^{2r-2-m}/a,aq^{1+m};q,p)_{k_i}}$$
$$=q^{-4\binom{r}{3}}\left(\frac{a}{bcdq}\right)^{\binom{r}{2}}\prod_{i=1}^r(q,b,c,d,a^2q^{3-2r+m}/bcd;q,p)_{i-1}$$
$$\times\prod_{i=1}^r\frac{(q,aq;q,p)_m(aq^{2-i}/bc,aq^{2-i}/bd,aq^{2-i}/cd;q,p)_{m+1-r}}{(q,aq/b,aq/c,aq/d,aq^{2-2r+i}/bcd;q,p)_{m+1-i}}.$$

Note that the Vandermonde determinant-like factor appearing in the summand of (3.4) is squared. This distinctive feature is reminiscent of certain Schur function and multiple q-series identities with similar property (which can also be proved by the machinery of nonintersecting lattice paths), see e.g. [**11**, Thms. 5 and 6] and [**3**, Thms. 27–29].

The following result is the natural generalization of Theorem 3.1 to the higher level of transformations. It is a special case of a multivariate ${}_{12}V_{11}$ transformation formula conjectured by Warnaar (let $x=q$ in [**23**, Conj. 6.1]) which has subsequently been proved (in more generality) by Rains [**15**] and, independently, by Coskun and Gustafson [**5**].

Theorem 3.2 (A multivariate extension of Frenkel and Turaev's ${}_{12}V_{11}$ transformation formula). *Let a,b,c,d,e,f be indeterminates, let m be a nonnegative integer, and $r\ge 1$. Then we have*

$$\sum_{0\le k_1<k_2<\dots<k_r\le m} q^{\sum_{i=1}^r(2i-1)k_i}\prod_{1\le i<j\le r}\theta(q^{k_i-k_j},aq^{k_i+k_j};p)^2\tag{3.5}$$
$$\times\prod_{i=1}^r\frac{\theta(aq^{2k_i};p)(a,b,c,d,e,f,\lambda aq^{2-r+m}/ef,q^{-m};q,p)_{k_i}}{\theta(a;p)(q,aq/b,aq/c,aq/d,aq/e,aq/f,efq^{r-1-m}/\lambda,aq^{1+m};q,p)_{k_i}}$$
$$=\prod_{i=1}^r\frac{(b,c,d,ef/a;q,p)_{i-1}(aq;q,p)_m(aq/ef;q,p)_{m+1-r}(\lambda q/e,\lambda q/f;q,p)_{m+1-i}}{(\lambda b/a,\lambda c/a,\lambda d/a,ef/\lambda;q,p)_{i-1}(\lambda q;q,p)_m(\lambda q/ef;q,p)_{m+1-r}(aq/e,aq/f;q,p)_{m+1-i}}$$
$$\times\sum_{0\le k_1<k_2<\dots<k_r\le m} q^{\sum_{i=1}^r(2i-1)k_i}\prod_{1\le i<j\le r}\theta(q^{k_i-k_j},\lambda q^{k_i+k_j};p)^2$$
$$\times\prod_{i=1}^r\frac{\theta(\lambda q^{2k_i};p)(\lambda,\lambda b/a,\lambda c/a,\lambda d/a,e,f,\lambda aq^{2-r+m}/ef,q^{-m};q,p)_{k_i}}{\theta(\lambda;p)(q,aq/b,aq/c,aq/d,\lambda q/e,\lambda q/f,efq^{r-1-m}/\lambda,\lambda q^{1+m};q,p)_{k_i}},$$

where $\lambda=a^2q^{2-r}/bcd$.

The $r = 1$ case of Theorem 3.2 is Frenkel and Turaev's ${}_{12}V_{11}$ transformation theorem [**7**], an elliptic extension of Bailey's ${}_{10}\phi_9$ transformation [**8**, Eq. (2.9.1)]. Again, the Vandermonde determinant-like factor appearing in the summand of (3.5) is squared. (Similar identities but with a simple Vandermonde determinant-like factor appearing in the summand have been derived in [**17**].) Due to symmetry the range of summations on both sides of (3.5) can also be taken over all integers $0 \le k_1, \dots, k_r \le m$. If we let $c = aq/b$ in (3.5), the left-hand side reduces to a multivariate ${}_{10}V_9$ series. On the right-hand side, since $\lambda d/a = q^{1-r}$, the sum boils down to just a single term, with the indices $k_i = i-1$, $1 \le i \le r$. The result, after simplifications, is of course Theorem 3.1.

It would be particularly interesting to find a combinatorial proof of (3.5) involving nonintersecting lattice paths. Even for $r = 1$ we so far failed to find a lattice path proof.

4. Future perspectives

4.1. Tableaux and plane partitions. It is quite clear how one can enumerate objects such as tableaux or (various classes of) plane partitions with respect to elliptic weights. First, one has to translate the respective combinatorial objects via a standard bijection into a set of nonintersecting lattice paths (see [**10**] or [**21**]). The translation back, in order to obtain an explicit definition for the weight of the corresponding combinatorial object, is not difficult. In the simplest cases the elliptic generating function is then expressed, by Theorem 1.1, as a determinant which may be computed by Proposition 2.1. If the starting and/or end points of the lattice paths are not fixed, one applies instead of Theorem 1.1 a result by Okada [**14**] (see also Stembridge [**21**]), which expresses the generating function as a Pfaffian. Since the square of a Pfaffian is a determinant of a skew symmetric matrix, this again involves the computation of a determinant. It needs to be explored which of the classical results can be extended to the elliptic setting. Some elliptic determinant evaluations, other than Warnaar's in Lemma 2.2, which might be useful in this context have been provided by Rosengren and present author [**18**].

4.2. Elliptic Schur functions. One can replace (2.2) by the more general weight

$$w(x;n,m) := \frac{\theta(ax_m^2q^n, bq^{2n}, bq^{2n-1}, aq^{1-n}/b, aq^{-n}/b;p)}{\theta(aq^n, bx_mq^{2n}, bx_mq^{2n-1}, ax_mq^{1-n}/b, ax_mq^{-n}/b;p)} x_m \tag{4.1}$$

(defined on horizontal steps $(n-1,m) \to (n,m)$ of $\mathbb{Z}^2$), and enumerate nonintersecting lattice paths, corresponding to tableaux, with respect to (4.1). The result is an elliptic extension of Schur functions (which may no longer be orthogonal) which, when "principally specialized" ($x_i \mapsto q^i$, $i \ge 0$) factors into closed form in view of Proposition 2.1. On one hand it should be investigated whether these elliptic Schur functions have other nice properties (as they do have in the classical case, see [**13**]). It appears that they are *not* related to any of the BC-symmetric functions considered in [**5**] or [**15**].

4.3. Other weight functions. We were able to disguise Frenkel and Turaev's ${}_{10}V_9$ summation formula as a convolution identity of elliptic binomial coefficients (see also Rains [**15**, Sec. 4] and Coskun and Gustafson [**5**]). In our case this involved lattice paths with respect to elliptic weights. Similarly, it should also be feasible to reproduce other known convolution formulae (such as Abel's generalization of the binomial theorem or the Hagen–Rothe summation, cf. [**19**], or others) using lattice paths with appropriately chosen weights. The three types of convolutions, displayed in (2.12), (2.15), and (2.16), still hold, but may then lead to mutually different identities. One can also try to work with *bibasic* weights (either elliptic or non-elliptic), in order to recover some of the identities in [**8**, Secs. 3.6 and 3.8] and in [**23**]. It seems likely that in the non-elliptic case (here we mean that there is no nome p, or $p = 0$) Bill Gosper used exactly this method to first derive his "strange evaluations" (which were later subsumed/generalized in [**8**, Secs. 3.6 and 3.8]). Of course, whatever identities or other results one obtains by lattice path interpretation, one can check for possible related determinant evaluations. Also the other direction should be investigated, e.g. does Warnaar's quadratic elliptic determinant in [**23**, Thm. 4.17] correspond to a specific set of nonintersecting lattice paths with quadratic elliptic weight function?

4.4. "Elliptic" combinatorics. We believe that the results presented in this paper do not stand alone, i.e., that elliptic enumeration is not necessarily restricted to lattice paths. In the same way as the generating functions for various classes of combinatorial objects (most notably, of partitions, which correspond to paths) can be expressed in terms of q-series, closed form elliptic generating functions for several of these classes should exist as well. The main idea would be to replace q-weights by suitable elliptic weights (and then make the further analysis works out). There are certainly restrictions to the elliptic approach (besides that the objects counted should be finite). Already when considering paths in $\mathbb{Z}^2$, techniques involving André's reflection principle (cf. [**4**, p. 22]) or shifting paths (as in [**9**, Prop. 1]) are not applicable as they are not anymore weight invariant. A good area where to look for elliptic extensions would be a general combinatorial theory such as Viennot's theory of heaps [**22**].

References

[1] G. E. Andrews, *q-Series: Their development and application in analysis, number theory, combinatorics, physics and computer algebra*, CBMS Regional Conference Lectures Series **66**, Amer. Math. Soc., Providence, RI, 1986.

[2] R. J. Baxter, "Eight-vertex model in lattice statistics and one-dimensional anisotropic Heisenberg chain, II: Equivalence to a generalized ice-type model", *Ann. Phys.* **76** (1973), 193–228.

[3] M. Ciucu, T. Eisenkölbl, C. Krattenthaler, and D. Zare, "Enumeration of lozenge tilings of hexagons with a central triangular hole", *J. Combin. Theory Ser. A* **95** (2001), 251–334.

[4] L. Comtet, *Advanced combinatorics*, D. Reidel, Dordrecht, Holland, 1974.

[5] H. Coskun and R. A. Gustafson, "Well-poised Macdonald functions W_λ and Jackson coefficients ω_λ on BC_n", Proceedings of the workshop on Jack, Hall–Littlewood and Macdonald polynomials, *Contemp. Math.*, to appear.

[6] E. Date, M. Jimbo, A. Kuniba, T. Miwa, M. Okado, "Exactly solvable SOS models: local height probabilities and theta function identities", *Nuclear Phys. B* **290** (1987), 231–273.

[7] I. B. Frenkel and V. G. Turaev, "Elliptic solutions of the Yang–Baxter equation and modular hypergeometric functions", in V. I. Arnold et al. (eds.), *The Arnold–Gelfand Mathematical Seminars*, 171–204, Birkhäuser, Boston, 1997.

[8] G. Gasper and M. Rahman, *Basic hypergeometric series*, 2^{nd} ed., Encyclopedia of Mathematics And Its Applications 96, Cambridge University Press, Cambridge, 2004.

[9] I. M. Gessel and C. Krattenthaler, "Cylindric partitions", *Trans. Amer. Math. Soc.* **349** (2) (1997), 429–479.

[10] I. M. Gessel and X. Viennot, "Binomial determinants, paths, and hook length formulae", *Adv. Math.* **58** (1985), 300–321.

[11] C. Krattenthaler, "Schur function identities and the number of perfect matchings of holey Aztec rectangles", *Contemp. Math.* **254** (2000), 335–335.

[12] B. Lindström, "On the vector representations of induced matroids", *Bull. London Math. Soc.* **5** (1973), 85–90.

[13] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Clarendon Press, second edition, Oxford, 1995.

[14] S. Okada, "On the generating functions for certain classes of plane partitions", *J. Combin. Theory Ser. A* **51** (1989), 1–23.

[15] E. Rains, "BC_n-symmetric abelian functions", *Duke Math. J.*, to appear.

[16] H. Rosengren, "A proof of a multivariable elliptic summation formula conjectured by Warnaar", *Contemp. Math.* **291** (2001), 193–202.

[17] H. Rosengren and M. Schlosser, "Summations and transformations for multiple basic and elliptic hypergeometric series by determinant evaluations", *Indag. Math. (N.S.)* **14** (2003), 483–514.

[18] H. Rosengren and M. Schlosser, "Elliptic determinant evaluations and the Macdonald identities for affine root systems", *Compositio Math.*, to appear.

[19] M. Schlosser, "Some new applications of matrix inversions in A_r", *Ramanujan J.* **3** (1999), 405–461.

[20] V. P. Spiridonov, "Theta hypergeometric series", in V. A. Malyshev and A. M. Vershik (eds.), *Asymptotic Combinatorics with Applications to Mathematical Physics*, 307–327, Kluwer Acad. Publ., Dordrecht, 2002.

[21] J. R. Stembridge, "Nonintersecting paths, pfaffians and plane partitions", *Adv. Math.* **83** (1990), 96–131.

[22] G. X. Viennot, "Heaps of pieces. I: Basic definitions and combinatorial lemmas", in *Combinatoire énumérative*, Proc. Colloq., Montréal, Canada 1985, Lect. Notes Math. **1234** (1986), 321-350.

[23] S. O. Warnaar, "Summation and transformation formulas for elliptic hypergeometric series", *Constr. Approx.* **18** (2002), 479–502.

[24] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed., Cambridge University Press, Cambridge, 1962.

FAKULTÄT FÜR MATHEMATIK, UNIVERSITÄT WIEN, NORDBERGSTRASSE 15, A-1090 VIENNA, AUSTRIA
E-mail address: `michael.schlosser@univie.ac.at`
URL: `http://www.mat.univie.ac.at/~schlosse`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Octahedrons with equally many lattice points and generalizations

Thomas Stoll and Robert F. Tichy

Abstract. While counting lattice points in octahedra of different dimensions n and m, it is an interesting question to ask, how many octahedra exist containing equally many such points. This gives rise to the Diophantine equation $P_n(x) = P_m(y)$ in rational integers x, y, where $\{P_k(x)\}$ denote special Meixner polynomials $\{M_k^{(\beta,c)}(x)\}$ with $\beta = 1$, $c = -1$. We join the purely algebraic criterion of Y. Bilu and R. F. Tichy (*The Diophantine equation* $f(x) = g(y)$, Acta Arith. **95** (2000), no. 3, 261–288) with a famous result of P. Erdös and J. L. Selfridge (*The product of consecutive integers is never a power*, Illinois J. Math. **19** (1975), 292–301) and prove that

$$M_n^{(\beta,c_1)}(x) = M_m^{(\beta,c_2)}(y)$$

with $m, n \geq 3$, $\beta \in \mathbb{Z} \setminus \{0, -1, -2, -\max(n, m) + 1\}$ and $c_1, c_2 \in \mathbb{Q} \setminus \{0, 1\}$ only admits a finite number of integral solutions x, y. Some more results on polynomial families in three-term recurrences are presented.

Résumé. Dans l'étude du dénombrement de sommets d'octaèdres de dimensions n et m se pose la question intéressante de connaître combien d'octaèdres existent possédant le même nombre de sommets. Ce problème se traduit par l'équation diophantienne $P_n(x) = P_m(y)$, avec x, y entiers relatifs et où $\{P_k(x)\}$ sont les polynômes spéciaux de Meixner avec $\beta = 1$, $c = -1$. Nous joignons au critère purement algébrique de Y. Bilu et R. F. Tichy (*The Diophantine equation* $f(x) = g(y)$, Acta Arith. **95** (2000), no. 3, 261–288) un fameux résultat dû à P. Erdös et J. L. Selfridge (*The product of consecutive integers is never a power*, Illinois J. Math. **19** (1975), 292–301) et prouvons que

$$M_n^{(\beta,c_1)}(x) = M_m^{(\beta,c_2)}(y)$$

avec $m, n \geq 3$, $\beta \in \mathbb{Z} \setminus \{0, -1, -2, -\max(n, m) + 1\}$ et $c_1, c_2 \in \mathbb{Q} \setminus \{0, 1\}$ n'admet qu'un nombre fini de solutions entières x, y. De plus, nous présentons quelques résultats portant sur des familles polynômiales avec triple récurrence.

1. Introduction

An n-dimensional octahedron of radius r is the convex body in $\mathbb{R}^n$ defined by $|x_1| + \cdots + |x_n| \leq r$. In this talk we investigate the following problem and some algebraic generalizations:

Problem: *Given distinct positive integers n, m, how often can two octahedrons of dimensions n and m, respectively, contain equally many integral points?*

Obviously, it is sufficient to consider octahedrons of integral radius r. Also, any positive odd number can occur as the number of integers in the "one-dimensional octahedron" $[-r, r]$. Hence, it is natural to assume that $n, m \geq 2$.

Denote by $P_n(r)$ the number of integral points $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ satisfying $|x_1| + \cdots + |x_n| \leq r$. In 1967, Erhardt [**5**] proved that $P_n(r)$ is a polynomial in r of degree n indeed for any general lattice polytope described by

$$\frac{|x_1|}{a_1} + \frac{|x_2|}{a_2} + \cdots + \frac{|x_n|}{a_n} \leq r,$$

2000 *Mathematics Subject Classification.* Primary 11D41; Secondary 11D45, 33C05.
Key words and phrases. counting lattice points, Diophantine equations, Meixner polynomials, orthogonal polynomials.

where $a_1, \ldots, a_n$ are positive integers. In general, the Ehrhart polynomial is difficult to access and its coefficients involve Dedekind sums and their higher analogues [**1**]. However, in the special case of symmetric octahedra, which we are dealing with here, Kirschenhofer, Pethö and Tichy [**10**] could show that $P_n(r)$ can be made explicit, namely

$$P_n(r) = \sum_{i=0}^{n} 2^i \binom{n}{i}\binom{r}{i} = {}_2F_1\left[\begin{matrix} -n, -r \\ 1 \end{matrix}\; ; 2\right], \tag{1.1}$$

where

$${}_2F_1\left[\begin{matrix} a, b \\ c \end{matrix}\; ; z\right] = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k}\, z^k$$

is the Gauss hypergeometric function with $(a)_0 = 1$ and $(a)_k = a(a+1)\ldots(a+k-1)$ the Pochhammer symbol. Thus, the original combinatorial counting problem can be restated by means of a polynomial Diophantine equation:

Problem, restated: *How many solutions $x, y \in \mathbb{Z}$ can the equation $P_n(x) = P_m(y)$ have?*

According to the modern Askey-scheme [**14**] and (1.1), we note that

$$P_k(x) = M_k^{(1,-1)}(x), \tag{1.2}$$

where

$$M_k^{(\beta,c)}(x) = {}_2F_1\left[\begin{matrix} -k, -x \\ \beta \end{matrix}\; ; 1 - \frac{1}{c}\right]$$

denote the well-known *Meixner* polynomials.

2. Historical remarks

Hajdu [**7, 8**] studied the problem for small n and m. For the cases

$$(n, m) \in \{(3,2), (4,2), (6,2), (4,3), (6,4)\}$$

he completely determined all integral solutions of $P_n(x) = P_m(y)$. He also conjectured that the equation has finitely many solutions when $n > m = 2$. This was confirmed by Kirschenhofer, Pethő and Tichy [**10**], who reduced it to the Siegel–Baker theorem about the hyperelliptic equation $y^2 = f(x)$ in order to give a computable bound for integral solutions x, y of the equation $P_n(x) = P_2(y)$. Moreover, finiteness is also shown in the following three cases: $m = 4$; $2 \le m < n \le 103$; $n \not\equiv m \bmod 2$. The two latter results are no longer effective (i.e., no upper bound for x, y can be retrieved from the proof), because they depend on the non-effective Davenport–Lewis–Schinzel [**4**] theorem about the Diophantine equation $f(x) = g(y)$. The general answer to the problem has been obtained in [**2**]:

Theorem 2.1 (Bilu-Stoll-Tichy, 2000). *Let n and m be distinct integers satisfying $m, n \ge 2$. Then the equation*

$$P_n(x) = P_m(y)$$

has only finitely many solutions in rational integers x, y.

In other words, sufficiently large octahedra of distinct dimensions n, m cannot have equally many lattice points. The proof of Theorem 2.1 is based on a non-effective result of Bilu and Tichy [**3**], thus, we cannot make "sufficiently large" more explicit.

3. Generalizations

Several new questions arise in this context. For instance, it is well-known that the general family $\{M_k^{(\beta,c)}(x)\}$ defines a discrete orthogonal polynomial family if and only if $\beta > 0$ and $0 < c < 1$. Since the original case $\beta = 1$, $c = -1$ (see (1.2)) does not fit in, we are interested in a more general statement, which handles both the original and the orthogonal case.

Question 1: *Is it possible to derive a similar result to Theorem 2.1 for more general β and c, including the orthogonal case?*

Furthermore, one may also ask, whether it is possible to replace the family of Meixner polyomials by some other polynomial family $\{p_k(x)\}$. Since orthogonal polynomials are closely related to polynomials in three-term recurrences by Favard's theorem, the following question seems of interest.

Question 2: *Let $\{p_k(x)\}$ be a sequence of polynomials defined by*

$$\begin{aligned} p_0(x) &= 1 \\ p_1(x) &= x - c_0 \\ p_{k+1}(x) &= (x - c_k)p_n(x) - d_k p_{k-1}(x), \qquad k = 1, 2, \dots, \end{aligned} \tag{3.1}$$

where c_k and d_k are parameters depending only on k. For which c_k, d_k the equation $p_n(x) = p_m(y)$ only has finitely many integral solutions x, y?

Note again, that by the Askey scheme, the Meixner polynomials satisfy a normalized recurrence relation with $c_k = (k + (k + \beta)c)/(c - 1)$ and $d_k = (k(k + \beta - 1)c)/(c - 1)^2$.

Diophantine equations of the form $p_m(x) = p_n(y)$ with polynomials in three-term recurrences have been studied recently by Kirschenhofer and Pfeiffer [**11, 12**]. They point out several striking connections to enumeration problems (for instance, to permutations with coloured cycles).

4. Main results

4.1. Concerning 'Question 1'. Question 1 is settled by the following result [**17**]:

Theorem 4.1. *Let n and m be distinct integers satisfying $m, n \geq 3$, further let $c_1, c_2 \in \mathbb{Q} \setminus \{0, 1\}$ and $\beta \in \mathbb{Z} \setminus \{0, -1, -2, -\max(n, m) + 1\}$. Then the equation*

$$M_n^{(\beta,c_1)}(x) = M_m^{(\beta,c_2)}(y)$$

has only finitely many solutions in integers x, y.

Denote by $K_n^{(p,N)}(x)$ the two-parametric *Krawtchouk* polynomials given in [**14**]:

$$K_n^{(p,N)}(x) = {}_2F_1\left[\begin{matrix} -n, -x \\ -N \end{matrix} ; \frac{1}{p}\right] \qquad n = 0, 1, 2, \dots, N.$$

Since

$$K_n^{(p,N)}(x) = M_n^{(-N,p/(p-1))}(x),$$

we also have

Theorem 4.2. *Let n and m be distinct integers satisfying $m, n \geq 3$, further let $N \geq \max(m, n)$ and $p_1, p_2 \in \mathbb{Q} \setminus \{0, 1\}$. Then the equation*

$$K_n^{(p_1,N)}(x) = K_m^{(p_2,N)}(y) \tag{4.1}$$

has only finitely many solutions in integers x, y.

4.2. Concerning 'Question 2'. We obtain sufficient conditions on c_k and d_k in order to state an again more general finiteness theorem [**18**]:

Theorem 4.3. *Let $\{p_k(x)\}$ be a polynomial sequence satisfying (3.1). Assume one of the following conditions ($A, B, C \in \mathbb{Q}$)*

1. *$c_0 = A, \quad c_k = A, \quad d_k = B$ with $A \neq 0$ and $B > 0$,*
2. *$c_0 = A + B, \quad c_k = A, \quad d_k = B^2$ with $B \neq 0$,*
3. *$c_0 = A, \quad c_k = Bk + A, \quad d_k = \frac{1}{4}B^2k^2 + Ck$ with $C > -\frac{1}{4}B^2$.*

Then the Diophantine equation

$$\mathcal{A}p_m(x) + \mathcal{B}p_n(y) = \mathcal{C}$$

with $m > n \geq 4$, $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{Q}$, $\mathcal{AB} \neq 0$ has at most finitely many solutions in rational integers x, y.

Note that, for instance, in case (3) there are the six rational parameters $\mathcal{A}, \mathcal{B}, \mathcal{C}, A, B, C$ involved, thus, the generality of Theorem 4.3 should well fit specific combinatorial applications. Furthermore, well-known orthogonal families are covered by the statement. So, for example, in the first case of Theorem 4.3 we deal with (shifted) *Jacobi* polynomials, while the third case corresponds to modified *Hermite* and *Laguerre* polynomials.

5. Methods and tools

5.1. The Bilu-Tichy method. The proofs of Theorem 4.1, Theorem 4.2 and Theorem 4.3 are basically algebraic, as they are based on an explicit algorithmic criterion of Bilu and Tichy [**3**], which only involves knowledge of the coefficients of the polynomials under consideration. In order to state that result, we have to introduce some more notation.

Let $\gamma, \delta \in \mathbb{Q}\setminus\{0\}$, $q, s, t \in \mathbb{Z}_{>0}$, $r \in \mathbb{Z}_{\geq 0}$ and $v(x) \in \mathbb{Q}[x]$ a non-zero polynomial (which may be constant). Further let $D_s(x,\gamma)$ denote the *Dickson* polynomials which can be defined via

$$D_s(x,\gamma) = \sum_{i=0}^{\lfloor s/2 \rfloor} d_{s,i} x^{s-2i} \qquad \text{with} \qquad d_{s,i} = \frac{s}{s-i}\binom{s-i}{i}(-\gamma)^i.$$

The pair $(f(x), g(x))$ or viceversa $(g(x), f(x))$ is called a *standard pair over* $\mathbb{Q}$ if it can be represented by an explicit form listed below. In such a case we call (f,g) a standard pair of the *first*, *second*, *third*, *fourth*, *fifth kind*, respectively.

kind	explicit form of (f,g) resp. (g,f)	parameter restrictions
first	$(x^q, \gamma x^r v(x)^q)$	with $0 \leq r < q,\ (r,q)=1,\ r + \deg v > 0$
second	$(x^2, (\gamma x^2 + \delta) v(x)^2)$	–
third	$(D_s(x,\gamma^t), D_t(x,\gamma^s))$	with $(s,t)=1$
fourth	$\left(\gamma^{-s/2} D_s(x,\gamma), -\delta^{-t/2} D_t(x,\delta)\right)$	with $(s,t)=2$
fifth	$\left((\gamma x^2 - 1)^3, 3x^4 - 4x^3\right)$	–

These standard pairs are important in view of the following characterization result [**3**].

THEOREM 5.1 (Bilu-Tichy, 2000). *Let $p(x), q(x) \in \mathbb{Q}[x]$ be non-constant polynomials. Then the following two assertions are equivalent:*

(a) *The equation $p(x) = q(y)$ has infinitely many rational solutions with a bounded denominator.*
(b) *We can express $p \circ \kappa_1 = \phi \circ f$ and $q \circ \kappa_2 = \phi \circ g$ where $\kappa_1, \kappa_2 \in \mathbb{Q}[x]$ are linear, $\phi(x) \in \mathbb{Q}[x]$, and (f,g) is a standard pair over $\mathbb{Q}$.*

If we are able to get contradictions for decompositions of p and q as demanded in (b) of Theorem 5.1 then finiteness of number of integral solutions x, y of the original Diophantine equation $p(x) = q(y)$ is guaranteed. Note that this approach is basically an algebraic one and does involve an accurate comparison of coefficients.

5.2. Erdös-Selfrdige tool. As an additional tool, we restate a well-known result obtained by Erdös and Selfridge [**6**]:

THEOREM 5.2 (Erdös-Selfridge, 1975). *The equation*

$$x(x+1)\cdots(x+k-1) = y^l$$

has no solution in rational integers $x > 0$, $k > 1$, $l > 1$, $y > 1$.

Interestingly, simple comparison of the leading coefficients of the *Meixner* polynomials gives an equation very similar to that of Theorem 5.2. Therefore, there are no parameters that satisfy such a coefficient equation. In other words, we can easily derive a contradiction if we suppose a higher degree polynomial representation in Theorem 5.1.

5.3. Lesky tool. There is a close connection beween three-term recurrences and Sturm-Liouville differential equations [**13**]:

THEOREM 5.3 (Koepf-Schmersau, 2002). *The following conditions are equivalent:*

(1) *The second-order Sturm-Liouville differential equation ($k \geq 0$)*

$$\sigma(x) p_k''(x) + \tau(x) p_k'(x) - k((k-1)a + d) p_k(x) = 0, \tag{5.1}$$

with $\sigma(x) = ax^2 + bx + c \not\equiv 0$, $\tau = dx + e$, $a, b, c, d, e \in \mathbb{R}$, $d \neq -ta$ for all $t \in \mathbb{Z}_{\geq 0}$ has a (up to a factor depending on k) unique infinite polynomial family solution $\{p_k(x)\}$ of exact degree k.

(2) *The family* $\{p_k(x)\}$ *satifies a three-term recurrence of type (3.1) with*

$$
\begin{aligned}
c_0 &= -\frac{e}{d},\\
c_k &= -\frac{2kb((k-1)a+d)-e(2a-d)}{(2ka+d)((2k-2)a+d)},\\
d_k &= \frac{k((k-2)a+d)}{((2k-1)a+d)((2k-3)a+d)}\left(-c+\frac{((k-1)b+e)(((k-1)a+d)b-ae)}{((2k-2)a+d)^2}\right).
\end{aligned}
$$

The properties of Theorem 5.3 are shared by all classical orthogonal polynomials (Jacobi, Laguerre, Hermite). On the other hand, one has by Favard's Theorem (see for instance [**19**]), that all polynomial families defined by a three-term recurrence of shape (3.1) are orthogonal with respect to some moment functional. If one demands orthogonality with respect to a positive definite moment functional (in order to use all known facts about zeros of orthogonal polynomials etc.), then one exactly gets only Jacobi, Laguerre and Hermite up to a linear transformation $x \mapsto \nu_1 x + \nu_2$ with $\nu_1, \nu_2 \in \mathbb{R}$ (see the results of Lesky in [**15**]). Hence, one can completely characterize all positive definite orthogonal solutions of (5.1) just by looking at the coefficients a, b, c, d, e (see [**9**]). This can be translated into conditions on c_k and d_k for the general equation

$$\mathcal{A}p_m(x) + \mathcal{B}p_n(y) = \mathcal{C}.$$

References

[1] M. Beck, *Counting lattice points by means of the residue theorem*, Ramanujan J. **4** (2000), no. 3, 299–310.

[2] Yu. F. Bilu, Th. Stoll, and R. F. Tichy, *Octahedrons with equally many lattice points*, Period. Math. Hungar. **40** (2000), no. 2, 229–238.

[3] Yu. F. Bilu and R. F. Tichy, *The Diophantine equation* $f(x) = g(y)$, Acta Arith. **95** (2000), no. 3, 261–288.

[4] H. Davenport, D. J. Lewis and A. Schinzel, *Equations of the form* $f(x) = g(y)$, Quat. J. Math. Oxford **12** (1961), 304–312.

[5] E. Ehrhart, *Sur un problème de géométrie diophantienne linéaire. II. Systèmes diophantiens linéaires*, J. Reine Angew. Math. **227** (1967), 25–49.

[6] P. Erdős and J. L. Selfridge, *The product of consecutive integers is never a power*, Illinois J. Math. **19** (1975), 292–301.

[7] L. Hajdu, *On a Diophantine equation concerning the number of integer points in special domains. II*, Publ. Math. Debrecen **51** (1997), no. 3-4, 331–342.

[8] L. Hajdu, *On a Diophantine equation concerning the number of integer points in special domains*, Acta Math. Hungar. **78** (1998), no. 1-2, 59–70.

[9] Helmstädter-Rösner, Über Polynomlösungen Sturm-Liouvillescher Differential- bzw. Differenzengleichungen zweiter Ordnung - Dissertation, Polyfoto-Dr. Vogt KG, (Stuttgart, 1986).

[10] P. Kirschenhofer, A. Pethő, and R. F. Tichy, *On analytical and Diophantine properties of a family of counting polynomials*, Acta Sci. Math. (Szeged) **65** (1999), no. 1-2, 47–59.

[11] P. Kirschenhofer and O. Pfeiffer, *On a class of combinatorial Diophantine equations*, Sém. Lothar. Combin. **44** (2000), Art. B44h, 7 pp. (electronic).

[12] P. Kirschenhofer and O. Pfeiffer, *Diophantine equations between polynomials obeying second order recurrences*, Period. Math. Hungar. **47** (2003), no. 1-2, 119–134.

[13] W. Koepf and D. Schmersau, Recurrence equations and their classical orthogonal polynomial solutions, *Appl. Math. Comput.* **128(2-3)** (2002) 303–327.

[14] R. Koekoek and R. F. Swarttouw, *The Askey-Scheme of Hypergeometric Orthogonal Polynomials and its q-Analogue.*, Delft, Netherlands, Report **98-17** (1998).

[15] P. Lesky, Die Charakterisierung der klassischen orthogonalen Polynome durch Sturm-Liouvillesche Differentialgleichungen. *Arch. Rational Mech. Anal.* **10** (1962) 341–351.

[16] Th. Stoll and R. F. Tichy, *Diophantine equations for classical continuous orthogonal polynomials*, Indag. Math. (N.S.) **14** (2003), no. 2, 263–274.

[17] Th. Stoll and R. F. Tichy, *Diophantine equations involving general Meixner and Krawtchouk polynomials*, Quaest. Math. **28** (2005), no. 1, 105–115.

[18] Th. Stoll and R. F. Tichy, *Diophantine equations for Morgan-Voyce and other modified orthogonal polynomials*, submitted, see *http://dmg.tuwien.ac.at/stoll/publ.html*.

[19] W. Van Assche, Orthogonal polynomials in the complex plane and on the real line, in: *Special functions, q-series and related topics (Toronto, ON, 1995)*, Amer. Math. Soc., Providence, RI, (1997) 211–245.

Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Wiedner Hauptstrasse 8–10/104, A-1040 Vienna, Austria
E-mail address: stoll@dmg.tuwien.ac.at
URL: http://dmg.tuwien.ac.at/stoll/

Department of Mathematics A, Graz University of Technology, Steyrergasse 30, A-8010 Graz, Austria
E-mail address: tichy@tugraz.at
URL: http://finanz.math.tu-graz.ac.at/tichy/

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Hankel Determinants for Some Common Lattice Paths

Robert A. Sulanke and Guoce Xin

ABSTRACT. For a single value of ℓ, let $f(n,\ell)$ denote the number of lattice paths that use the steps $(1,1)$, $(1,-1)$, and $(\ell,0)$, that run from $(0,0)$ to $(n,0)$, and that never run below the horizontal axis. Equivalently, $f(n,\ell)$ satisfies the quadratic functional equation $F(x)=\sum_{n\geq 0} f(n,\ell)x^n = 1+x^\ell F(x)+x^2F(x)^2$. Let H_n denote the n by n Hankel matrix, defined so that $[H_n]_{i,j}=f(i+j-2,\ell)$. Here we investigate the values of such determinants where $\ell=0,1,2,3$. For $\ell=0,1,2$ we are able to employ the Gessel-Viennot-Lindström method. For the case $\ell=3$, the sequence of determinants forms a sequence of period 14, namely,

$$(\det(H_n))_{n\geq 1}=(1,1,0,0,-1,-1,-1,-1,-1,0,0,1,1,1,1,1,0,0,-1,-1,-1,\ldots)$$

For this case we are able to use the continued fractions method recently introduced by Gessel and Xin. We also apply this technique to evaluate Hankel determinants for other generating functions satisfying a certain type of quadratic functional equation.

RÉSUMÉ. Pour une seule valeur de ℓ, soit $f(n,\ell)$ le nombre des chemins treillis que utilise les pas $(1,1)$, $(1,-1)$, et $(\ell,0)$, vient de $(0,0)$ á $(n,0)$, et que ne vient jamais dessous l'axis horizontale. Équivalentement, le $f(n,\ell)$ satisfié l'équation fonctionnelle quadratique $F(x)=\sum_{n\geq 0} f(n,\ell)x^n = 1+x^\ell F(x)+x^2F(x)^2$. Soit H_n le n par n matrice de Hankel, définit pour que $[H_n]_{i,j}=f(i+j-2,\ell)$. Nous examinons de tels déterminants oú $\ell=0,1,2,3$. Pour $\ell=0,1,2$ nous pouvons employer la méthode de Gessel-Viennot-Lindström. Pour le cas $\ell=3$, ls séquence de déterminants forme une séquence de période 14, á savoir

$$(\det(H_n))_{n\geq 1}=(1,1,0,0,-1,-1,-1,-1,-1,0,0,1,1,1,1,1,0,0,-1,-1,-1,\ldots)$$

Pour ce cas que nous pouvons utiliser la méthode de fractions continuée récemment introduit par Gessel et Xin. Nous appliquons aussi cette technique pour évaluer les déterminants de Hankel pour l'autres fonctions generatrices quie satisfait un certain type d'équation fonctionnelle qudratique.

1. Introduction

We will consider lattice paths that use the following three steps: $U=(1,1)$, the up diagonal step; $H=(\ell,0)$, the horizontal step of length ℓ, where ℓ is a single nonnegative integer; and $D=(1,-1)$, the down diagonal step. Further, each H step will be weighted by t, and the others by 1. The weight of a path is the product of the weights of its steps. The weight of a path set is the sum of the weights of its paths.

Let $f(n,t,\ell)$ denote the weight of the path set of paths running from $(0,0)$ to $(n,0)$ that never run below the x-axis. When $t=1$, weight becomes cardinality. For example,

- $f(n,0,0)$, equivalently $f(n,0,\ell)$, is the weight of a set of Dyck paths, counted by the aerated Catalan numbers:
  $(f(0,0,0),f(1,0,0),f(2,0,0),\ldots)=(1,0,1,0,2,0,5,0,14,0,42,0,132,0,429,\ldots)$.
- $f(n,t,1)$ is the weight of a set of Motzkin paths, counted by the Motzkin numbers:
  $(f(0,1,1),f(1,1,1),f(2,1,1),\ldots)=(1,1,2,4,9,21,51,127,323,835\ldots)$.
- $f(n,t,2)$ is the weight of a set of large Schröder paths, counted by the aerated large Schröder numbers:
  $(f(0,1,2),f(1,1,2),f(2,1,2),\ldots)=(1,0,2,0,6,0,22,0,90,0,394,0,)$.
- $(f(0,1,3),f(1,1,3),f(2,1,3),\ldots)=(1,0,1,1,2,3,6,10,20,36,72,136,273,532,\ldots)$

2000 *Mathematics Subject Classification.* Primary 05A15; Secondary 15A36.
Key words and phrases. Hankel determinants, lattice paths, continued fractions.

Previously, Pergola, *et al* [**9**] and Sulanke [**11**] have considered such generalized Motzkin paths for various values of ℓ and have given additional references. Letting

$$F(x) = \sum_{n \geq 0} f(n, t, \ell) x^n$$

denote the generating function for $f(n, t, \ell)$, we find by a common combinatorial decomposition that $F(x)$ satisfies the functional equation

$$F(x) = 1 + tx^\ell F(x) + x^2 F(x)^2.$$

Any sequence $A = (a_0, a_1, a_2 \ldots)$ defines a sequence of Hankel matrices, $H_1, H_2, H_3 \ldots$, where H_n is an n by n matrix with entries $(H_n)_{i,j} = a_{i+j-2}$. For instance, the sequence $(f(n, 1, 3))_{n \geq 0}$ yields

$$H_1 = \begin{bmatrix} 1 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad H_4 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 6 \end{bmatrix}$$

Our interest is to consider, for nonnegative integer ℓ, the corresponding sequence of determinants $\det(H_n)$ where each matrix H_n has entries

$$(H_n)_{i,j} = f(i + j - 2, t, \ell).$$

The following propositions constitute our main results:

PROPOSITION 1.1. *For $n \geq 0$, $\ell = 1$, and arbitrary t (including $t = 0$, yielding the Dyck path case)*

$$\det(H_n) = 1.$$

PROPOSITION 1.2. *For $n \geq 0$, $\ell = 2$, and arbitrary t (including $t = 0$, yielding the Dyck path case),*

$$\det(H_n) = \begin{cases} (1+t)^{n^2/4} & \text{if } n \text{ is even} \\ (1+t)^{(n-1)(n+1)/4} & \text{if } n \text{ is odd} \end{cases}$$

PROPOSITION 1.3. *For $t = 1$ and $\ell = 3$,*

$$(\det(H_n))_{n \geq 1}^{14} = (1, 1, 0, 0, -1, -1, -1, -1, -1, 0, 0, 1, 1, 1).$$

Moreover, if $m, n \geq 0$ with $n - m = 0 \mod 14$ then $\det(H_m) = \det(H_n)$.

In Section 2, using the well-known combinatorial method of Gessel-Viennot-Lindström [**3**] [**5**] [**13**], we will prove Propositions 1.1 and 1.2. Our proof of Propositions 1.1 is essentially that of Viennot [**13**] who also used the method to calculate various other Hankel determinants relating to Motzkin paths. Aigner [**1**] also studied such determinants. We note that earlier Shapiro [**10**] demonstrated that the Hankel determinants for the usual Catalan numbers is 1. For the large Schröder numbers $(r(n))_{n \geq 0} = 1, 2, 6, 22, 90, 394, \ldots$ whose generating function satisfies

$$R(x) = \sum_{k \geq 0} r(k) x^k = 1 + xR(x) + xR(x)^2,$$

we show that the n-order Hankel determinant is $2^{n(n-1)/2}$, as stated in Proposition 2.1.

We remark that the problem of evaluating Hankel determinants corresponding to a generating function has received significant attention as considered by Wall [**14**]. One of the basic tools for such evaluation is the method of continued fractions, either by J-fractions in Krattenthaler [**8**] or Wall [**14**] or by S-fractions in Jones and Thron [**7**, Theorem 7.2]. However, both of these methods need the condition that the determinant can never be zero, a condition not always present in our study. Recently, Brualdi and Kirkland [**4**] used the J-fraction expansion to calculate Hankel determinants for various sequences related to the Schröder numbers. A slight modification of their proof of [**4**, Lemma 4.7] proves our Proposition 2.1 for $t = 1$.

In Section 3 we establish the periodicity of 14 for the case $\ell = 3$ of Proposition 1.3, by the continued fraction method recently developed by Gessel and Xin [**6**]. In the final section, we review their technique more generally: it yields a transformation for generating functions, satisfying a certain quadratic functional equation, that also transforms the associated Hankel determinants in a simple manner. We apply this transformation to evaluate the Hankel determinants for the cases $\ell = 1, 2$ (again) and for other path enumeration sequences related to $\ell = 3$.

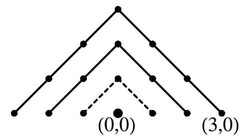


FIGURE 1. Some of the 4-tuples of paths for $\ell = 1$ and for I-T-CONFIG with $[(0,0),(-1,0),(-2,0),(-3,0)]$ and $[(0,0),(1,0),(2,0),(3,0)]$. In each of these 4-tuples there is a point path (a path of zero length) at $(0,0)$. The first 4-tuple is the only nonintersecting 4-tuple for this case. The second and third 4-tuples are intersecting only at the point $(0,1)$. The second 4-tuple corresponds to the permutation $\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 2 & 4 & 3\end{pmatrix}$, having sign of -1, while the third corresponds to the permutation $\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 3 & 4 & 2\end{pmatrix}$, having sign of 1. These two 4-tuples cancel one another under the Gessel-Viennot-Lindström method.

2. Employing the Gessel-Viennot-Lindström method

Assuming a rudimentary knowledge of the Gessel-Viennot-Lindström method, we reformulate it to our needs. All lattice paths use the three steps as previously defined. Given an n-tuple of lattice paths on the $\mathbb{Z} \times \mathbb{Z}$ plane, we say that it is *nonintersecting* if no steps from different paths share a common end point. Thus an nonintersecting n-tuple may have paths crossing or touching at points other than a common step end point.

Let $[(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)]$ and $[(x'_1,y'_1),(x'_2,y'_2),\ldots,(x'_n,y'_n)]$ denote two lists of distinct lattice points such that

$$x_{k+1} \le x_k \le 0 \text{ and } 0 \le y_k \le y_{k+1}$$

and

$$0 \le x'_k \le x'_{k+1} \text{ and } 0 \le y'_k \le y'_{k+1}.$$

We will refer to such a pair of lists as an "I-T-CONFIG" of order n as their points will be the initial and terminal points for each n-tuple of paths being considered.

Let $P_{i,j}$ denote the set of all paths running from (x_i,y_i) to (x'_j,y'_j) that never run below the x-axis, with $|P_{i,j}|$ denoting the sum of the weights of its paths. Let S_n denote the set of permutations on $\{1,2,3,\ldots,n\}$. For any permutation $\sigma \in S_n$, let P_σ denote the set of all n-tuples of paths $(p_1,p_2,\ldots,p_n)$, where $p_i \in P_{i,\sigma(i)}$ for $1 \le i \le n$. The *signed weight* of $(p_1,p_2,\ldots,p_n) \in P_\sigma$ is defined to be $\mathrm{sgn}(\sigma)$ times the product of the weights of the n paths. See Figures 1 and 2.

For our purpose the Gessel-Viennot-Lindström method is formulated in a form similar to that in Viennot's notes [**13**]:

LEMMA 2.1. *Given an* I-T-CONFIG *of order n, the sum of the signed weights of the nonintersecting n-tuples in $\cup_{\sigma\in S_n} P_\sigma$ is equal to* $\det(\ (|P_{i,j}|)_{1\le i,j\le n}\ )$.

PROOF OF PROPOSITION 1.1. (A similar proof appears in [**13**].) By Lemma 2.1 $\det(H_n)$ is equal to the sum of the signed weights of the nonintersecting n-tuples in $\cup_{\sigma\in S_n} P_\sigma$ for the I-T-CONFIG where $(x_i,y_i) = (-i+1,0)$ and $(x'_i,y'_i) = (i-1,0)$, for $1 \le i \le n$. Thus, for this I-T-CONFIG, we seek the nonintersecting n-tuples. First, the 1-tuple $P_{1,1}$ contains just the point path beginning and ending at $(0,0)$. Next, any nonintersecting path from $(-i+1,0)$, for $1 < i \le n$, must begin with an U step, while any nonintersecting path to $(j-1,0)$, for $1 < j \le n$, must end with an D step. Repeating this analysis at each integer-ordinate level k, shows the nonintersecting path from $(-i+1,0)$, $1 \le i \le k$, is forced to be a sequence of U steps followed by a sequence of D steps; moreover, it shows that any nonintersecting path from $(-i+1,0)$ to $(j-1,0)$, $k < i,j$, must start with k U steps and end with k D steps. Inductively, each nonintersecting path is a sequence of U steps followed by a sequence of D steps. The n-tuple of such paths is the only nonintersecting n-tuple of $\cup_{\sigma\in S_n} P_\sigma$, and it has weight equal 1. $\square$

We will use the following in proving Proposition 1.2:

LEMMA 2.2. *For the lattice paths that use the steps U, $H = (2,0)$, and D, that never run below the x-axis, and that have the* I-T-CONFIG,

$$(x_i, y_i) = (-2i+2, 0) \text{ and } (x_i', y_i') = (2i-2, 0)$$

for $1 \le i \le n$, the sum of the signed weights of the nonintersecting n-tuples in $\cup_{\sigma \in S_n} P_\sigma$ equals $(1+t)^{n(n-1)/2}$.

PROOF. For $(p_1, p_2, \ldots, p_n) \in (P_{1,\sigma(1)}, P_{2,\sigma(2)}, \ldots, P_{n,\sigma(n)})$, suppose that $(p_1, p_2, \ldots, p_n)$ is a nonintersecting n-tuple of paths for some permutation σ. Since the points in the I-T-CONFIG are spaced two units apart, the horizontal distance at any integer ordinate between any two paths of $(P_{1,\sigma(1)}, P_{2,\sigma(2)}, \ldots, P_{n,\sigma(n)})$ must be even. It follows inductively that, for $1 \le i \le n$, any path of the path set $P_{i,\sigma(i)}$ must begin with a sequence of $i-1$ U-steps and finish with a sequence of $\sigma(i)-1$ D-steps. Thus, computing the weight of the nonintersecting n-tuples is equivalent to computing the weight of the nonintersecting n-tuples for the new ("V" shaped) initial-terminal configuration, denoted by I-T-CONFIG-NEW, defined by

$$(x_i, y_i) = (-i+1, i-1) \text{ and } (x_i', y_i') = (i-1, i-1)$$

for $1 \le i \le n$.

Before continuing, we notice, for example when $t = 1$ and $n = 4$, that the matrix $M(0)$ define by $(M(0)_{i,j})_{1 \le i,j \le 4} = (|P'_{i,j}|)_{1 \le i,j \le 4}$ for I-T-CONFIG-NEW is an array of Delannoy numbers. (See [**2**], [**12**].) When $t = 0$, $M(0)$ is the initial array from Pascal's triangle. In the following array for $t = 1$, the entries count the ways a chess king can move from the north-west corner if it uses only east, south, or south-east steps. Momentarily we will see the role of the argument 0 in $M(0)$.

$$M(0) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 5 & 7 \\ 1 & 5 & 13 & 25 \\ 1 & 7 & 25 & 63 \end{bmatrix}$$

Now for arbitrary t and n, let $M(0)$ be the n by n matrix defined recursively by

$$M(0)_{i,j} = M(0)_{i-1,j} + tM(0)_{i-1,j-1} + M(0)_{i,j-1}$$

for $1 < i$ and $1 < j$ with $M(0)_{1,j} = 1$ and $M(0)_{i,1} = 1$ for $1 \le i$ and $1 \le j$. By Lemma 2.1 $M(0) = (|P_{ij}|)_{1 \le i,j \le n}$ for I-T-CONFIGNEW. Thus $\det(M(0))$ is equal to the weight of the nonintersecting n-tuples for I-T-CONFIG. The proof is completed once we show

$$\det(M(0)) = (1+t)^{n(n-1)/2}.$$

Given $M(0)$, we recursively define a sequence of n by n matrices

$$M(0),\ M(1),\ M(2),\ \ldots, M(n-1)$$

where, for $1 \le k \le n-1$,

$$M(k)_{ij} = \begin{cases} M(k-1)_{i,j} & \text{for } 1 \le i \le k \\ M(k-1)_{i,j} - M(k-1)_{i-1,j} & \text{for } k+1 \le i \le n \end{cases}$$

With CLAIM(k) denoting the claim that

$$\begin{aligned} M(k)_{i,j} &= M(k)_{i-1,j} + tM(k)_{i-1,j-1} + M(k)_{i,j-1} \text{ for } i,j > k, \\ M(k)_{i,i} &= (1+t)^{i-1} \text{ for } i \le k, \\ M(k)_{i,j} &= 0 \text{ for } i > j \text{ and } j \le k, \\ M(k)_{i,k+1} &= (1+t)^k \text{ for } i \ge k+1, \end{aligned}$$

one can establish CLAIM(k) for $1 \le k \le n-1$ by induction. Since $M(n-1)$ is upper triangular, we observe that

$$\det(M(n-1)) = (1+t)^{n(n-1)/2}.$$

By the type of row operations used to obtain the sequence $M(0), M(1), M(2), \ldots, M(n-1)$, their determinants are equal. $\square$

Since, by the I-T-CONFIG of Lemma 2.2, $(H)_{i,j} = |P_{i,j}|$ counts the large Schröder paths from $(0,0)$ to $(2i+2j-2,0)$, immediately we have the the following corollary for the Hankel determinants of the weighted non-aerated Schröder numbers:

PROPOSITION 2.1. *Let f_n denote the weight of the path set of paths from $(0,0)$ to $(2n,0)$ which never run beneath the x-axis and where $H = (2,0)$ is weighted by t. Equivalently, let f_n satisfy*

$$F(x) = \sum_{n\geq 0} f_n x^n = 1 + txF(x) + xF(x)^2.$$

Then the determinant of the n-th order Hankel matrix equals $(1+t)^{n(n-1)/2}$.

As a second corollary to Lemma 2.2, we have

LEMMA 2.3. *For the lattice paths that use the steps U, $H = (2,0)$, and D, that never run below the x-axis, and that have the I-T-CONFIG with*

$$(x_i, y_i) = (-2i+1, 0) \text{ and } (x'_i, y'_i) = (2i-1, 0)$$

for $1 \leq i \leq n$, the sum of the signed weights for the nonintersecting n-tuples in $\cup_{\sigma\in S_n} P_\sigma$ is $(1+t)^{n(n+1)/2}$.

PROOF OF LEMMA 2.3. We first translate all paths upwards one unit and then prepend a U-step and append a D-step to every path. Next we add the point path at $(0,0)$. The sum of the signed weights of the nonintersecting n-tuples in the original configuration equals that of the nonintersecting $n+1$-tuples in this new configuration, which in turn is given by Lemma 2.2. □

PROOF OF PROPOSITION 1.2. Suppose that n is even; the proof when n is odd is similar. Here the Hankel matrix $(|P_{i,j}|)_{1\leq i,j\leq n}$ corresponds to the I-T-CONFIG with

$$(x_i, y_i) = (-i+1, 0) \text{ and } (x'_i, y'_i) = (i-1, 0) \text{ for } 1 \leq i \leq n.$$

Since $\ell = 2$, no endpoint of a step on a path that originates from an oddly indexed initial point (i.e., a point $(-i+1,0)$ for odd i) will intersect an endpoint of a step on a path that originates from an evenly indexed initial point. Moreover, for any permutation σ corresponding to a nonintersecting n-tuple, $\sigma(i) - i$ must be even for each i, and hence $\mathrm{sgn}(\sigma) = 1$. Thus the weight of the nonintersecting n-tuples is the product of the weight of those originating from oddly indexed initial points times the weight of those originating from evenly indexed initial points.

Hence, with $m = n/2$, let I-T-CONFIGA have

$$(x_i, y_i) = (-2i+2, 0) \text{ and } (x'_i, y'_i) = (2i-2, 0) \text{ for } 1 \leq i \leq m,$$

and let I-T-CONFIGB have

$$(x_i, y_i) = (-2i+1, 0) \text{ and } (x'_i, y'_i) = (2i-1, 0) \text{ for } 1 \leq i \leq m.$$

Applying Lemmas 2.2 and 2.3 to these configurations yields the weight of nonintersecting n-tuples of the original configuration as

$$(1+t)^{m(m-1)/2}(1+t)^{m(m+1)/2} = (1+t)^{m^2} = (1+t)^{n^2/4}.$$

□

Next we consider Hankel determinants for sequences of path weights that ignore the initial term. For the sequence $f(1,t,\ell), f(2,t,\ell), \ldots$, we will let H^1_n denote the matrix where the entries satisfy $(H^1_n)_{i,j} = f(i+j-1,t,\ell)$. See Figure 2.

PROPOSITION 2.2. *For $\ell = 1$ (Motzkin case again), the sequence of determinants satisfies the recurrence*

$$\det(H^1_n) = t\ \det(H^1_{n-1}) - \det(H^1_{n-2})$$

subject to $\det(H^1_1) = t$ *and* $\det(H^1_2) = (t-1)(t+1)$.

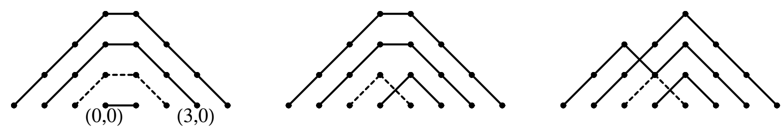


FIGURE 2. Three of the 4-tuples of paths for $\ell = 1$ and for I-T-CONFIG with $[(0,0),(-1,0),(-2,0),(-3,0)]$ and $[(1,0),(2,0),(3,0),(4,0)]$. The first and second 4-tuples are both nonintersecting. The first has a signed weight of t^4 while the second has a signed weight of $-t^2$. The third is intersecting only at the point $(0,1)$.

PROOF. Aigner [**1**] considered the case for $t = 1$. For arbitrary t, our proof considers how the particular paths must look in the nonintersecting case. Observe that $\det(H_n^1)$ is the sum of the weights of the nonintersecting n-tuples for the I-T-CONFIG(n) taken as

$$[(0,0),(-1,0),\ldots,(n-1,0)] \text{ and } [(1,0),(2,0),\ldots,(n,0)].$$

Each of these nonintersecting n-tuples belongs to one of two types: (1) those containing the path from $(0,0)$ to $(1,0)$ with all other paths forced to begin with U, end with D, and have ordinate at least one elsewhere; (2) those containing the path UD from $(0,0)$ to $(2,0)$ and the path UD $(-1,0)$ to $(1,0)$ with all other paths forced to begin with UU, end with DD, and have ordinate at least two elsewhere. The set of the first type has a total weight t times the sum of the weights of the nonintersecting $(n-1)$-tuples on the I-T-CONFIG(n-1), which is $t\det(H_{n-1}^1)$. Since each n-tuple of the second type has the defined crossing of the path from $(0,0)$ with that from $(-1,0)$, the set has total weight is the sign of the corresponding permutation times the sum of the weights of the nonintersecting $(n-2)$-tuples on the I-T-CONFIG(n-2), which is $-\det(H_{n-2}^1)$. □

For $\ell = 2$, we will indicate how Lemma 2.3 proves

PROPOSITION 2.3. *For $n \geq 0$, $\ell = 2$, and arbitrary t, the sequence of determinants satisfies*

$$\det(H_n^1) = \begin{cases} 0 & \textit{if } n \textit{ is odd} \\ (-1)^{n/2}(1+t)^{n(n+2)/4} & \textit{if } n \textit{ is even} \end{cases}$$

PROOF. Here the Hankel matrix can correspond to I-T-CONFIG with

$$(x_i, y_i) = (-i+1, 0) \text{ and } (x_i', y_i') = (i, 0) \text{ for } 1 \leq i \leq n.$$

Since $\ell = 2$, if there is a path from (x_i, y_i) to (x_j', y_j'), then $i - j$ is odd. It follows that, if n is odd, there can be no n-tuples of paths for the configuration. If n is even and $m = n/2$, the sign of any permutation for an nonintersecting n-tuples can be shown to be $(-1)^m$. Thus the weight of the nonintersecting n-tuples is $(-1)^m$ times the weight of those originating from oddly indexed initial points times the weight of those originating from evenly indexed initial points. The proof is completed by applying 2.3 to I-T-CONFIGA with

$$(x_i, y_i) = (-2i+2, 0) \text{ and } (x_i', y_i') = (2i, 0) \text{ for } 1 \leq i \leq m,$$

and to I-T-CONFIGB with

$$(x_i, y_i) = (-2i+1, 0) \text{ and } (x_i', y_i') = (2i-1, 0) \text{ for } 1 \leq i \leq m.$$

□

3. Periodicity fourteen and continued fractions

Here we will repeated apply the "continued fractions method" recently developed by Gessel and Xin [**6**] to determine the periodicity of the sequence of Hankel determinants for $\ell = 3$ and $t = 1$. This method, presented more formally in the next section, transforms both generating functions and corresponding determinants. In this section we will concentrate on the specific generating function $F(x)$ satisfying

$$F(x) = 1 + x^3F(x) + x^2F(x)^2.$$

From this functional equation, or from the related recurrence for its coefficients, there appears to be no clue why the associated sequence of Hankel determinants should have a period of 14.

For an arbitrary generating function $D(x,y)=\sum_{i,j=0}^{\infty} d_{i,j}x^iy^j$, let $[D(x,y)]_n$ denote the n by n determinant $\det((d_{i,j})_{0\le i,j\le n-1})$. For any $A(x)=\sum_{n\ge 0} a_nx^n$, define the *Hankel matrix for A of order n*, $n\ge 1$, by $H_n(A)=(a_{i+j-2})_{1\le i,j\le n}$. It is straight forward to show that the Hankel determinant $\det(H_n(A))$ can be expressed as

$$\det(H_n(A))=\left[\frac{xA(x)-yA(y)}{x-y}\right]_n.$$

We will use an easily-proven "product rule" of [**6**] for transforming the generating functions: *If $u(x)$ is a formal power series with $u(0)=1$, then*

$$[u(x)D(x,y)]_n=[D(x,y)]_n=[u(y)D(x,y)]_n.$$

We will make five transformations showing, for $n\ge 8$,

$$\det(H_n(F))=\det\left(\operatorname{diag}\left([1],[1],\begin{bmatrix}0&0&1\\0&1&0\\1&0&-2\end{bmatrix},[1],[1],H_{n-7}(F)\right)\right),$$

where the right side is the determinant of a block-diagonal matrix consisting of six blocks along the diagonal, four of which are 1 by 1 identity matrices, and having entry 0 elsewhere. It then follows that $\det(H_n(F))=-\det(H_{n-7}(F))$. This implies that the period for $\det(H_n(F))$ is 14, and Proposition 1.3 will be proved.

We start with $F_0(x)=F(x)$, and define $F_i(x)$ from $F_{i-1}(x)$ according to a transformation where each Hankel determinant for $F_i(x)$ are derived from one for $F_{i-1}(x)$ with the aid of the product rule, which is not always mentioned. In the following, $F_i(x)$ will always satisfy a quadratic functional equation

$$a(x)F_i(x)^2+b(x)F_i(x)+c(x)=0,$$

which is equivalent to the continued fraction

$$F_i(x)=\frac{-c(x)}{b(x)+a(x)F_i(x)}.$$

In particular, for $\ell=3$,

$$F_0(x)=\frac{1}{1-x^3-x^2F_0(x)}.$$

Transformation 1: Using this continued fraction of F_0, substitution, and simplification we obtain

$$\det(H_n(F))=\left[\frac{xF_0(x)-yF_0(y)}{x-y}\right]_n=\left[\frac{-xy^2F_0(y)+yx^2F_0(x)+(x-y)\left(yx^2+xy^2+1\right)}{(1-x^3-x^2F_0(x))(1-y^3-y^2F_0(y))(x-y)}\right]_n.$$

Multiplying by $(1-x^3-x^2F_0(x))(1-y^3-y^2F_0(y))$, which will not affect the value of the determinant by the product rule, we can write the determinant as

$$\left[1+xy\frac{xF_1(x)-yF_1(y)}{x-y}\right]_n$$

where

$$F_1(x)=F_0(x)+x. \tag{3.1}$$

The associated matrix is block-diagonal with two blocks: the matrix [1] and the Hankel matrix for $F_1(x)$. Certainly,

$$\det(H_n(F_0))=\det(H_{n-1}(F_1)).$$

From (3.1) and the functional equation for $F_0(x)$, we obtain the functional equation

$$F_1(x)=\frac{1+x}{1+x^3-x^2F_1(x)}.$$

Transformation 2: Using this continued fraction for F_1, substituting in $\dfrac{xF_1(x)-yF_1(y)}{x-y}$, and multiplying by $(1+x^3-x^2F_1(x))(1+y^3-y^2F_1(y))$ yields

$$\left[\frac{xF_1(x)-yF_1(y)}{x-y}\right]_n=\left[\frac{-xy^2\,(x+1)\,F_1(y)+yx^2\,(y+1)\,F_1(x)-(y+1)\,(x+1)\,(xy-1)\,(x-y)}{x-y}\right]_n.$$

Upon multiplying by $(1+x)^{-1}(1+y)^{-1}$, the determinant is equal to

$$\left[1+xy\frac{xF_2(x)-yF_2(y)}{x-y}\right]_n,$$

where

$$F_2(x)=F_1(x)/(1+x)-1. \tag{3.2}$$

The associated matrix being block diagonal shows

$$\det(H_{n-1}(F_1))=\det(H_{n-2}(F_2)).$$

From (3.2) and the functional equation for $F_1(x)$, we obtain

$$F_2(x)=\frac{x^2}{1-2x^2-x^3-(x^3+x^2)\,F_2(x)}.$$

Transformation 3: Substituting for F_2 with the above fraction, simplifying, and multiplying by $(1+x)(1-x-x^2-x^2F_2(x))(1+y)(1-y-y^2-y^2F_2(y))$ shows that the determinant $\left[\frac{xF_2(x)-yF_2(y)}{x-y}\right]_n$ equals

$$\left[\frac{y^2x^3\,(y+1)\,F_2(y)-x^2y^3\,(x+1)\,F_2(x)-(x-y)\left(2\,y^2x^2-x^2-xy-y^2\right)}{x-y}\right]_n$$

which can be rewritten as

$$\left[x^2+xy+y^2-2x^2y^2+x^3y^3\frac{xF_3(x)-yF_3(y)}{x-y}\right]_n,$$

where $F_3(x)$ is indeed a power series satisfying

$$F_3(x)=(x+1)F_2(x)/x^2. \tag{3.3}$$

This time the corresponding matrix is a block-diagonal matrix with the block $\begin{bmatrix}0&0&1\\0&1&0\\1&0&-2\end{bmatrix}$ followed by the Hankel matrix for $F_3(x)$. Hence

$$\det(H_{n-2}(F_2))=-\det(H_{n-5}(F_3)).$$

From (3.3) and the functional equation for $F_2(x)$, we obtain

$$F_3(x)=\frac{1+x}{1-2x^2-x^3-x^4F_3(x)}.$$

Transformation 4: Substituting for F_3 with the fraction, simplifying, and multiplying by $(1-2x^2-x^3-x^4F_3(x))((1-2y^2-y^3-y^4F_3(y))$ the determinant $\left[\frac{xF_3(x)-yF_3(y)}{x-y}\right]_n$ equals

$$\left[\frac{-xy^4\,(x+1)\,F_3(y)+yx^4\,(y+1)\,F_3(x)+(y+1)\,(x+1)\,(xy+1)\,(x-y)}{x-y}\right]_n.$$

By multiplying the generating function by $(1+x)^{-1}(1+y)^{-1}$, this determinant becomes

$$\left[1+xy\frac{xF_4(x)-yF_4(y)}{x-y}\right]_n,$$

where

$$F_4(x)=1+x^2F_3(x)/(1+x). \tag{3.4}$$

Therefore,

$$\det(H_{n-5}(F_3))=\det(H_{n-6}(F_4)).$$

From (3.4) and the functional equation for $F_3(x)$, we obtain

$$F_4(x)=\frac{1}{1+x^3-(x^3+x^2)\,F_4(x)}.$$

Transformation 5: Substituting for F_4 with the above fraction, simplifying, and multiplying by $(1-x^2F_4(x))(1-y^2F_4(y))$ the determinant $\left[\frac{xF_4(x)-yF_4(y)}{x-y}\right]_n$ equals

$$\left[\frac{-xy^2\,(y+1)\,F_4(y)+x^2y\,(x+1)\,F_4(x)-(x-y)\left(yx^2+xy^2-1\right)}{x-y}\right]_n=\left[1+xy\frac{xF_5(x)-yF_5(y)}{x-y}\right]_n,$$

where $F_5(x)=(1+x)F_4(x)-x$. Hence, $\det(H_{n-6}(F_4))=\det(H_{n-7}(F_5))$.

Finally, it is routinely checked that $F_5(x)=F_0(x)$.

4. The quadratic transformation for Hankel determinants

One can use the method introduced in the previous section to evaluate the Hankel determinants for generating functions satisfying a certain type of quadratic functional equation. The generating functions $F(x)$ in this section are the unique solution of a quadratic functional equation satisfying

$$F(x)=\frac{x^d}{u(x)+x^kv(x)F(x)}, \tag{4.1}$$

where $u(x)$ and $v(x)$ are rational power series with nonzero constants, d is a nonnegative integer, and k is a positive integer. Note that if $k=0$, $F(x)$ is not unique. Our task now is to derive a transformation $\mathcal{T}$ so that $\det(H_n(F))=a\det(H_{n-d-1}(\mathcal{T}(F)))$ for some value a and nonnegative integer d. In addition to Hankel matrices for the power series $A=\sum_{n\geq 0}a_ix^i$, we will consider *shifted Hankel matrices*: $H_n^k(A)$ denotes the matrix $(a_{i+j+k-2})_{1\leq i,j\leq n}$. Shifted matrices have appeared in Proposition 2.2 and 2.3.

The first proposition is elementary:

Proposition 4.1. *If F satisfies (4.1), then $G=u(0)F$ satisfies*

$$\det(H_n(G))=u(0)^n\det(H_n(F)),\ \ \textit{and}\quad G(x)=\frac{x^d}{u(0)^{-1}u(x)+x^ku(0)^{-2}v(x)G(x)}.$$

Proposition 4.2. *Suppose F satisfies (4.1) with $u(0)=1$. We separate $u(x)$ uniquely as $u(x)=u_L(x)+x^{d+2}u_H(x)$, where $u_L(x)$ is a polynomial of degree at most $d+1$ and $u_H(x)$ is a power series.*

(i) *If $k=1$, then there is a unique G such that*

$$G(x)=\frac{-v(x)-xu_L(x)u_H(x)}{u_L(x)-x^{d+2}u_H(x)-x^{d+1}G(x)},$$

Moreover,

$$G(x)=-xu_H(x)-x^{-d}v(x)F(x)$$

and a shifted matrices appears with

$$\det(H^1_{n-d-1}(G(x)))=(-1)^{d(d+1)/2}\det(H_n(F(x))).$$

(ii) *If $k\geq 2$, then there is a unique G such that*

$$G(x)=\frac{-x^{k-2}v(x)-u_L(x)u_H(x)}{u_L(x)-x^{d+2}u_H(x)-x^{d+2}G(x)},$$

Moreover,

$$G(x)=-u_H(x)-x^{k-d-2}v(x)F(x)$$

and

$$\det(H_{n-d-1}(G(x)))=(-1)^{d(d+1)/2}\det(H_n(F(x))).$$

Proof. We prove only part (ii) as part (i) is similar. The generating function for $H_n(F)$ is given by

$$\begin{aligned}\frac{xF(x)-yF(y)}{x-y}&=\frac{1}{x-y}\left(\frac{x^{d+1}}{u(x)+x^kv(x)F(x)}-\frac{y^{d+1}}{u(y)+y^kv(y)F(y)}\right)\\&=\frac{-y^{d+1}u(x)-y^{d+1}x^kv(x)F(x)+x^{d+1}u(y)+x^{d+1}y^kv(y)F(y)}{\left(u(x)+x^kv(x)F(x)\right)\left(u(y)+y^kv(y)F(y)\right)(x-y)}\end{aligned}$$

We can multiply by $(u(x)+x^k v(x)F(x))$ and by $(u(y)+y^k v(y)F(y))$ without changing the above determinant by the product rule. Next we observe that x^d divides $F(x)$, and write $u(x) = u_L(x) + x^{d+2}u_H(x)$ as in the proposition. The resulting generating function can be written as

$$\frac{-y^{d+1}u_L(x) + x^{d+1}u_L(y)}{x-y} + (xy)^{d+1}\frac{-x(u_H(x) - x^{k-d-2}v(x)F(x)) + y(u_H(y) + y^{k-d-2}v(y)F(y))}{x-y}.$$

We now set $G(x) = -u_H(x) - x^{k-d-2}v(x)F(x)$, which can be straightforwardly shown to agree with the defining functional equations. Suppose that $u_L(x) = 1 + a_1 x + \cdots a_{d+1}x^{d+1}$, then $\left[\frac{xF(x) - yF(y)}{x-y}\right]_n$ is equal to the determinant of the block-diagonal matrix

$$\operatorname{diag}\left(\begin{pmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \vdots & a_{d-1} \\ 1 & a_1 & \cdots & a_d \end{pmatrix}, H(G(x))\right).$$

The determinant of the first block is easily seen to be $(-1)^{d(d+1)/2}$. □

Given these propositions and that $H^1(A) = H(x^{-1}(A(x) - A(0)))$ for any series A, we can now define our transformation $\mathcal{T}(F)$: *For F satisfying* (4.1),

- *if $u(0) \neq 1$, then $\mathcal{T}(F) = G$, as given in Proposition 4.1.*
- *if $u(0) = 1$ and $k = 1$, then $\mathcal{T}(F) = x^{-1}(G(x) - G(0))$, with G given in Proposition 4.2(i).*
- *if $u(0) = 1$ and $k \geq 2$, then $\mathcal{T}(F) = G$, as given in Proposition 4.2(ii).*

Moreover, the relation between $\det(H_n(F))$ *and* $\det(H_n(\mathcal{T}(F)))$ *is given in Propositions 4.1 and 4.2.*

Example 1: OTHER PROOFS OF PROPOSITIONS 1.1 AND 2.2. For Motzkin paths with arbitrary t, the generating function $F(x)$ satisfies

$$F(x) = \frac{1}{1 - tx - x^2 F(x)}.$$

Applying Proposition 4.2 so $F_1 = \mathcal{T}(F)$ gives

$$\det(H_{n-1}(F_1)) = \det(H_n(F)) \text{ where } F_1(x) = \frac{1}{1 - tx - x^2 F_1(x)}.$$

Hence, $F(x) = F_1(x)$, and consequently $\det(H_n(F(x))) = 1$ for all n.

Whereas the Gessel-Viennot-Lindström method leads to a proof in the shifted case for arbitrary t, as in Proposition 2.2, we have been able to use the continued fractions technique only for $t = 1$ and $t = 2$.

For $t = 1$ we will show that $(\det(H_n^1(F)))_{n\geq 1} = (1, 0, -1, -1, 0, 0, 1, 1, \ldots)$, continuing with period 6. Let $G_1(x) = (F(x) - 1)/x$, so that $\det(H_n^1(F)) = \det(H_n(G_1))$. Let $G_2 = \mathcal{T}(G_1)$ and $G_3 = \mathcal{T}(G_2)$, both under Proposition 4.2(ii). Since

$$G_1(x) = \frac{1 + x}{1 - x - 2x^2 - x^3 G_1(x)}$$

with $d = 0$, $k = 3$, $u(x) = u_L(x) = 1 - 2x$, $u_H = 0$, and $v(x) = -(1 + x)^{-1}$, we find that

$$G_2(x) = \frac{x}{1 - x - 2x^2 - x^2(1 + x)G_2(x)}$$

with $d = 1$, $k = 2$, $u(x) = u_L(x) = 1 - x - 2x^2$, $u_H = 0$, and $v(x) = -(1 + x)$. Applying Proposition 4.2(ii) shows

$$\begin{aligned} G_3(x) &= -x^{-1}(-(1 + x))G_2(x) \\ &= -x^{-1}(-(1 + x))(-x(-(1 + x)^{-1}))G_1(x) \\ &= G_1(x) \end{aligned}$$

and $\det(H_{n-3}(G_3)) = -\det(H_{n-1}(G_2)) = -\det(H_n(G_1))$, which yields the periodicity of the sequence of determinants.

For $t = 2$ we will show that $\det(H_n^1(F)) = n+1$ for $n \geq 1$. Define, G_1 to satisfy,

$$G_1(x) = \frac{2+x}{1-2x-2x^2-x^3G_1(x)}.$$

One can easily see that $G_1(x) = (F(x)-1)/x$ with $G_1(0) = u_1(0)^{-1} = \det(H_1(G_1)) = \det(H_1^1(F)) = 2$. For $n \geq 2$, define, G_n to satisfy,

$$G_n(x) = \frac{(n-1)^2(n^2+n+x)}{(n^2-n)(n^2-2n^2x-2x^2)-n^2(n^2-n+x)x^2G_n(x)}.$$

By induction one can show that $G_n = \mathcal{T} \circ \mathcal{T}(G_{n-1})$ (under Prop. 4.1 then under Prop. 4.2), and that $G_n(0) = u_n(0)^{-1} = (n-1)(n+1)/n^2$. Also by induction and Proposition 4.1, for $n \geq 2$,

$$\begin{aligned} \det(H_n(G_1)) &= \left[2^n \prod_{i=2}^{n-1} \left(\frac{(i-1)(i+1)}{i^2}\right)^{n+1-i}\right] \det(H_1(G_n)) \\ &= 2^n \prod_{i=2}^{n} \left(\frac{(i-1)(i+1)}{i^2}\right)^{n+1-i} \end{aligned}$$

which simplifies to $\det(H_n(G_1)) = n+1$.

Example 2: Another proof of Proposition 1.2. For large Schröder paths arbitrary t, we have

$$F(x) = \frac{1}{1-tx^2-x^2F(x)}.$$

Applying $\mathcal{T}$ gives

$$\det(H_{n-1}(F_1)) = \det(H_n(F)), \text{ where } F_1(x) = \frac{1+t}{1+tx^2-x^2F_1(x)}.$$

Applying $\mathcal{T}$ again, we obtain

$$(1+t)^n \det(H_{n-1}(F_2)) = \det(H_n(F_1)), \text{ where } F_2(x) = \frac{1}{1-tx^2-x^2F_2(x)}.$$

This implies $F_2 = F$, and hence the recurrence $\det(H_n(F)) = (1+t)^{n-1}\det(H_{n-2}(F))$, with initial condition $\det(H_1(F)) = 1$, and $\det(H_2(F)) = 1+t$.

Example 3: Another proof of Proposition 2.1. Consider the continued fraction

$$F(x) = \frac{1}{1-tx-xF(x)},$$

where $F(x)$ is the generating function for the Catalan numbers for $t=0$ and the large Schröder numbers for $t=1$.

Under Proposition 4.2(i) we have a unique G_1 such that $G_1(x) = F(x)$ and $\det(H_{n-1}^1(G_1)) = \det(H_n(F))$. Taking $G_2 = (G_1(x)-1)/x = (F(x)-1)/x$, we have

$$G_2(x) = \frac{(1+t)}{1-(2+t)x-x^2G_2(x)}$$

where $\det(H_{n-1}(G_2)) = \det(H_{n-1}^1(F))$ and $u(x) = (1-(2+t)x)/(1+t)$.

Under Proposition 4.1 we have a unique G_3

$$G_3(x) = \frac{1}{1-(2+t)x-(1+t)x^2G_3(x)},$$

with $G_3(x) = G_2/(1+t)$ and $\det(H_{n-1}(G_3)) = (1+t)^{-(n-1)}\det(H_{n-1}(G_2))$.

Under Proposition 4.2(ii) we have a unique G_4 such that $G_4(x) = (1+t)G_3(x)$ and $\det(H_{n-2}(G_4)) = \det(H_{n-1}(G_3))$.

We see that $G_4(x) = G_2(x)$; thus $\det(H_{n-1}(G_2)) = (1+t)^{n-1}\det(H_{n-2}(G_2))$ with $\det(H_1(G_2)) = 1+t$. Hence $\det(H_n(F)) = \det(H_{n-1}(G_2)) = (1+t)^{n(n-1)/2}$.

Example 4: ANOTHER PROOF OF PROPOSITION 2.3. To compute $\det(H_n^1(F))$, first we consider

$$H_n^1(F) = H_n(F_1), \text{ where } F_1 = \frac{(t+1)x}{1-(2+t)x^2 - x^3 F_1}.$$

Applying $\mathcal{T}$ shows that $\det(H_n(F_1)) = -(1+t)^n \det(H_{n-2}(F_1))$.

Example 5: For $\ell = 3$, recall the functional equation

$$F_0(x) = \frac{1}{1 - tx^3 - x^2 F_0(x)}.$$

For arbitrary t, our transformation gives more and more complicated expressions. This is not surprising since the Hankel determinants do not factor nicely. However, for $t = 1$ and for $k = 1, 2, 3$, the transformation gives nice results similar to that of Proposition 1.3: indeed, sequences of $\det(H_n^k(F_0))$ also have period 14. For $k = 4$ there is an interesting result.

Subexample 5i: The sequence for $\det(H_n^1(F_0))$ starts with $0, -1, 0, 1, 1, 0, -1, 0, 1, 0, -1, -1, 0, 1$. If we define F_1 so that $F_0(x) = 1 + xF_1(x)$, then

$$\det(H_n(F_1)) = \det(H_n^1(F_0)), \text{ with } F_1 = \frac{x(x+1)}{1 - 2x^2 - x^3 - x^3 F_1} \text{ and } d = 1.$$

Then applying $\mathcal{T}$ repeatedly so $\mathcal{T}(F_i) = F_{i+1}$, we obtain

$$\begin{aligned}
\det(H_{n-2}(F_2)) &= -\det(H_n(F_1)), \text{ where } F_2 = \frac{x}{(x+1)(1 - x - x^2 - x^3 F_2)} \text{ and } d = 1;\\
\det(H_{n-2}(F_3)) &= -\det(H_n(F_2)), \text{ where } F_3 = \frac{1 + x - x^2}{1 - 2x^2 + x^3 - x^3 F_3} \text{ and } d = 0;\\
\det(H_{n-1}(F_4)) &= \det(H_n(F_3)), \text{ where } F_3 = \frac{x}{(1 + x - x^2)(1 - x - x^2 F_3)} \text{ and } d = 1;\\
\det(H_{n-2}(F_5)) &= -\det(H_n(F_4)), \text{ where } F_5 = \frac{x(x+1)}{1 - 2x^2 - x^3 - x^3 F_5}.
\end{aligned}$$

The periodicity is established by noticing that $F_5 = F_1$ and $\det(H_{n-7}(F_5)) = -\det(H_n(F_1))$.

Subexample 5ii: The sequence for $\det(H_n^2(F_0))$ starts with $1, 1, 1, 1, 0, 0, -1, -1, -1, -1, -1, 0, 0, 1,$. If we define G_0 so that $F_0(x) = 1 + x^2 G_0(x)$, then $\det(H_n(G_0)) = \det(H_n^2(F_0))$,

$$G_0 = \frac{1 + x}{1 - 2x^2 - x^3 - x^4 G_0}.$$

One can establish the periodicity using Proposition 4.2. However, this generating function has appeared in Transformation 3 of section 3, where one can see that

$$\det(H_n(G_0)) = -\det(H_{n+5}(F_0)). \tag{4.2}$$

Subexample 5iii: The sequence for $\det(H_n^3(F_0))$ starts with $1, -1, -1, 0, 0, 0, -1, -1, 1, 1, 0, 0, 0, 1$ and continues with period 14. The verification for this case uses Proposition 4.2(ii) occasionally interspersed with Proposition 4.1. Here we will only sketch the verification. By defining F_1 so that $F_0(x) = 1 + x^2 + x^3 F_1(x)$, one finds that

$$F_1 = \frac{1 + 2x + x^2 + x^3}{1 - 2x^2 - x^3 - 2x^4 - x^5 F_1}.$$

For the first transformation, with $F_2 = \mathcal{T} F_1$, we find

$$F_2 = \frac{1 - 2x + x^3}{-1 + 4x^2 + x^3 + 2x^4 - x^2(1 + 2x + x^2 + x^3) F_2},$$

in which $u(x) = (-1 + 4x^2 + x^3 + 2x^4)/(1 - 2x + x^3)$. Now, since $u(0) = -1$, one needs to apply Proposition 4.1 for the next transformation. One proceeds until a generating function equal to F_1 appears to establish the periodicity. We remark that $d = 0$ for each transformation until the final one which uses Proposition 4.2(ii) with $d = 3$ (This corresponds to a fourth order block).

Subexample 5iv: The sequence for $\det(H_n^4(F_0))$ begins with

$$2, 3, 4, 0, 0, -4, -5, -6, -7, -8, 0, 0, 8, 9, 10, 11, 12, 0, 0, -12, -13, -14, -15, -16, 0, 0, 16, \ldots.$$

For $n \geq 8$, an essence of periodicity can be gleaned from the recurrence

$$\det(H_n^4(F_0)) = 4\det(H_{n-1}(F_0)) - \det(H_{n-7}^4(F_0)),$$

for which we sketch a proof, often omitting the functional equations.

We will be applying the transformation $\mathcal{T}$ eight times, alternating its definition to be first under Proposition 4.1 and then under Proposition 4.2(ii). Let F_1 satisfy $F_0 = 1 + x^2 + x^3 + x^4 F_1$. Hence, $\det(H_n(F_1)) = \det(H_n^4(F_0))$, and

$$F_1 = \frac{2 + 3x + 2x^2 + 2x^3 + x^4}{1 - 2x^2 - x^3 - 2x^4 - 2x^5 - x^6 F_1}.$$

Here $u(0) = \frac{1}{2}$, where $u(x)$ is for F_1. Thus, with $F_2 = \mathcal{T}F_1$, $\det(H_n(F_2)) = (\frac{1}{2})^n \det(H_n(F_1))$. Now $d = 0$, where d is for F_2. With $F_3 = \mathcal{T}F_2$, $\det(H_{n-1}(F_3)) = \det(H_n(F_2))$.

Here $u(0) = \frac{4}{3}$, where $u(x)$ is for F_3. Thus, with $F_4 = \mathcal{T}F_3$, $\det(H_{n-1}(F_4)) = (\frac{4}{3})^{n-1} \det(H_{n-1}(F_3))$. Now $d = 0$, where d is for F_4. With $F_5 = \mathcal{T}F_4$, $\det(H_{n-2}(F_5)) = \det(H_{n-1}(F_4))$.

Here $u(0) = \frac{9}{8}$, where $u(x)$ is for F_5. Thus, with $F_6 = \mathcal{T}F_5$, $\det(H_{n-2}(F_6)) = (\frac{9}{8})^{n-2} \det(H_{n-2}(F_5))$. Now $d = 0$, where d is for F_6. With $F_7 = \mathcal{T}F_6$, $\det(H_{n-3}(F_7)) = \det(H_{n-2}(F_6))$.

Here $u(0) = \frac{4}{3}$, where $u(x)$ is for F_7. Thus, with $F_8 = \mathcal{T}F_7$, $\det(H_{n-3}(F_8)) = (\frac{4}{3})^{n-3} \det(H_{n-3}(F_7))$.

Now $d = 2$, where d is for F_8. With $F_9 = \mathcal{T}F_8$, $\det(H_{n-6}(F_9)) = -\det((H_{n-3}(F_8)) =$ $(\frac{5}{4}, \frac{6}{4}, \frac{7}{4}, \frac{8}{4}, 0, 0, -\frac{8}{4}, -\frac{9}{4}, -\frac{10}{4}, \ldots)$.

Thus, (surprisingly)

$$\begin{aligned}
\det(H_{n-6}(F_9)) &= -(\frac{1}{2})^n (\frac{4}{3})^{n-1} (\frac{9}{8})^{n-2} (\frac{4}{3})^{n-3} \det(H_n^4(F_0)) \\
&= -\frac{1}{4}\det(H_n^4(F_0)) \qquad (4.3)
\end{aligned}$$

Moreover,

$$F_9 = \frac{20 + 16x - 8x^2 - 4x^3 + x^4}{8(2 - 4x^2 - 2x^3 + x^4) - 16x^4 F_9} = \frac{5}{4} + x + 2x^2 + 3x^3 + 6x^4 + 10x^5 + \cdots$$

It is easily verified that $F_9(x)$ and $\frac{1}{4} + G_0(x)$, where G_0 appears in Subexample 5ii, satisfy the same functional equation, and hence are equal. Therefore,

$$\begin{aligned}
\det(H_{n-6}(F_9)) &= \left[\frac{xF_9(x) - yF_9(y)}{x - y}\right]_{n-6} \\
&= \left[\frac{1}{4} + \frac{xG_0(x) - yG_0(y)}{x - y}\right]_{n-6} \\
&= \frac{1}{4}\det(H_{n-7}^4(F_0)) + \det(H_{n-6}(G_0))
\end{aligned}$$

where $\frac{1}{4}\det(H_{n-7}^4(F_0))$ is $\frac{1}{4}$ times the determinant of the 1,1-minor of $H_{n-6}(G_0)$, equivalently of $H_{n-6}^2(F_0)$. Combining this with identity (4.3) and noting $\det(H_n(G_0)) = -\det(H_{n+5}(F_0))$ from (4.2) proves the initial recurrence of this subexample.

Acknowledgments. We, particularly Xin, are most grateful to Ira Gessel. We also appreciate the suggestions made by Christian Krattenthaler, Lou Shapiro, and referees.

References

[1] M. Aigner, Motzkin Numbers, *Europ. J. Comb.* 19 (1998), 663-675.
[2] C. Banderier and S. Schwer, Why Delannoy numbers?, *J. Stat. Plan. and Infer.*, 135 (2005), 40-54.
[3] D. Bressoud, *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture*, Cambridge University Press, 1999.

[4] R. A. Brualdi and S. Kirkland, Aztec diamonds and digraphs, and Hankel determinants of Schröder numbers, *J. Comb. Th., ser. B*, 94 (2005), 334-351.
[5] I. M. Gessel and G. X. Viennot, Binomial determinants, paths, and hook length formulae. *Adv. in Math.* 58 (1985), no. 3, 300-321.
[6] I. M. Gessel and G. Xin, *The generating function of ternary trees and continued fractions*, Electron. J. Combin., to appear, ArXiv:math.CO/0505217.
[7] W. B. Jones and W. J. Thron, *Continued Fractions: Analytic Theory and Applications*, Encyclopedia of Mathematics and its Applications, vol. 11, Addison-Wesley Publishing Co., Reading, Mass., 1980.
[8] C. Krattenthaler, Advanced determinant calculus: a complement, *Linear Algebra Appl.* 411 (2005), 68-166
The Andrews Festschrift (Maratea, 1998). *Sem. Lothar. Combin.* 42 (1999), Art. B42q, 67 pp. (electronic).
[9] E. Pergola, R. Pinzani, S. Rinaldi, and R. A. Sulanke, Lattice Path Moments by Cut and Paste. *Adv. in Appl. Math.* 30 (2003), no. 1-2, 208-218.
[10] Shapiro, L. W. A Catalan triangle. *Discrete Math.* 14 no. 1 (1976), 83–90.
[11] R. A. Sulanke, Moments of generalized Motzkin paths, *J. Int. Seq.*, 3 (2000), Article 00.1.1
[12] R. A. Sulanke, Objects counted by the central Delannoy numbers. *J. Int. Seq.* 6 (2003), no. 1, Art. 03.1.5, 19 pp.
[13] G. Viennot, Une théorie combinatoire des polynômes orthogonaux généraux, Lecture notes, Univ. Quebec, Montreal, Que., 1983.
[14] H. S. Wall, *Analytic Theory of Continued Fractions*, Van Nostrand, New York, 1948.

Department of Mathematics, Boise State University, Boise, ID, USA 83725
E-mail address: `sulanke@math.boisestate.edu`

Department of Mathematics, University of Kentucky, Lexington, KY, USA 40506
E-mail address: `gxin@ms.uky.edu`

Part IV
Software
Logiciels

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Mu-PAD Combinat

Nicolas M. Thiéry

MuPAD-Combinat is an open-source algebraic combinatorics package for the computer algebra system MuPAD. Its main purpose is to provide an extensible toolbox for computer exploration, and to foster code sharing between researchers in this area.

http://mupad-combinat.sf.net http://www.mupad.de

The development of MuPAD-Combinat started in 2001, and it is part of the standard distribution of MuPAD since 2003. It played a central role in 15+ publications with more than a dozen contributors. You are more than welcome to join the team!

During FPSAC 2006, several of the contributors will demonstrate, as part of their research talks, their own use of MuPAD-Combinat (F. Hivert, N. Thiry, J.C. Novelli, F. Descouens, J. Nzeutchap, M. Rey, ...). We will also present a more formal software demonstration with an overview presentation and discussions about the design and development model. It will be followed by an on-computer tutorial for those who would like to get a first hand experience.

The software will be included in the proceedings CD-Rom along with its full online documentation (600 pages), which includes a step by step introductory tutorial and design notes. It requires the MuPAD computer algebra system, a demonstration version of which we will try to include as well.

MuPAD-Combinat est une bibliothèque de combinatoire algébrique pour le système de calcul formel MuPAD, développée sous licence libre (LGPL). Son objectif est de fournir un cadre de développement qui permette aux chercheurs du domaine de mettre en commun les routines qu'ils sont amenés à développer dans le cadre de leur recherche.

http://mupad-combinat.sf.net http://www.mupad.de

Le développement de MuPAD-Combinat a commencé en 2001, et est partie intégrante de la distribution standard de MuPAD depuis 2003. Il a joué un rôle central dans plus de quinze publications, avec plus d'une douzaine de contributeurs. Vous-êtes les bienvenus dans l'équipe!

Durant FPSAC 2006, certains des contributeurs présenterons succinctement, à l'occasion de leur exposé de recherche, leur propre utilisation dans la pratique de MuPAD-Combinat (F. Hivert, N. Thiéry, J.C. Novelli, F. Descouens, J. Nzeutchap, M. Rey, ...). Nous proposerons aussi une démonstration logicielle plus formelle, où nous décrirons la structure générale de cette bibliothèque, tout en expliquant sa conception et son modèle de développement. Cette démonstration sera suivi par un tutoriel sur ordinateur, pour ceux qui souhaiteront se faire leur propre idée.

Le logiciel sera inclus dans le CD-Rom des comptes-rendus, avec sa documentation complète (600 pages) qui inclue entre autres un tutoriel pas-à-pas, ainsi que des notes sur la conception. Le système de calcul formel MuPAD est requis; nous essaierons d'en fournir une version de démonstration.

Laboratoire de Mathématiques, Université Paris Sud, B<E2>t 425, 91405 Orsay Cedex, FRANCE
E-mail address: `nthiery@users.sf.net`

Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

A user manual for CrystalView

Philip Sternberg

Abstract. A description of CrystalView, a software package for visualizing crystal bases and carrying out calculations on them.

Résumé. Une description de CrystalView, un progiciel pour visualizer les bases cristallines et mise en oeuvre des calculs sur eux.

1. Overview

CrystalView is a software package for visualizing crystals for irreducible highest weight modules of simple Lie algebras. Based on user input, the program will produce an image file with the requested crystal graph. The program will automatically produce an .epsf file, which can be included directly in a LaTeX file. Using the web interface, the user can see the tableau associated a vertex of the crystal by moving the mouse pointer over it.

When running this program locally (as opposed to via the web), the program produces a list of all tableaux in the specified crystal. This can be used to carry out calculations on crystals, including searches for tableaux with specified properties. Additionaly, Kashiwara operators may be applied to the tableaux.

All calculations on tableaux are carried out using python. Image files are automatically generated PostScript. The web interface uses html, javascript, and css.

2. Requirements

CrystalView may be run using a web interface or from source code. The web interface can, in principle, be used from any browser that supports form input (any "modern" browser). However, some aspects of the interface will only work with a browser that complies with standard html, javascript, and css. In particular, Internet Explorer is known to have issues with some dynamic aspects of the web interface. Firefox is a recommended alternative.

To run the software from source code, the user must have Python 2.4 installed on their local machine. Other versions of Python (both older and newer) may not run CrystalView properly. See `http://www.python.org` for further information regarding python. Python is available free of charge for all major operating systems, and is included pre-installed on many modern computers, including almost all distributions of Unix/Linux.

If used to generate image files locally, the user is advised that for large crystals, these images can get quite large. See section 7.

3. User input

The user may specify the following Lie theoretic data:

- symmetry type of the algebra being represented;
- rank of the algebra being represented;

2000 *Mathematics Subject Classification.* Primary 17B37; Secondary 05C75.
Key words and phrases. crystal bases, Lie algebras, software.
Supported in part by NSF grants DMS-0135345, DMS-0200774, and DMS-0501101.

- highest weight of the representation.

Additionally, the user may specify how the edges of the crystal graph will be drawn. There are two default settings, color and grayscale, as well as an option for custom settings. If the custom option is selected, the user may specify the following:

- red/green/blue values for each edge color on a scale from 0 to 1 (Default = 1),
- line width on a scale from 1 to 5 (Default = 5),
- dash pattern; none, short, long (Default = none).

There are numerous resources on the web and preinstalled on many computers to assist the user in finding red/green/blue values for their desired colors.

The user may also choose to have the output converted to .pdf, .jpg, .gif, and/or .tiff formats.

4. Limitations

Currently, only types A and B (i.e., $\mathfrak{sl}_n$ and $\mathfrak{so}_{2n+1}$) are supported. Furthermore, in the case of type B, only even multiples of the highest fundamental weight (corresponding to the short root) may be specified. These limitations are due to the current stage of the development cycle; future versions of the software will add support for types C and D and all dominant weights.

The rank of the algebra is currently restricted by the web interface to be no larger than 5. This is an artificial limitation; any rank of algebra may be specified when running CrystalView from source.

The web interface only allows crystals with as many as 4,000 vertices to be calculated to prevent excessive strain on the server. Considering the resolution at which these images can be viewed/printed, it is unlikely that producing crystals larger than this would be useful to most users. However, this limitation is artificial; when running the program on a local machine, the user is limited only by their own patience and hard disk space.

5. How it works

The web interface for CrystalView is written in dhtml; i.e., html enhanced by javascript and css. The Weyl dimension formula is used to calculate the number of vertices in the currently specified crystal.

The tableaux are produced by generating the list of all column tableaux for the column lengths appearing in the shape specified by the dominant weight. To determine what constitutes a legal tableaux, the criteria of [**1**] are used. The columns are then compared pairwise in order of decreasing length to build the set of all legal tableaux. In the case of type B crystals, the "split form" criterion of [**2**] is used to determine which columns can be adjacent in a legal tableau.

The graph is ensured to have a reasonable number of edge crossings by ordering the vertices of the graph as follows, starting from the top and going down the rows, and proceeding through each row from left to right. First, the tableaux are collected into rows according to content. There is only one tableau in the first row (the highest weight tableau), so the first row is in order. Now, given that row n is in order, row $n+1$ is put in the following order. The leftmost vertices in row $n+1$ will be the non-zero images of the Kashiwara operators f_i on the leftmost vertex in row n, taken in order from f_1 up through f_r, where r is the rank of the algebra. The next leftmost vertices in row $n+1$ are those tableaux that result from applying f_i to the next vertex from row n, excluding those that have been placed to the left already. This process continues until all vertices have been placed in their final position.

6. How to get the software

This software can be accessed at the following url:

`http://www.math.ucdavis.edu/ sternberg/crystalview/`

The source code is available upon request from the author.

7. Legal disclaimer

This software is provided without warranty; use is at the user's sole risk. Under no circumstances will any user hold the author of this software liable for any damages that may result from the use of this software. Use of the software implies that the user agrees to these terms.

References

[1] J. Hong and S.-J. Kang, *Introduction to quantum groups and crystal bases*, Graduate Studies in Mathematics, 42. American Mathematical Society, Providence, RI, 2002. xviii+307 pp. ISBN: 0-8218-2874-6.

[2] C. Lecouvey, *Schensted-type correspondences and plactic monoid for types B_n and D_n*, J. Algebraic Combin. **18** (2003), no. 2, 99–133.

Department of Mathematics, University of California, One Shields Avenue, Davis, CA 95616-8633, U.S.A.
E-mail address: `sternberg@math.ucdavis.edu`
URL: `http://www.math.ucdavis.edu/~sternberg`